Oxford Resources for IB
Diploma Programme

2025 EDITION

PSYCHOLOGY

COURSE COMPANION



Alexey Popov

OXFORD

Oxford Resources for IB
Diploma Programme

2025 EDITION

PSYCHOLOGY

COURSE COMPANION

Alexey Popov

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP, United Kingdom

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries.

© Oxford University Press 2025

The moral rights of the author have been asserted.

First published in 2025

Index: James Helling

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, transmitted, used for text and data mining, or used for training artificial intelligence, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by licence or under terms agreed with the appropriate reprographics rights organization. Enquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above.

You must not circulate this work in any other form and you must impose this same condition on any acquirer.

British Library Cataloguing in Publication Data

Data available

9781382056663

9781382056670 (ebook)

10 9 8 7 6 5 4 3 2 1

The manufacturing process conforms to the environmental regulations of the country of origin.

Printed in China by Shanghai Offset Ltd.

Acknowledgements

The "In cooperation with IB" logo signifies the content in this textbook has been reviewed by the IB to ensure it fully aligns with current IB curriculum and offers high quality guidance and support for IB teaching and learning.

The Publisher wishes to thank the International Baccalaureate Organization for permission to reproduce their intellectual property.

The publisher and authors would like to thank the following for permission to use photographs and other copyright material:

Cover: JethuynhCan / Getty Images. **Photos: pX:** EreborMountain / Shutterstock; **p2:** golero / E+ / Getty Images; **p6:** Chronicle / Alamy Stock Photo; **p52:** jannoon028 / Shutterstock; **p57:** Keystone Press / Alamy Stock Photo; **p59:** Edwin Butter / Shutterstock; **p68:** Dream01 / Shutterstock; **p76(t):** VectorMine / Shutterstock; **p76(b):** anananas / Shutterstock; **p77(b):** zizou7 / Shutterstock; **p82:** Piscine26 / Shutterstock; **p98:** Alex Potemkin / E+ / Getty Images; **p102:** Zuri Swimmer / Alamy Stock Photo; **p106(t, b):** Science History Images / Alamy Stock Photo; **p119:** Krakenimages.com / Shutterstock; **p135:** dpa picture alliance archive / Alamy Stock Photo; **p138:** Altaf Shah / 500px / Getty Images; **p140:** Vitalii Vodolazskyi / Shutterstock; **p146:** Chris Dent / Oxford University Press; **p153:** Chris WM Willemsen / Shutterstock; **p157:** Jacques Jangoux / Alamy Stock Photo; **p162:** Daniel de la Hoz / Getty images; **p164:** BLACKDAY / Shutterstock; **p170:** Real Valdés, Santiago & Araujo, Alvaro. (2019). "Navigation Systems for the Blind and Visually Impaired: Past Work, Challenges, and Open Problems." *Sensors. 19*. 3404. 10.3390/s19153404.; **p171:** mark reinstein / Shutterstock; **p173:** Willett et al. (2021). "High-performance brain-to-text communication via handwriting." *Nature, 593(7858)*, 249–254. https://doi.org/10.1038/s41586-021-03506-2; **p178:** Flashpop / Stone / Getty Images; **p181:** Reprinted from "Development of the child's brain and behavior," Bryan Kolb, Bryan D. Fantie, Springer Nature, 2009, with permission of Springer Nature; **p184:** Basu Kshitiz / Shutterstock; **p187:** Svetlana Lazarenka / Alamy Stock Photo; **p191(t):** 2p2play / Shutterstock; **p191(b):** ARIMAG / Shutterstock; **p192:** Gado Images / Alamy Stock Photo; **p199:** Album / Alamy Stock Photo; **p201:** GAS-photo / Shutterstock; **p203:** Tatiana Buzmakova / Shutterstock; **p211:** Chester Voyage / Alamy Stock Photo; **p218:** Smit / Shutterstock; **p224:** Ariadne Van Zandbergen / Alamy Stock Photo; **p230:** Premack D, Woodruff G. "Chimpanzee problem-solving: a test for comprehension." *Science*. 1978 Nov 3;202(4367):532–5. doi: 10.1126/science.705342. PMID: 705342.; **p231:** Reprinted from *Trends in Cognitive Sciences*, Vol number 12, Josep Call, Michael Tomasello, "Does the chimpanzee have a theory of mind? 30 years later," 187–192, 2008, with permission from Elsevier; **p234:** Reprinted from *NeuroImage*, Vol number 19, R Saxe, N Kanwisher, "People thinking about thinking people: The role of the temporo-parietal junction in 'theory of mind'," 1835–1842, 2003, with permission from Elsevier; **p238, 239:** "The nature of love." Harlow, Harry F., *American Psychologist*, Vol 13(12), Dec 1958, 673–685/Public Domain; **p242:** Nick Greaves; Agfa Awards Winner / Alamy Stock Photo; **p246:** jose carlos cerdeno martinez / Moment / Getty Images; **p264:** grandbrothers / Shutterstock; **p267:** Sipa US / Alamy Stock Photo; **p308:** Designua / Shutterstock; **p317:** Prostock-studio / Shutterstock; **p324:** sogane / Shutterstock; **p332:** aaaaimages / Moment / Getty Images; **p338:** NFNArts / Shutterstock; **p355:** PRASANNAPIX / Shutterstock; **p357:** a katz / Shutterstock; **p358:** Attila JANDI / Shutterstock; **p364(t):** Adisha Pramod / Alamy Stock Photo; **p364(b):** Steve Lillie / Alamy Stock Photo; **p366:** Justas in the wilderness / Shutterstock; **p367:** Salomons, N., Sebo, S. S., Qin, M., & Scassellati, B. (2021). "A minority of one against a majority of robots." *ACM Transactions on Human–Robot Interaction, 10(2)*, 1–22. https://doi.org/10.1145/3442627; **p389:** Japan's Fireworks / Shutterstock; **p390:** *Journal of comparative neurology* / John Wiley & Sons; **p392(l):** Riju Sathyan / Alamy Stock Photo; **p392(r):** Maylandia / Shutterstock; **p396:** eggeegg / Shutterstock; **p410:** dpa picture alliance / Alamy Stock Photo; **p414:** Oscar Wong / Moment / Getty Images; **p444:** Rijksmuseum; **p453:** Prathankarnpap / Shutterstock; **p455:** Used with permission of University of Chicago Press – Journal, from Jennifer Shepard Payne, (2012). "Influence of Race and Symptom Expression on Clinicians' Depressive Disorder Identification in African American Men," *Journal of the Society for Social Work and Research 3(3)*, 162–177. © 2012 by the Society for Social Work and Research. All rights reserved. Permission conveyed through Copyright Clearance Center, Inc.; **p459:** Rubens Alarcon / Alamy Stock Photo; **p489:** sutadimages / Shutterstock; **p492:** We Are / DigitalVision / Getty Images; **p504:** Fuse / Corbis / Getty Images.
Artwork by Q2A Media and Thomson.

Text: Extract from Vaterlaus, J.M., Tulane, S., Porter, B.D., & Beckert, T.E. (2018). "The Perceived Influence of Media and Technology on Adolescent Romantic Relationships." *Journal of Adolescent Research, 33(6)*, 651–671. Permission conveyed through Copyright Clearance Center, Inc.
Extract from Heider, F. & Simmel, M. (1944). "An experimental study of apparent behavior." *The American Journal of Psychology, 57(2)*, 243–259. Permission conveyed through Copyright Clearance Center, Inc.
Extract from Pike, K.L. (1967). *Language in relation to a unified theory of the structure of human behaviour*. The Hague: Mouton, p. 41. Permission conveyed through Copyright Clearance Center, Inc.
Extract from Berry, J.W. (1989). "Imposed etics—emics-derived etics: The operationalization of a compelling idea." *International Journal of Psychology, 24,* 726. Permission conveyed through Copyright Clearance Center, Inc.
Extract from Bach-y-Rita, P., Collins, C.C., Saunders, F.A., White, B., & Scadden, L. (1969). "Vision substitution by tactile image projection." *Nature, 221,* 963–964. Permission conveyed through Copyright Clearance Center, Inc.
Extract from Kolb, B., & Fantie, B. (2009). "Development of the child's brain and behavior." In C.R. Reynolds, & E. Fletcher-Janzen (Eds.), *Handbook of Clinical Child Neuropsychology. Third Edition* (pp. 19–46). Springer Science+Business Media. Permission conveyed through Copyright Clearance Center, Inc.
Extract from Doidge, N. (2007). *The brain that changes itself: Stories of personal triumph from frontiers of brain science*. New York, USA: Penguin Books, pp. 63–64.
Extract from Blakemore, C., & Cooper, G. (1970). "Development of the Brain depends on the Visual Environment." *Nature, 228,* 477. Permission conveyed through Copyright Clearance Center, Inc.
Extract from World Health Organization. (17 June 2022). "Mental Health."
Extract from Zhao, N., & Zhou, G. (2021). "COVID-19 stress and addictive social media use (SMU): Mediating role of active use and social media flow." *Frontiers in Psychiatry, 12,* 85. Permission conveyed through Copyright Clearance Center, Inc.
Extract from Hunt, M.G., Marx, R., Lipson, C., & Young, J. (2018). "No more FOMO: Limiting social media decreases loneliness and depression." *Journal of Social and Clinical Psychology, 37(10),* 751–768. Used with permission of Guilford Publications, Inc.; permission conveyed through Copyright Clearance Center, Inc.
Extract from Morey, J.N., Gentzler, A.L., Creasy, B., Oberhauser, A.M., & Westerman, D. (2013). "Young adults' use of communication technology within their romantic relationships and associations with attachment style." *Computers in Human Behavior, 29(4),* 1771–1778. Used with permission of Elsevier Ltd; permission conveyed through Copyright Clearance Center, Inc.
Extract from Brevers, D., & Turel, O. (2019). "Strategies for self-controlling social media use: Classification and role in preventing social media addiction symptoms." *Journal of Behavioral Addictions, 8(3),* 554–563. Used with permission of Akademiai Kiado; permission conveyed through Copyright Clearance Center, Inc.
Extract from Fischer, K.W., & Silvern, L. (1985). "Stages and individual differences in cognitive development." *Annual Review of Psychology, 36,* 614. Used with permission of Annual Reviews, Inc.; permission conveyed through Copyright Clearance Center, Inc.
Extract from Bandura, A. (1989). "Human agency in social cognitive theory." *American Psychologist, 44,* 1175. Used with permission of American Psychological Association; permission conveyed through Copyright Clearance Center, Inc.
Extract from Bower, G.H., Black, J.B., & Turner, T.J. (1979). "Scripts in memory for text." *Cognitive Psychology, 11,* 190. Used with permission of Elsevier Inc.; permission conveyed through Copyright Clearance Center, Inc.

Every effort has been made to contact copyright holders of material reproduced in this book. Any omissions will be rectified in subsequent printings if notice is given to the publisher.
Links to third party websites are provided by Oxford in good faith and for information only. Oxford disclaims any responsibility for the materials contained in any third party website referenced in this work.

Contents

Introduction ..... **vi**
Key concepts in IB Psychology ..... **x**
1 Research methods and data interpretation **3**
1.1 What is psychology? 5
1.2 Research methods in psychology (overview) 9
1.3 Analysing research (methodological considerations) 13
1.4 The experiment 16
1.5 Correlational studies 31
1.6 Qualitative research 36
1.7 Surveys and questionnaires 46
1.8 Ethics in psychological research 54
1.9 Research with animals 58
1.10 Descriptive statistics 64
1.11 Plotting and graphing 71
1.12 Inferential statistics 80
1.13 Analysing qualitative data 94
2 Learning and cognition **99**
2.1 Classical and operant conditioning 101
2.2 Schema theory 110
2.3 Biological factors in cognitive processes 117
2.4 Cognitive models 124
2.5 Cognitive biases and the dual processing model of thinking and decision-making 134
2.6 Social learning theory 143
2.7 Cultural factors in cognitive processes 152
2.8 Environmental influences on cognitive processes 161
2.9 Potential for improving cognitive processes 168
3 Human development **179**
3.1 Brain development 181
3.2 Stage theories and continuous models of human development 190
3.3 Sociocultural factors in development 198
3.4 Enculturation of social norms 206
3.5 Peer influence 214

3.6 Role of childhood experiences..... 221
3.7 Theory of mind..... 229
3.8 Attachment..... 237

4 Health and well-being..... 247
4.1 Biological explanations of mental disorders..... 249
4.2 Biological treatment..... 258
4.3 Cognitive models of mental disorders..... 266
4.4 Psychological treatment..... 275
4.5 Cultural differences in mental disorders..... 283
4.6 Environmental factors in mental disorders..... 291
4.7 Prevalence of health problems..... 299
4.8 Stress and health..... 307
4.9 Social learning and health problems..... 315
4.10 Prevention of one health problem..... 323

5 Human relationships..... 333
5.1 Social identity theory..... 335
5.2 Cultural dimensions..... 343
5.3 Acculturation..... 352
5.4 Conformity..... 362
5.5 Compliance techniques..... 370
5.6 Social learning in group behaviour..... 378
5.7 The role of chemical messengers in interpersonal relationships..... 386
5.8 Cognitive explanations for interpersonal relationships..... 395
5.9 The role of communication and language, and strategies for improving relationships..... 403

6 HL extensions..... 415
6.1 Motivation..... 417
6.2 Motivation in human relationships..... 421
6.3 Motivation and mental health..... 426
6.4 Motivation and human development..... 432
6.5 Motivation in learning and cognition..... 438
6.6 Culture..... 444

6.7 Culture in human relationships..... 447
6.8 Culture and mental health..... 453
6.9 Culture and human development..... 458
6.10 Culture in learning and cognition..... 463
6.11 Technology..... 468
6.12 Technology in human relationships..... 471
6.13 Effects of technology on health and well-being..... 476
6.14 Effects of digital technology on child development..... 481
6.15 Effects of technology on learning and cognition..... 487

7 Internal assessment..... 493
7.1 Overview of the task..... 493
7.2 Criterion A: Introduction..... 494
7.3 Criterion B: Research methodology..... 496
7.4 Criterion C: Data collection..... 498
7.5 Criterion D: Discussion..... 501
7.6 Ethical considerations in the research proposal..... 502
7.7 The presentation of the research proposal..... 503

8 External assessment..... 505
8.1 Structure of exam papers..... 505
8.2 General exam tips..... 510
8.3 Sample paper..... 516

Bibliography..... 523

Answers: www.oxfordsecondary.com/ib-psych-support

Introduction

Overview of the course

The IB Psychology course is structured around three key elements: content, contexts, and concepts.

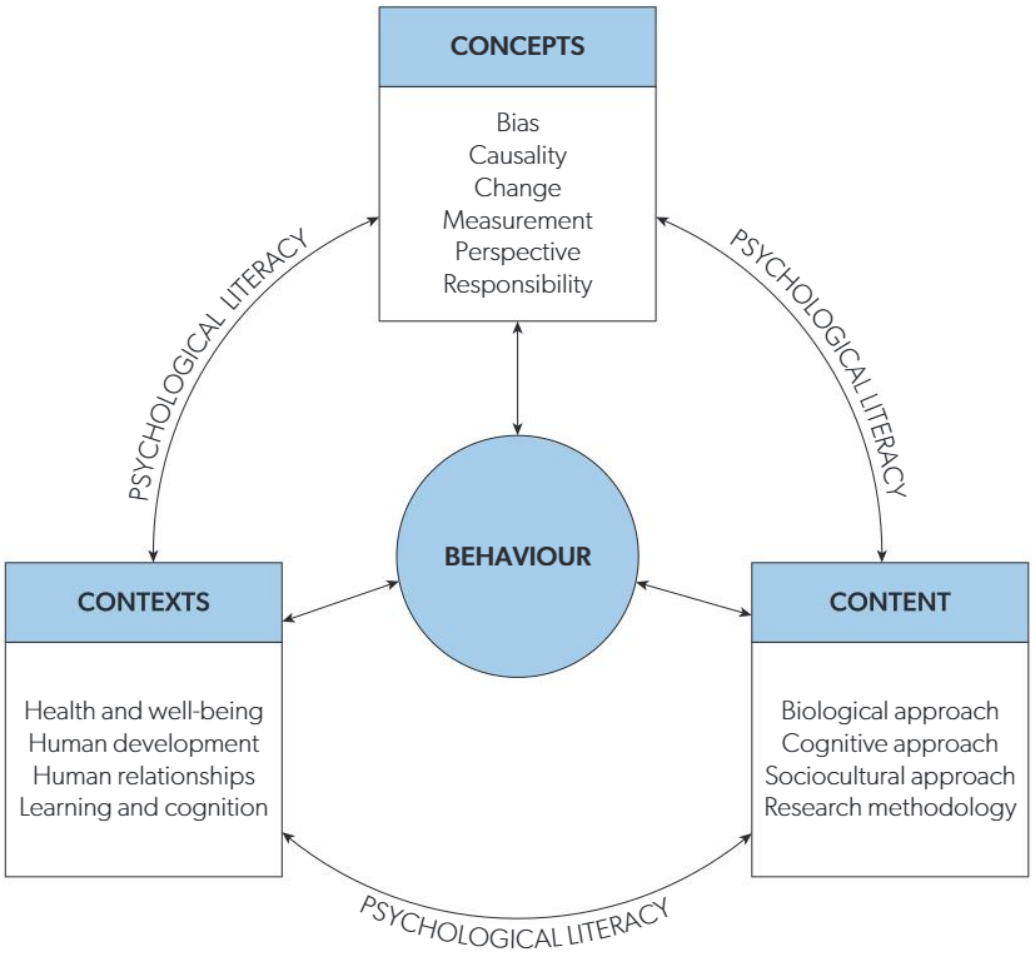
Content refers to specific psychological ideas, theories, and research studies. Social identity theory, operant conditioning, and models of memory are all examples of content. Content can be further organized into the following four groups:

- Biological approach to behaviour
- Cognitive approach to behaviour
- Sociocultural approach to behaviour
- Research methodology

The first three represent the triad (biological–cognitive–sociocultural) that you will come across frequently. It refers to the three main perspectives on behaviour. It is widely recognized that the most holistic understanding of behaviour is achieved when perspectives are combined, and various factors are considered in interaction with each other. However, the long history associated with each of these approaches—as well as the distinctly different ways of conducting research in interpreting results within each of them—makes it necessary to keep this distinction.

The fourth group—research methodology—is added in recognition of the fact that research is how we gain knowledge about behaviour, which makes an understanding of how research works fundamental to understanding psychology.

We can compare the biological, cognitive, and sociocultural approaches to "lenses" through which we interpret the results of research studies in an attempt to explain behaviour.



▲ **Figure 1** Concepts, contexts, and content in IB Psychology

Exam tip

Note: the overview provided here is just an introduction into how material in IB Psychology is structured. There are some further nuances when it comes to exams. For example, there is a distinction between content required for:

- Paper 1 Sections A and B
- Paper 1 Section C.

See Chapter 8: External assessment for a full explanation of how exams are structured and what the requirements are.

Content is embedded within **contexts**. Contexts are specific areas of research that have both theoretical and practical significance. The four contexts in IB Psychology are:

- Learning and cognition
- Health and well-being
- Human development
- Human relationships

There is always an integrated analysis of behaviour within each of the contexts. For example, mental disorders (health and well-being) are considered from the perspective of the three approaches (biological, cognitive, sociocultural), both separately and in combination. All three groups of factors contribute to the development of mental disorders, and one needs to understand their complex interplay in order to comprehend the nature of these phenomena. Of course, it is only possible through a careful consideration of research methodology and its strengths and limitations in each particular study and research area.

Finally, the **concepts** are the big overarching ideas that manifest in a variety of ways in the specific content and contexts. In the following section—Key concepts in IB Psychology— we will provide some initial explanation of the six key concepts in IB Psychology.

Apart from these elements, there are also components of the course that are only required for higher level students (**HL extensions**). They are:

- Culture (the role of culture in shaping behaviour)
- Motivation (the role of motivation in shaping behaviour)
- Technology (the role of technology in shaping behaviour)
- Data analysis and interpretation.

Data analysis and interpretation has a somewhat separate status because it is recommended that every student is familiar with it, but it is only HL students who will be directly tested on it.

Overview of the book

Chapter 1

This book begins with a comprehensive chapter on research methodology and data interpretation. It lays the foundation for all further chapters and supports you in being able to understand and critically evaluate research in psychology. It is expected that you will frequently revisit this chapter and apply the skills that you are gaining from it to new content as you keep studying through the course. Research methodology in this chapter is separated from data interpretation to make clear which part of the chapter is for HL students, although of course these two sections reinforce each other. Chapter 1 is also aligned with the four class practicals that you will be required to carry out at various points throughout your study.

Chapters 2–5

The book is based on the four contexts (Chapters 2–5). Each context is further broken down into units which are aligned with the content required for exams (see Chapter 8: External assessment). Each unit includes the following repeated elements:

- Inquiry questions. These provide some starting points for the inquiry process.
- "What you will learn in this section": Key learning and Key terms. This feature provides a concise overview of the main take-away messages from each unit.
    - The statements under "Key learning" represent the main conceptual understandings. The idea is that, if you can look at these statements and explain each of them, then you have deeply understood the material in the unit.
    - Key terms are just that—the most important terms that have been used in the unit. Ensure you understand the meaning of all of them.
- "In a wider context". This is a short section that each unit starts with—to place the material in the context of what you already know from other parts of the course. Each unit represents one relatively independent, but small, component in a much larger picture.
- Various other sections of the unit that are specific for each topic.
- Various in-text features, such as Activity, Discussion, Chat with AI, Exam tip, TOK (explained on page ix).
- Conceptual analysis. This section contains some initial suggestions on how the content in the unit could be related to each of the six concepts (Bias, Perspective, Measurement, Causality, Change, Responsibility). This is not an exhaustive analysis and not a prescribed way to link content to

concepts—it would be impossible to map out all possible connections of all concepts to all content. It would also go against the philosophy of the course: you are supposed to acquire some key transferable skills that will enable you to answer questions about any possible combination of concepts and content, rather than memorizing these possible combinations. However, the suggestions may give you an idea about what kinds of links could be used, to further your own thinking in each case.

Concepts, content, and contexts are all interlinked in these four chapters. You study content through the lens of concepts, and your understanding of concepts gradually emerges through the study of multiple content points in which they are manifested. This is all against the backdrop of a specific area of theoretical and practical significance—a context. However, in all this closely intertwined, integrated construction it is important not to lose clarity when it comes to exam requirements.

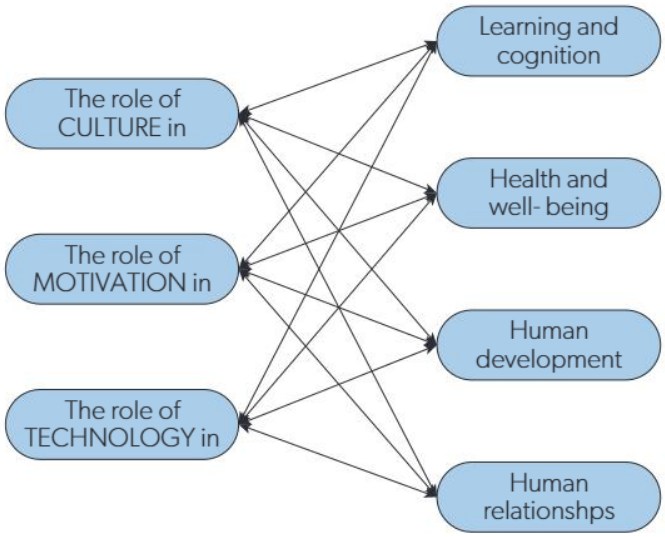
Some content—most of it—is optional, in the sense that it is not a requirement to remember it and your knowledge of factual details (i.e., details of research studies) will not be directly assessed. However, some content is considered the basic requirement and is tested directly in such exam components as Paper 1 Section A and Paper 1 Section B. See further details in Chapter 8: External assessment.

Although this basic content is embedded in the study of contexts, we have clearly separated it into special SAQ (short-answer questions) features. Table 1 provides a full list of these basic content points required for Papers 1A and 1B and where they can be found in the book.

Biological approach	Page	Cognitive approach	Page	Sociocultural approach	Page
Brain imaging techniques	121	Classical conditioning	103	Models of acculturation	353
Chemical messengers	387	Operant conditioning	105	Cognitive dissonance	148
Neurotransmission	253	Schema theory	112	Compliance techniques	371
Neuroplasticity	183	Cognitive models	127	Conformity	363
Localization of function	120	Dual-processing theory	135	Cultural dimensions	344
Genetic inheritance	250	Cognitive load theory	131	Enculturation	207
Diathesis–stress model	256	Anchoring bias	139	Social identity theory	336
Animal research/animal models	58	Confirmation bias	137	Social learning theory	144
Biological reductionism	118	–		Emic approach	154
–		–		Etic approach	154

▲ **Table 1** List of content for Paper 1A and 1B. See further details in Chapter 8: External assessment.

Chapter 6



▲ **Figure 2** HL extensions

Chapter 6 is focused on HL extensions. It is further broken down into units, one for each combination of the HL extension (culture, motivation, technology) and the context (learning and cognition, health and well-being, human relationships, human development). For example, one of the sections is the role of technology in human development. HL extensions do not represent additional content; they should be viewed as additional "lenses" to enhance your study of the contexts. As you will learn from exam requirements, you are not directly tested on your knowledge of material from these HL extensions, rather it is the skills of data interpretation and evaluating research conclusions that are tested.

For this reason, Chapter 6 is organized differently. It models the kind of thinking that will be required to successfully answer questions in Paper 3 by presenting several research sources and critically analysing their conclusions. You could consider this chapter as an opportunity to practise your skills with a variety of specific research examples.

Chapter 7

Chapter 7 is focused on the Internal Assessment. In this task you need to prepare a research proposal to investigate an issue of practical significance that can potentially impact the lives of people in positive ways. The chapter analyses all relevant requirements in detail and gives you examples of how to approach the task.

Chapter 8

Chapter 8 is an outline of external assessments. It presents the details of all external assessment components (Papers 1–3). There is some crucial information on what exactly you are expected to know and to be able to do, and to what level of depth. Study this chapter carefully and keep coming back to it as you learn new material.

At the end of the book, you will find a bibliography giving references to all sources that were used in the book.

About in-text features

Throughout this book, you will find a number of in-text features that will help you deepen and broaden your understanding of the material. Here is an overview of the common in-text features with an explanation of the reasoning behind each of them.

Activity

This in-text feature suggests activities that you can do either by yourself or with your classmates during lessons. They include small group projects, debates, conducting research online, brainstorming sessions, sharing your thoughts with each other, and many more. They are tailored to the five key ATL skills that are fundamental to any IB course.

Discussion

This feature suggests points of discussion with a partner or a group. When you discuss an issue with someone, you can compare various perspectives and points of view to enable a richer consideration of the problem, while considering its various aspects and sides.

Chat with AI

Generative AI has become an integral part of our lives. Throughout this book, the feature "Chat with AI" teaches you the principles and tricks of using such systems in a way that will deepen your knowledge and make your learning more effective. You will learn to recognize when to trust the output of such systems and when to be sceptical and double-check.

TOK

This in-text feature links the material you are studying to some relevant ideas or concepts from Theory of Knowledge (TOK).

Where relevant, the feature also introduces elements of TOK itself, especially when our discussion of psychology comes in close proximity to more generic concepts of knowledge, such as justification, falsification, or paradigm shift.

SAQ

The content required for Paper 1 Sections A and B is presented in this book as embedded in a wider discussion of the four contexts. However, we have also clearly separated and highlighted this content throughout the text to make it easier for you to see what exactly is required and to what depth. This in-text feature serves the purpose of highlighting the content that you are required to know for the exams.

There is one such in-text feature for each of the SAQ content points on the list. However, they are scattered around the book and are used when they are most relevant in the context. There is a full index of these in-text features in Table 1, opposite.

Exam tip

Sometimes there are tricks and nuances that are good to know for exams. This feature appears as and when necessary, especially in those sections of the course where you are more likely to make mistakes or decisions that are not optimal.

Key concepts in IB Psychology

Modern theories of education often bring up the idea of "conceptual learning" or "concept-based teaching". You might have heard these words from your subject teachers as these ideas are becoming more and more deeply integrated into IB subjects. So what does it mean to learn "conceptually"?

What is a concept?

A concept is a general idea that can be applied to a large number of specific instances.

Think about various types of furniture: chair, sofa, bed, desk. You might say that all these are specific objects. But there is something that connects them all—the concept of "furniture".

Imagine knowing what a chair is, what a sofa is, what a bed is—but not knowing what furniture is. That would be a weird, inefficient form of knowledge, limited to being able to recognize specific examples but not understanding the nature of things.

On the contrary, suppose you know what furniture is but don't know what a sofa is. You already know the general idea. Once you learn that a sofa is a piece of furniture, you already know so much about it because all the properties of the concept (furniture) apply to this particular instance. All that is left for you to understand is what makes a sofa different from other types of furniture that you are already familiar with.

What is a hierarchy of concepts?

In each discipline there is always a hierarchy of concepts. Closer to the bottom are very specific terms that refer to a narrow set or particular examples of objects or phenomena. Closer to the top are broad concepts that cover a large number of specific terms.

DOMAIN	Eukarya
KINGDOM	Animalia
PHYLUM	Chordata
CLASS	Mammalia
ORDER	Carnivora
FAMILY	Canidae
GENUS	*Canis*
SPECIES	*Canis lupus*



▲ **Figure 3** Hierarchy of concepts

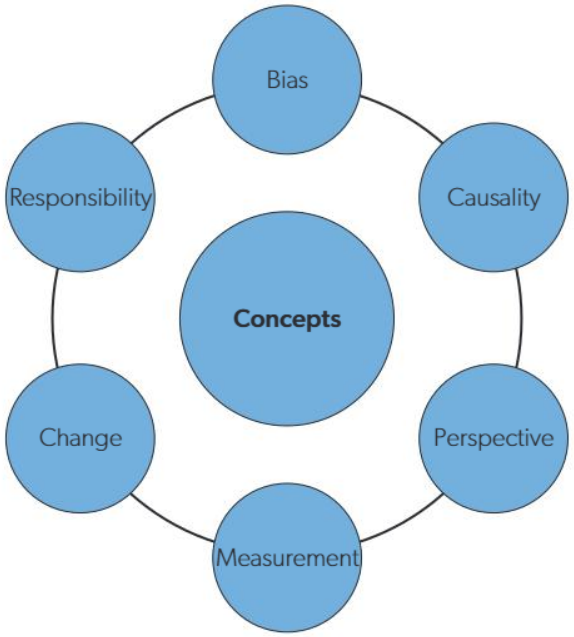
How does this apply to IB Psychology?

Like any other discipline, psychology also includes a hierarchy of concepts. The bottom of this hierarchy includes specific terms that denote very specific ideas, theories, models, or behaviours. Examples of these are the multi-store model of memory (MSMM), serotonin (a neurotransmitter), visual cortex, and attention deficit. Somewhere in the middle are more generic concepts with more explanatory power encompassing a larger number of specific ideas. For example, chemical messengers, localization of function, cognitive processes, externalizing behaviours. Close to the top are key concepts that capture the essence of many crucial problems in today's study of psychology. There are six such concepts in IB Psychology: bias, causality, perspective, measurement, change, and responsibility (see Figure 4).

What follows is a brief introduction into each of the six concepts in the IB Psychology course.

Exam tip

Make sure you occasionally return to this section and review the general descriptions of the six concepts—to get some ideas on how you could link the content you are studying to one or more of these concepts.



▲ **Figure 4** IB Psychology concepts

Bias

In a general sense, bias is a systematic deviation from the truth. Being systematic is what makes bias different from a mistake or a random error. If you measure the weight of an object using scales 10 times in a row, you will likely get 10 slightly different results each time. Some results will be above the object's true weight, some will be below. This is random error. However, if your scales are biased, they will systematically deviate in the same direction, i.e., they will underestimate or overestimate the true weight of the object. Every bias is caused by something—in other words, bias has a source. Knowing sources of bias is important if we want to eliminate it or at least take it into account.

In psychology, bias can manifest in a variety of ways:

1. Bias may be found in research. Examples include researcher bias, participant bias, sampling bias, publication bias, confirmation bias, gender bias, cultural bias, and dominant respondent bias. You will find these and other examples discussed in this book—introduced in Chapter 1 and applied to the analysis of specific research studies throughout the other chapters. Not all of them have the helpful word "bias" in the name. Order effects, experimental mortality, social desirability, and demand characteristics are all examples of sources of bias in a research study. The flipside of bias in research is its credibility: research is credible if it is free of bias. Therefore, every time you are analysing a research study through the lens of such ideas as credibility, trustworthiness, validity, reflexivity, and reliability, your analysis is relevant to the concept of bias.
2. Bias may occur in the way we interpret research findings. This is related to the researcher's beliefs or theoretical orientations. The same research findings may be interpreted differently depending on the theoretical perspective. Examples of this sort of bias include biological and environmental reductionism—theoretical positions that inflate the importance of some variables in human behaviour and downplay the importance of others.
3. Bias may occur not only in our study of behaviour, but in behaviour itself. Human behaviour may be biased. In this course, you will study such phenomena as cognitive biases, stereotypes, discrimination—these and others provide examples of how human cognition and behaviour can deviate from objectivity and rationality. A related example is bias in diagnosing mental disorders.

These are all examples of how the concept of bias links to different specific content that you will study in this course. There could be other links. All are acceptable as long as they are relevant and justified.

Causality

Causality is the relationship between variables when one variable is the cause of another.

In science, it is very important for us to understand the causes of things because it is only then that we can know their true nature. Science has four functions: to describe, to explain, to predict, and to control. We describe when we carefully document the behaviour of the observed reality (i.e., in astronomy when we document the orbits of planets). We explain when we can answer the question "why?" Planetary movements were explained when the laws of gravity and laws of classical mechanics were discovered. We understood that planets move the way they do *because* they are driven by gravity and inertia. After we have explained, we can now predict. For example, we can tell exactly when and where the next solar eclipse will take place. If we can predict, we can also control. For example, we can now create explosions on an asteroid approaching Earth to deflect it from a collision.

When psychology was emerging as a science, researchers approached it with a similar desire to study human behaviour in a way that uncovers its causes and therefore gives us the ability to predict and control. Over the course of time, it became obvious that the study of human behaviour cannot be approached in the same way as the study of an asteroid. However, the desire to uncover causal relationships wherever possible remained.

The problem with human behaviour is that it is complex and multi-determined (that is, caused by an interaction of multiple variables). Untangling these changing and interacting causal relationships is not an easy task.

This course is organized around a study of three factors that can influence human behaviour: biological, cognitive, and sociocultural. To better understand the complex causality of these factors, researchers try to isolate each of them and study their effects one by one. This links to the method of experiment—the only scientific method that allows us to make cause–effect inferences. Issues of validity, extraneous variables, and experimental controls become relevant to this discussion. Does the experiment really demonstrate that A influences B? As you can see, this aspect of causality overlaps with the concept of bias: we are more confident that an experiment allows cause–effect inferences if we know that we have controlled potential biases.

Apart from judging whether or not a study is designed in a way that allows us to make cause–effect inferences, the concept of causality is also central to interpreting findings. As already mentioned, the picture of causality in human behaviour is usually complex, and researchers try to gradually reconstruct it. For example:

- Is there bidirectional ambiguity in results?
- If we are confident that there is a cause–effect relationship, what kind of causality is it?

- Is it a direct influence or mediated by some other variables?
- Is it short term or long term?
- Are there side effects?
- Is it "domino causality" where there is a chain of events leading to the final outcome?

Finally, there is the issue of complexity and interaction. Do the factors influencing behaviour interact with each other and if they do, then how?

Perspective

A perspective is a way of looking at things. One and the same object or phenomenon can be looked at from different perspectives, and this would reveal different sides of it. To fully understand an object or a phenomenon, it is usually necessary to look at it from various perspectives and then combine them into a holistic picture of reality.

The IB Psychology course is organized around the idea that human behaviour can be studied using three broad approaches: biological, cognitive, and sociocultural. These three approaches represent the three main perspectives in psychology.

The triad "biological–cognitive–sociocultural" is very prominent in IB Psychology, but it is not the only link to the concept of perspective.

Various psychological theories or psychological models that provide different explanations of the same behaviour can also be considered perspectives.

Even on a more specific level, alternative explanations and interpretations of the same research finding could also be linked to the concept of perspective. It is a very common situation in psychology that findings of a research study may be interpreted in light of different theoretical perspectives, providing different interpretations of the result obtained in the study, and arriving at different conclusions.

Measurement

In IB Psychology, the concept of measurement refers to using research to obtain data about the behaviour of interest. Fundamentally, measurement is how we know something about behaviour. The quality of measurement determines the depth and objectivity of our knowledge.

Note: the word "measurement" has a connotation of something quantitative, expressing some observation in the form of numbers. Measurement as a concept in IB Psychology is broader than that and can equally apply to qualitative research and data that is not numerical.

Human behaviour is difficult to observe and objectively measure (e.g., think about how you could observe and measure the emotion of embarrassment). Selecting an appropriate method of measurement is a crucial task and an important challenge for a researcher.

Broadly speaking, the concept of measurement would be relevant to almost any discussion related to research methods. It will be explicitly or implicitly present in any topic because everything we know in psychology, we know because we studied it using some research methods, obtaining data and analysing it.

Some more specific examples of content that is directly relevant to the concept of measurement include the following:

- Selecting an appropriate research method to study a specific behaviour.
- Data collection tools and how they influence the credibility of the study as well as our conclusions.
- Evaluation of the quality of a research study that was used to obtain knowledge about a certain behaviour.
- Measurement in qualitative research: the role of reflexivity, ways to increase trustworthiness of findings.
- Brain imaging techniques, their strengths and limitations, and their usefulness in the context of research of a particular behaviour.
- Issues related to interpretation of research data, for example, effect size and statistical significance.

Change

Speaking very broadly, change means a process through which something becomes different. A great variety of specific examples may be related to this broad definition.

In this psychology course, there are two examples of change that you will come across most frequently:

- Development that occurs naturally: for example, the cognitive development of a child, the development of symptoms of depression in response to stress, or the development of romantic relationships over time.
- Change that is the result of a purposeful intervention: for example, treatment of a mental disorder, using technology to improve a cognitive process, or using communication strategies to repair a friendship.

Although less frequently, you may also come across other examples that may be related to the concept of change, such as prevalence rates of health problems that change not only over time, but from one population to another.

Keeping all this in mind, some of the ideas that may be relevant to the concept of change in the first sense (as development that occurs naturally) are: cognitive and social development, maturation, brain development, enculturation, acculturation, social learning, environmental influences on cognitive processes, and many others.

Some of the ideas that may be relevant to the concept of change in the second sense (as the result of a purposeful intervention) are: treatment for mental disorders, prevention of health problems, prevalence of health problems, compliance techniques, strategies for improving relationships, improving a cognitive process, operant conditioning, and so on.

Exam tip

Importantly, although some topics in IB Psychology naturally lend themselves to an exploration through the lens of a particular concept (e.g., the concept of change), any topic can be analysed through the lens of any concept. For example, the topic "Biological treatment for one disorder" is obviously related to change because it deals with psychiatric medication which we use to change symptoms of mental disorders. On the other hand, the topic "Biological explanations of mental disorders" is more obviously related to the concepts of perspective and causality. However, it does not mean that you cannot discuss biological explanations of mental disorders through the angle of change. In fact, as we mentioned earlier, explaining something (causality) gives us the opportunity to predict it and control it (change). Ultimately, all concepts are interconnected, and all topics are interconnected through them.

Responsibility

Psychologists conduct research with living beings. They also influence decisions and policies that affect the lives of people in various ways. This places psychologists in a position of power, which implies that they bear responsibility for both the process and consequences of research.

The concept of responsibility in IB Psychology mostly links to ethical considerations: in conducting a research study, in reporting the results, and in ensuring that research findings are ethically implemented.

Like many other concepts, ideas related to responsibility can be found in any topic or content that you will study. Psychological knowledge is based on research, and where there is research, there are ethical considerations. There is a lot of overlap between topics because ethical principles are universal and apply to all research equally. For example, the ethics of conducting a study of stereotypes will not be fundamentally different from the ethics of investigating the influence of chemical messengers on human behaviour.

Having said that, some ethical considerations are more prominent or relevant in some areas of study compared to others. A study of stereotypes will most probably touch upon some sensitive topics, and published results may affect public opinion or public policies in certain ways. So one needs to think carefully about the possible discomfort created in participants by the research procedure, and also the possible adverse social consequences of publishing certain results. A study of chemical messengers will likely involve some kind of ingestion (for example, taking a medicine that affects the level of neurotransmitters in the brain). Therefore, one needs to carefully think about side effects and possible physical harm, balancing this against potential benefits that can be created by the study.

Some examples that are relevant to the concept of responsibility in this course include, but are not limited to:

- Ethical considerations involved in conducting a specific research study (e.g., the use of deception, informed consent, protection from harm, right to withdraw, or debriefing).
- Ethical considerations involved in conducting research with animals.
- Decision-making of ethics committees that consider research proposals and decide if they can be approved.
- Ethical considerations involved in reporting the results of a study, both to participants themselves and to a wider community (e.g., through publication).
- Ethical considerations regulating how psychologists handle data: for example, giving access to data for independent researchers to verify the results, removing identifying information from datasets to protect the identity of participants, and so on.
- Ethical considerations involved in disseminating and applying the results of research, publishing research findings about socially sensitive issues.
- Ethical considerations involved in social advocacy. For example, ensuring that the public understands the level of uncertainty associated with research results, making sure that research results are not oversimplified in public perception and not misinterpreted, protecting minority groups from stigma, advocating for change in social policies based on results of research.



1 Research methods and data interpretation

Introduction

This chapter deals with research methods and data interpretation in psychology.

In any discipline, knowledge of research methods greatly increases our ability to understand a topic. Psychology is not an exception. The ability to evaluate psychological knowledge critically on the basis of how it was obtained is essential to avoid misconceptions.

Speaking of misconceptions, there are plenty of them in this field. Psychology is a popular discipline which makes it vulnerable to numerous popular interpretations. Therefore, it is important to clearly understand what psychology is and what it is not.

Data obtained in psychological research has to be interpreted to be meaningful. Understanding what the data means and being able to read it are important skills that add to your psychological competency.

Psychology is a special discipline. On the one hand, it is scientific, which means that psychologists, just like physicists or chemists, rigorously test hypotheses and eliminate competing explanations in an attempt to achieve objective knowledge. On the other hand, unlike natural sciences that study "nature", psychology studies humans, inherently subjective creatures. Therefore, psychology is an attempt to study the subjective (for example, the mind) objectively. Not an easy task, if you think about it.

In many respects this chapter is foundational. The skills related to analysing research methodology and interpreting data obtained in a psychological study will be crucial in all examination components.

Research methodology:

- Paper 1 (Section A and B): assessed indirectly through your ability to understand the extent to which particular examples support a concept or a theory
- Paper 1 (Section C): assessed indirectly through your ability to critically analyse and evaluate psychological ideas based on their empirical support
- Paper 2 (Section A): directly assessed through a series of questions about research methodology, rooted in the context of your class practicals
- Paper 2 (Section B): directly assessed because discussing a research study through the lens of a concept almost always requires discussing the link between the evidence obtained in the study and the theoretical conclusions derived from it
- Paper 3 (HL): assessed indirectly through your ability to understand how the findings obtained in the study should be interpreted in the context of the research method used to obtain it
- Internal Assessment: assessed directly because your research proposal reflects your understanding of how research methods work and which method is appropriate in which context, as well as an appreciation of the strengths and weaknesses of each of them.

Data analysis and interpretation:

- Paper 3: assessed directly through a series of questions based on a given stimulus material
- Paper 2 (Section A) and Internal Assessment: assessed on a basic level through your ability to understand what kind of data will be acquired as a result of each class practical, as well as the advantages and disadvantages of this type of data
- Other components: not assessed, but helpful in deepening your understanding of research findings and their typical limitations.

Additionally, research methodology and data interpretation are profoundly linked to most of the key concepts:

- Causality: because some research findings may be interpreted in terms of cause–effect relationships and some may not, and it requires an understanding of research methods to make a decision in each specific situation
- Perspective: because qualitative and quantitative methodology may be viewed as different perspectives on the investigation of human behaviour and experiences
- Measurement: because implementing a research method and interpreting its results is what constitutes the nature of measurement
- Bias: because many biases originate from measurement, and many biases are associated with sampling, the choice of research procedure, control over variables, and so on
- Change: because some research designs allow a longitudinal analysis and therefore enable us to make conclusions about how behaviour changes over time, and some do not
- Responsibility: because ethical considerations always underpin our research efforts and need to be considered in combination with methodological considerations.

Although this chapter is presented separately in the book, it is recommended that it becomes your point of reference throughout the rest of the course. A good approach would be to study this chapter first using some basic examples. Then, as you study the course, apply what you have learned here to new contexts and specific research studies. This way you will constantly exercise your skills and deepen your understanding.

1.1 What is psychology?

Inquiry questions

- Is psychology a science?
- How is a psychological investigation conducted?
- What is the focus of IB Psychology?

What you will learn in this section

Key learning:

- Psychology is the scientific study of behaviour and mental processes.
- What is a scientific study?
- What is the difference between behaviour and mental processes?
- The focus of IB Psychology.

Key terms: empirical evidence, falsifiable/falsifiability, replicate, scientific study, behaviour, mental processes

"Psychology is the scientific study of behaviour and mental processes." This is the definition we are going to use throughout this book. Although it is quite a short definition, there are a lot of implications in it. We will try to uncover them one by one.

What is a scientific study?

What does it mean that psychology is a "**scientific study**"?

This part of the definition excludes such areas as popular psychology—simple and appealing explanations that are not backed up by research evidence. Where is the line between science and non-science? This is largely a TOK question and you will return to it throughout the book, but here are some major points.

- It should be based on **empirical evidence**.
- It should be **falsifiable**—that is, it should be possible for a psychological theory to be proven wrong.
- There should be a history of independent attempts to test the theory or **replicate** the study.

In the early 20th century, the public was very interested in animal intelligence: what exactly are animals capable of? The case of Clever Hans sparked a lot of interest. Hans was a horse. Its owner Wilhelm von Osten, a mathematics teacher, claimed that he had taught Hans to solve arithmetic problems (such as addition, subtraction, multiplication, division, fractions), read, spell, and understand some German. Questions could be asked verbally or in writing, and Hans would respond by tapping his hoof a certain number of times. Von Osten exhibited

TOK

Science versus non-science demarcation is one of the key topics in TOK. The following concepts are important in the discussion of demarcation criteria:

- empirical evidence
- falsification/falsifiability
- replication.

While reading this chapter, take a note of examples that illustrate these three concepts. Think of other similar examples from such areas of knowledge as human sciences, natural sciences, and mathematics.



▲ Figure 1.1 Wilhelm von Osten and Clever Hans

the horse frequently and gained a lot of public attention. A special committee was formed in Germany (called the Hans Commission). They ran a series of tests and concluded that the performance was not a fraud and Hans's abilities were officially recognized as phenomenal.

However, another independent investigation carried out later by Oskar Pfungst, a psychologist, yielded different results. It demonstrated that Hans could not actually perform mental operations such as multiplication, but the horse was very responsive to clues that were provided by unsuspecting humans. To arrive at these conclusions, Pfungst successively tested a number of alternative hypotheses.

1. What if spectators give the horse hints or clues? He tested the horse and the questioner in the absence of spectators, but the horse continued to solve tasks correctly.
2. What if von Osten himself gives the horse some clues? Another questioner was used during several trials, but the horse's performance did not worsen.
3. What if something in the questioner gives the correct answer away and the horse can feel that? Blinders were used to test this hypothesis. It turned out that when Hans was wearing blinders, responses (the number of hoof taps) were incorrect most of the time. So, it was something in the questioner after all.
4. Did the questioners consciously let the horse know the correct answer? Additional trials were organized so that the questioner either knew or did not know the answer to the questions. It turned out that Clever Hans could only answer the questions correctly when the questioner knew the answer in advance.

This changed the focus of the research from the horse to the questioner. When Pfungst carried out his observations, it was concluded that questioners who knew the answers tended to become more tense as the hoof tapping approached the correct answer which would be reflected in their posture and facial expressions without them realizing it. This was probably the clue that the horse was using. This makes sense evolutionarily, as detection of small postural changes is important as a survival skill for horses in the wild. Clever Hans certainly was clever, but the nature of his abilities was not mathematical (Goodwin, 2010)!

This whole story shows how claims can and should be tested scientifically, that is, by conducting a systematic evidence-based investigation that puts forward one hypothesis after another and tests them in a rigorous fashion. Note also how the whole investigation attempted to falsify the existing theory rather than support it.

Discussion

How does Pfungst's investigation illustrate the concepts of empirical evidence, falsification, and replication?

ATL Thinking, Communication

What is the difference between behaviour and mental processes?

We agreed to define psychology as a scientific study of behaviour and mental processes. A scientific investigation requires an empirical approach to research—that is, relying on observation as a means of data collection. On the other hand, psychology (which comes from the Greek *psyche* = soul and *logos* = study, "the study of the soul") concerns itself with phenomena of which many are not directly observable. So we are trying to observe something that is not observable.

Behaviour is everything that can be registered: it includes overt actions as well as gestures, facial expressions, verbal responses, endocrine reactions, and so on. What stays invisible are the **mental processes** such as attention, perception, memory, and thinking. We cannot observe them directly, but we can observe the indirect effects that mental processes have on one's behaviour. So by studying behaviour, we can infer something about the mental world as well.

Activity

Consider some behavioural indicators of the following:

- attention
- anxiety
- embarrassment.

To what extent do you think it is possible to use behavioural indicators to infer these "internal" phenomena? Would the inference be reliable?

ATL Research, Thinking

Throughout this book we will use the term "behaviour" to refer to external, observable manifestations, while the term "mental processes" will be used to denote internal patterns of information processing. However, you need to be aware of the fact that the term "behaviour" is often used in a more general sense, as an umbrella term for everything psychological. Therefore, sometimes you will encounter references to mental processes as types of "behaviour". This is not exactly accurate, but acceptable in IB Psychology.

The focus of IB Psychology

IB Psychology is an academic discipline with an emphasis on rigorous research and scientific knowledge, but psychology is broader than pure academics and research. When people think about psychology many imagine counsellors and psychotherapists, practitioners who work with individual clients. University workers in lab coats conducting research are not the first thing that comes to mind. However, IB Psychology focuses on academic knowledge and scientific research rather than counselling skills. This is because thorough understanding of psychological concepts and being able to think critically about psychological phenomena is of paramount importance in all spheres of psychology, including counselling. It makes perfect sense to start with building these skills, much like the need to study aerodynamics before you are allowed to pilot an airplane.

Chat with AI

Some say that the ability to ask good questions is the most important study skill, more important perhaps than being knowledgeable about a subject. If you know what to ask and how to ask it, you will always be able to acquire new knowledge and deepen your understanding no matter what the subject.

Generative AI chatbots are great for exercising these skills. The quality of their output greatly depends on the quality of your prompt. If you ask effective questions, they serve as a good introduction into a topic or subject that is new to you.

Use your favourite generative AI chatbot to gain more introductory knowledge about psychology. Here are some prompt fragments that you might consider using:

- Is there an official definition of psychology? Could you provide several "official" definitions, as well as who they belong to? Is there anything that is common across all definitions?
- If you were to define three to five key features that make psychology different from other disciplines, what would you say?
- Why exactly is psychology considered a science? Could you name and explain one crucial characteristic that sets psychology apart from non-sciences?

Remember that generative AI also works well to help you understand or clarify something in this book. Here are a few examples you could try as you are reading this chapter:

- What does "empirical" mean in psychology? Could you explain with some examples and non-examples, and what would be the opposite of "empirical"?
- Do you know the story of Clever Hans (the horse)? I am reading about it but I can't understand how a horse could detect the right answer from the behaviour of the human.
- Could you provide a one-paragraph explanation of psychoanalysis for a student who is only just starting to learn psychology? Then do the same for behaviourism and tell me how they are different from each other.

Remember also that you can use the ability of generative AI to remember the history of the conversation:

- I am a student who is beginning to study an introduction into psychology. Could you ask me a few questions to check my understanding? After I provide answers, please give me feedback and have a conversation with me to help me understand the topic better.
- I am studying an introduction to psychology and I would like to learn to ask good inquiry questions. Going forward, when I ask you a question about psychology, could you also suggest examples of some good follow-up questions that could deepen my inquiry? My end goal is to learn to ask such questions myself.

1.2 Research methods in psychology (overview)

Inquiry questions

- What is the difference between quantitative and qualitative research?
- What are the main broad types of research methods in psychology?

What you will learn in this section

Key learning:

- Quantitative versus qualitative research: they have different purposes and the nature of the data is different.
- Brief introduction into quantitative research methods: experiment, correlational study, and quantitative descriptive study.
- Brief introduction into qualitative research methods: observation, interview, and case study.
- A research method versus a tool of data collection.

Key terms: quantitative research, nomothetic approach, qualitative research, construct, operationalization, experiment, correlational study, quantitative descriptive study, idiographic approach

Quantitative research

Data in **quantitative research** comes in the form of numbers. The aim of quantitative research is usually to arrive at numerically expressed laws that characterize behaviour of large groups of individuals (i.e., universal laws). This is much like the aim of the natural sciences in which it has been the ideal for a long time to have a set of simple rules that describe the behaviour of all material objects throughout the universe (e.g., the laws of gravity in classic Newtonian physics). In philosophy of science, such orientation on deriving universal laws is called the **nomothetic approach**.

Quantitative research operates with variables. A variable ("something that can take on varying values") is any characteristic that is objectively registered and quantified. Since psychology deals with a lot of "internal" characteristics that are not directly observable, they need to be operationalized first. For this reason, there's an important distinction between **constructs** and **operationalizations**.

- A construct is any theoretically defined variable—for example, violence, aggression, attraction, memory, attention, love, anxiety. To define a construct, you give it a definition which delineates it from other similar (and dissimilar) constructs. Such definitions are based on theories. As a rule, constructs cannot be directly observed: they are called constructs for a reason—we have "constructed" them based on theory.

Activity

In small groups, think of operationalizations of the following constructs: belief in God, assertiveness, shyness, pain, love, friendship, prejudice, tolerance to uncertainty, intelligence, wisdom.

Is it equally easy to operationalize them?

Discuss each other's operationalizations and outline their strengths and limitations.

ATL Research, Communication, Social

- To enable research, constructs need to be operationalized. Operationalization of a construct means expressing it in terms of observable behaviour. For example, to operationalize verbal aggression you might look at "the number of insulting comments per hour" or "the number of swear words per 100 words in the most recent social media posts". To operationalize anxiety, you might look at a self-reported score on an anxiety questionnaire, the level of cortisol (the stress hormone) in the bloodstream, or weight loss. As you can see, there are usually multiple ways in which a construct may be operationalized—the researcher needs to use creativity in designing a good operationalization that captures the essence of the construct and yet is directly observable and reliably measurable. As you will see throughout the examples in this book, it is often a creative operationalization that makes research in psychology outstanding.

There are three types of quantitative research.

The first type is the **experiment**. The experiment in its simplest form includes one independent variable (IV) and one dependent variable (DV), while the other potentially important variables are controlled. The IV is the one manipulated by the researcher. The DV is expected to change as the IV changes. For example, if you want to investigate the effect of psychotherapy on depression, you might randomly assign participants to two groups: the experimental group will receive psychotherapy while the control group will not. After a while, you might measure the level of depression by conducting a standardized clinical interview (diagnosis) with each of them. In this case, the IV is psychotherapy. You manipulate the IV by changing its value: yes or no. The DV is depression—it is operationalized through the standardized diagnostic procedure. All the other potentially important variables are controlled by dividing participants into the two groups randomly. If the DV is different in the two groups, you may conclude that a change in the IV has "caused" a change in the DV. This is why the experiment is the only method that allows cause-and-effect inferences.

The second type is a **correlational study**. Correlational studies are different from experiments in that the researcher does not manipulate any variables (there are no IVs or DVs). Variables are measured and the relationship between them is quantified. For example, if you want to establish if there is any relationship between violent behaviour of adolescents and how much time they spend watching violent television shows, you may recruit a sample of adolescents and measure their violent behaviour (by self-report, by ratings from classmates, or even by observation in a natural setting) and the average number of hours per day spent watching violent television shows. Then you can correlate these two variables using a formula. Suppose you obtained a large positive correlation. This means that there is a trend in the data: the more time an adolescent spends watching violent shows, the more violent they are. However, you cannot make cause-and-effect inferences from correlational studies. Since you did not manipulate one of the variables, you do not know the direction of influence. It could be the case that watching violence influences violent behaviour (this would probably be the most popular, intuitive assumption). However, it is also possible that adolescents who behave violently choose to watch violent television programmes. Or there could even be a third variable (e.g., low self-esteem) that influences both violent behaviour and watching violence on television. What you observe "on the surface" is just that: "co-relation", the fact that one variable changes as the other one changes.

Activity

In small groups, come up with results of fictitious studies that would demonstrate either correlation or causation. Here are two examples:

1. In a group of adults, we measured their attitudes to horror films and the number of siblings they have. We found that the more siblings you have, the more you like horror films.
2. We told one group of astronauts that their mission would start in a month and the other group that the mission would start in a year. We measured anxiety and found that it was higher in the group of astronauts who expected the mission to start in a month.

As you go through your list of fictitious studies, the other groups will have to say whether the study shows correlation or causation.

ATL Research, Thinking, Communication

The third type of quantitative research is a **quantitative descriptive study**. In descriptive studies, relationships between variables are not investigated, and the variables are approached separately. An example of a quantitative descriptive study would be a public opinion survey. We ask questions (e.g., "Do you support the current policies of the president?") and we are interested in the distribution of answers to this particular question. Another example is using various measures such as clinical interviews or questionnaires to investigate prevalence rates of a mental disorder.

Note: surveys or questionnaires can be used in two ways:

1. As a research method in its own right. In this case, we are dealing with a quantitative descriptive study in which the end goal is to obtain a description of the population in terms of the construct that is measured by the survey or questionnaire. For example, a study of prevalence rates aims to establish what percentage of the population experience severe depression, mild depression, no depression, and so on.
2. As a data collection tool that is used as part of a broader study. In this case, the research method would be either the experiment or a correlational study, and the survey or questionnaire would be used to measure a variable, but not as a research method in its own right. For example, suppose we investigate the relationship between aggressiveness and insecurity among high school students. We might measure their aggressiveness and insecurity with special questionnaires, but the point of the research is to investigate the correlation between these two variables. This will be a correlational study in which questionnaires are used as a tool of data collection.

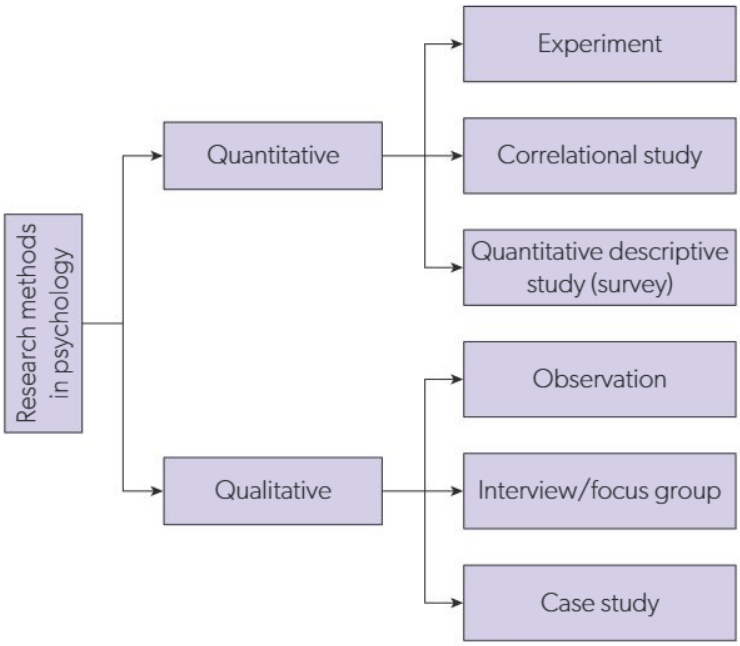
Qualitative research

Qualitative research is different. Its main focus is an in-depth study of a particular phenomenon. "In-depth" entails going beyond what can be objectively measured and quantified into the realm of human experiences, interpretations, and meanings. Qualitative research makes use of such methods as interviews or observations. Data usually comes in the form of texts: interview transcripts, observational notes, and so on. Interpretation of data involves a degree

of subjectivity, but analysis is deeper than we can usually achieve through quantitative approaches. In philosophy of science, such orientation on an in-depth analysis of a particular case or phenomenon (without trying to derive universally applicable laws) is called the **idiographic approach**.

Qualitative research methods that we will discuss in this chapter are:

- observation
- interview (including focus group)
- case study.



▲ **Figure 1.2** Research methods in psychology

Our brief introduction to methods of research in psychology is summarized in Table 1.1 and Figure 1.2.

Parameter	Quantitative research	Qualitative research
Aim	Nomothetic approach: derive universally applicable laws	Idiographic approach: in-depth understanding of a particular case or phenomenon
Data	Numbers	Texts
Focus	Behavioural manifestations (operationalizations)	Human experiences, interpretations, meanings
Objectivity	More objective (the researcher is eliminated from the studied reality)	More subjective (the researcher is included in the studied reality)

▲ **Table 1.1** Quantitative versus qualitative research

1.3 Analysing research (methodological considerations)

Inquiry questions

- What parameters can we use to make a judgement about the quality of a research study?
- Is this similar in qualitative and quantitative research?

What you will learn in this section

Key learning:

- The main parameters that are commonly used in making a judgement about the quality of a research study are generalizability, sampling, credibility, and bias.
- These parameters are approached differently in qualitative and quantitative research, sometimes with different terms referring to similar ideas.

Key terms: generalizability, sampling, credibility, bias

Overview

No research study is perfect. Every study has strengths and limitations. It is important to understand exactly what aspects of the study can and cannot be trusted. Understanding in what ways a study is limited is useful because you can design additional studies to address these specific limitations, and collectively this research will get you closer to knowing the truth.

In this short section, we will consider the main parameters that are commonly used in making a judgement about the quality of a research study. Later as we discuss research methods one by one, we will delve more deeply into the analysis of methodological considerations specific to those methods.

It is important to note that there are inconsistencies in terminology associated with the long history of the development of standards of scientific research in social sciences. It is especially visible between quantitative and qualitative research as they frequently use different terms to refer to the same or similar idea.

TOK

The term "methodological" refers to the procedures related to the methods of obtaining knowledge. Therefore, methodological analysis is always an analysis of the strengths and limitations that led to the acquisition of a certain knowledge, and the extent to which we can trust this knowledge based on this analysis.

What forms do you think a methodological analysis would take in various areas of knowledge other than human sciences?

Generalizability and sampling

Generalizability is the extent to which we can generalize the results of a research study beyond the study itself. There are at least three aspects of generalizability.

The first aspect of generalizability is generalizing results from the participants that actually took part in the study (the sample) to other people. **Sampling** is

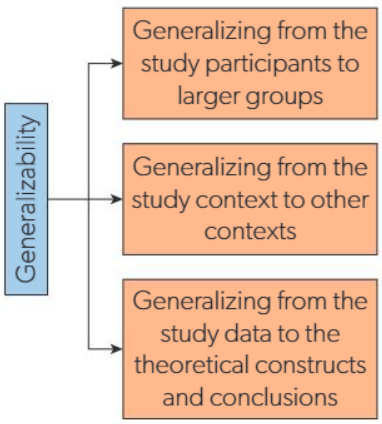
the process of finding and recruiting individuals for the study. There are different sampling techniques, and it is important to be aware of their strengths and limitations as sampling may affect the results of the study.

- In quantitative methods (such as the experiment), the idea that results of a study can be generalized to larger groups of people is referred to as population validity. We believe that population validity is high if the sample is representative of the target population—that is, it reflects all the essential characteristics of the target population in miniature.
- In qualitative research, a similar idea is known as sample-to-population generalizability, or, in some sources, inferential generalizability. Sampling in qualitative research is not the same and so the idea of representativeness is not applicable. Judgements about sample-to-population generalizability are based on different principles (see page 40).

The second aspect of generalizability is generalizing results of the study to other situations and settings. Again, slightly different terminology is used in qualitative and qualitative research to refer to this idea.

- In the experiment, the commonly used term is ecological validity. Sometimes experiments are conducted in an artificial setting (e.g., a laboratory), but we want to believe that people will behave the same way in their daily life as well. Ecological validity is the extent to which results of an experiment may be applied to natural real-life settings.
- In qualitative research, there is a similar idea known as transferability, or case-to-case generalizability. This is the extent to which we believe that the findings of a qualitative research study can be applied to a different, but similar context.

The third aspect of generalizability is the extent to which we can generalize from the data obtained in the study to the theoretical conclusion we infer from this data. In quantitative research (e.g., experiments, correlational studies, questionnaires), this idea is commonly referred to as construct validity. Construct validity is the extent to which we believe the study "measures" our hypothesized theoretical construct and not something else. In qualitative research, the term "construct validity" is not used, although a similar idea is referred to as theoretical generalizability.



▲ Figure 1.3 Aspects of generalizability

Credibility and bias

Credibility refers to the degree to which the results of the study can be trusted to reflect reality. It is closely linked to **bias**, because the results of the study do not reflect reality if there was some sort of bias in it. There are a lot of "traps" that a researcher can walk into. For example, in an interview, while the researcher believes the interviewee's responses to be true, participants may actually guess the aim of the study. Then they may respond in a way that they think the researcher is expecting them to. The credibility of a research study is believed to be high if it is shown that potential sources of bias were, to the best of our knowledge and abilities, controlled or eliminated.

Again, there is a terminological inconsistency between quantitative and qualitative research due to their separate histories of development. In the experiment, this idea is referred to as internal validity. Internal validity is high if we believe that it is the manipulation of the independent variable, and not

something else, that has led to the observable change in the dependent variable. In qualitative research, the term "validity" is not used. However, the terms "credibility" or "trustworthiness" are used to refer to a similar idea. We believe that the results of a qualitative study are credible if we think that the way we formulated conclusions reflects participants' actual experiences.

Table 1.2 gives an overview of the main concepts used to characterize sampling, generalizability, credibility, and bias in experimental as opposed to qualitative research. As you read on, you will understand these concepts better. Refer to Table 1.2 from time to time so that you can place the concepts clearly within the general framework.

Overarching idea	Quantitative research (the experiment)*	Qualitative research (interview, observation)
Generalizability		
Generalizing results to a larger group of people	Population validity**	Sample-to-population generalizability/ inferential generalizability
Generalizing results to real-life settings	Ecological validity**	Case-to-case generalizability/transferability
Generalizing results from the observation to the theoretical construct	Construct validity	Theoretical generalizability
Credibility		
The degree to which results of the study reflect reality	Internal validity	Credibility/trustworthiness
Sampling		
Types of sampling	Random Stratified Self-selected Opportunity	Quota Purposive Theoretical Snowball Convenience
Bias		
Types of bias	Threats to internal validity: selection, history, maturation, testing effect, instrumentation, regression to the mean, experimental mortality, experimenter bias, demand characteristics	**Participant bias:** acquiescence, social desirability, dominant respondent, sensitivity **Researcher bias:** confirmation bias, leading questions bias, question order bias, sampling bias, biased reporting

*Other quantitative research methods (correlational study and quantitative descriptive research) have a slightly different use of terminology—see later on in this chapter. There are also special types of validity in psychometrics associated with analysing the quality of a questionnaire.
**Together, the population validity and ecological validity of the experiment are also known as external validity.

▲ Table 1.2 A comparison of similar ideas of generalizability, sampling, credibility, and bias in quantitative (the experiment) and qualitative research

1.4 The experiment

Inquiry questions

- Why do experiments allow cause-and-effect inferences?
- How can bias in experimental research be prevented?
- How can findings from a small group of people be generalized to an entire population?
- How can experiments be designed?

What you will learn in this section

Key learning:

- Confounding variables.
- Sampling in the experiment: representativeness, random sampling, stratified sampling, convenience (opportunity) sampling, self-selected sampling.
- Experimental designs: independent measures design, matched pairs design (matching variable), repeated measures design (order effects; counterbalancing).
- Types of validity: construct validity, internal validity, external validity (population and ecological).
- Bias in experimental research: threats to internal validity: selection, maturation, testing effect, mortality, demand characteristics, experimenter bias.
- True experiments, quasi-experiments and non-experiments.
- Natural experiments.
- Laboratory and field experiments.

Key terms: cause-and-effect inferences, confounding variables, random sampling, stratified sampling, convenience (opportunity) sampling, self-selected sampling, independent measures design, matched pairs design, repeated measures design, counterbalancing, construct validity, internal validity, external validity, population validity, ecological validity, selection, maturation, testing effect, experimental mortality, demand characteristics, experimenter bias, double-blind design, quasi-experiment, true experiment, non-experiment, natural experiment, laboratory experiment, field experiment

Confounding variables

As we mentioned before, the experiment is the only method that allows researchers to make **cause-and-effect inferences**. This is achieved by defining the independent variable (IV) and the dependent variable (DV), manipulating the IV and observing how the DV changes in response to this manipulation.

Psychological reality is very complex—the trick is to isolate the IV so that when you manipulate it, nothing else changes.

Think about the following example: you are investigating the influence of sleep deprivation on memory performance. You manipulate sleep deprivation by waking up one group of participants every 15 minutes when they sleep, while the control group sleeps normally. You measure memory performance by a simple memory test in the morning. Sleep deprivation is the independent variable and memory performance is the dependent variable. Without realizing that this might be an important factor, you let the control group sleep at home while the experimental group sleeps in a laboratory being supervized by an experimenter. So there is another variable: sleep environment. It could be the case that in this experiment it was the unfamiliar environment that caused a reduction in memory performance, rather than sleep deprivation.

Variables that can potentially distort the relationship between the IV and the DV (e.g., sleep environment in the example above) are called **confounding variables**. They contribute to bias. These variables need to be controlled, either by eliminating them or keeping them constant in all groups of participants so that they do not affect the comparison.

Discussion

How could the researchers have controlled the confounding variable in this example?

ATL Research, Thinking

Sampling in the experiment

Being a truly nomothetic method, the experiment aims at discovering universal laws of behaviour applicable to large groups of people across a variety of situations. This makes relevant the distinction between the sample and the target population. The target population is the group of people to which the findings of the study are expected to be generalized. The sample is the group of people taking part in the experiment itself. How can we ensure that whatever results are obtained in the sample can be generalized to the target population? We do this through representativeness—the key property of a sample. A sample is said to be representative of the target population if it reflects all its essential characteristics.

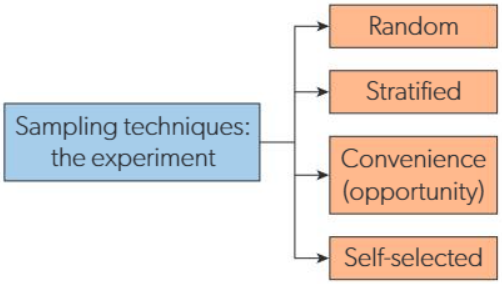
Activity

Imagine you are investigating the influence of praise on the school performance of teenagers. For this experiment, you need to have a sample of participants that you would split into two groups (experimental and control). In the experimental group, the teacher is instructed to praise every student three times a week while in the control group, the teacher is told to only praise the students once a week. At the end of the research period, performance grades in the two groups are compared.

Will you be able to generalize the findings to the target population—that is, teenagers in general? This depends on how representative your sample is. If the sample is representative, it must reflect the essential characteristics of the target population.

Given the aim of the study, how would you increase representativeness of your sample?

ATL Research, Thinking



▲ Figure 1.4 Sampling techniques used in experiments

Several sampling techniques can be used in an experiment. The choice depends on the aim of the research, the availability of resources, and the nature of the target population.

Random sampling. This is the ideal approach to make the sample representative. In random sampling, every member of the target population has an equal chance of becoming part of the sample. With a sufficient sample size this means that you consider all possible essential characteristics of the target population, even the ones you never suspected to play a role. Arguably, a random sample of sufficient size is a good representation of a population, making the results easily generalizable. However, random sampling is not always possible for practical reasons. An example of random sampling strategy is a pre-election telephone survey where participants are selected randomly from social media profiles. Even in this case, though, you have to admit that the target population is not all the citizens of a particular country—it is all the citizens of the country who have a particular type of social media profile.

Stratified sampling. This approach is more theory-driven. First you decide the essential characteristics the sample has to reflect. Then you study the distribution of these characteristics in the target population (for this you may use statistical data available from various agencies). Then you recruit your participants in a way that keeps the same proportions in the sample as is observed in the population. For example, imagine that your target population is all the students in your school. The characteristics you decide are important for the aim of the study are age (primary school, middle school, high school) and grade point average—GPA (low, average, high). You study school records and find out the distribution of students across these categories: see Table 1.3.

	Low GPA	Average GPA	High GPA	Total
Primary school	0%	10%	10%	20%
Middle school	5%	30%	15%	50%
High school	5%	20%	5%	30%
Total	10%	60%	30%	100%

▲ Table 1.3 An example to illustrate stratified sampling: grade point average (GPA) in various sub-populations of school students

For a stratified sample, you need to ensure that your sample has the same proportions. For every cell of Table 1.3, you can either sample randomly or use other approaches. In any case, what makes stratified sampling special is that it is theory-driven and it ensures that theory-defined essential characteristics of the population are fairly and equally represented in the sample. This may be the ideal choice when you are certain about essential participant characteristics and when available sample sizes are not large.

Convenience (opportunity) sampling. For this technique, you recruit participants that are more easily available. For example, university students are a very popular choice because researchers are usually university professors, so it is easy for them to find samples at their place of work. There could be several reasons for choosing convenience sampling. First, it is the technique of choice when financial resources and time are limited. Second, there could be reasons to believe that people are not that different in terms of the phenomenon under

study. For example, if you study the influence of caffeine on attention, you believe that this is something common to all people. Finally, convenience sampling is useful when you are conducting a pilot study and you are not sure the hypothesis will be supported by evidence. The limitation of convenience sampling is, of course, a lack of representativeness.

Self-selected sampling. This refers to recruiting volunteers. An example of this approach is advertising the experiment in a newspaper and using the participants who respond to the advert. The strength of self-selected sampling is that it is a quick and relatively easy way to recruit individuals while at the same time having wide coverage. The most essential limitation, again, is representativeness. People who volunteer to take part in experiments may be more motivated than the general population, or they may be looking for incentives. For example, in many studies participants are financially rewarded for their time.

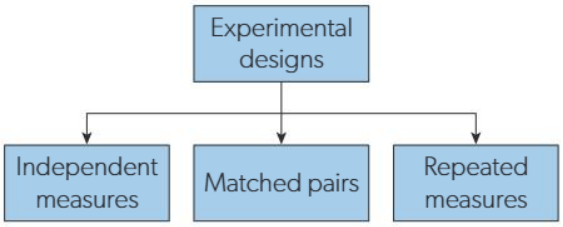
Discussion

Now that you know what sampling strategies can be used in an experiment, how would you change your approach to recruiting a sample for the investigation of the influence of praise on the academic performance of teenagers?

ATL Research, Thinking

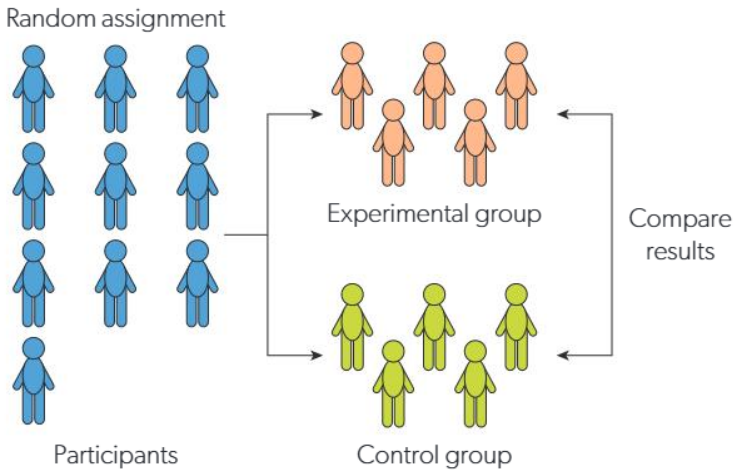
Experimental designs

Experiments always involve manipulating some variables and measuring the change in others. However, the specific ways in which this can be organized differ depending on the aims of the research. The organization of groups and conditions in an experiment is known as the experimental design, and there are three basic types of experimental design: independent measures, matched pairs, and repeated measures.



▲ Figure 1.5 Experimental designs

Independent measures design involves random allocation of participants into groups and a comparison between these groups. In its simplest form, you randomly allocate participants from your sample into the experimental group and the control group.



▲ Figure 1.6 Independent measures design

Exam tip

Be mindful of the difference between random sampling (selecting the sample from the target population) and random group allocation (dividing your sample into groups). It is possible to have random group allocation in non-random samples and vice versa.

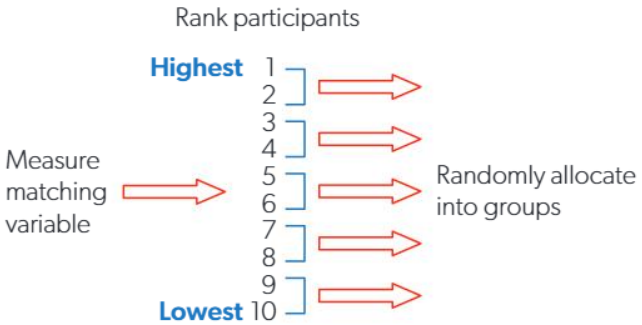
The rationale behind random group allocation is that all potential confounding variables cancel each other out. If the groups are not equivalent at the start of the experiment, it would be like comparing apples with oranges. Imagine that you are testing the hypothesis that praise at school improves students' performance and for this you take two existing groups of students, with one being rarely

praised by their teachers and the other one often praised. Arguably, the groups might not be equivalent: they have different experiences with the teachers, different ingroup values and habits, and so on. It is impossible to account for all these potentially important factors.

Conversely, when the group sizes are sufficiently large and allocation is completely random, it is likely that groups will be equivalent—the larger the sample, the higher the chance.

This allows us to assume that the groups are equivalent from the start so whatever difference we observe at the end of the experiment must have been caused by our experimental manipulation.

Matched pairs design is similar to independent measures. The only difference is that instead of completely random allocation, researchers use matching to form the groups. To illustrate matching, let us consider an example. Suppose you are conducting a study of the effect of sleep deprivation on memory. For this, you need two groups of participants. One of the groups will sleep peacefully in the laboratory and the other group will be woken up every 15 minutes. In the morning, you will give both groups a memory test and compare their performance. It is important that the two groups at the start of the experiment are equivalent in their baseline memory abilities. Random allocation will usually allow this, but you only have 20 participants (10 in each group). With a small sample like this there is a chance that random allocation will not work. So you want to control the equivalence of memory abilities "manually" while leaving everything else to random chance. For this, you test memory abilities in your participants prior to the experiment. Then you rank participants according to their memory abilities (e.g., from the highest to the lowest). Then you take the first two participants from the top of the list and randomly allocate one of them to the experimental group and the other one to the control group. You take the next two participants and repeat the procedure for the rest of the list. The two resulting groups are certainly equivalent in terms of memory abilities and probably (due to random chance) equivalent in all other characteristics.



▲ Figure 1.7 Matched pairs design

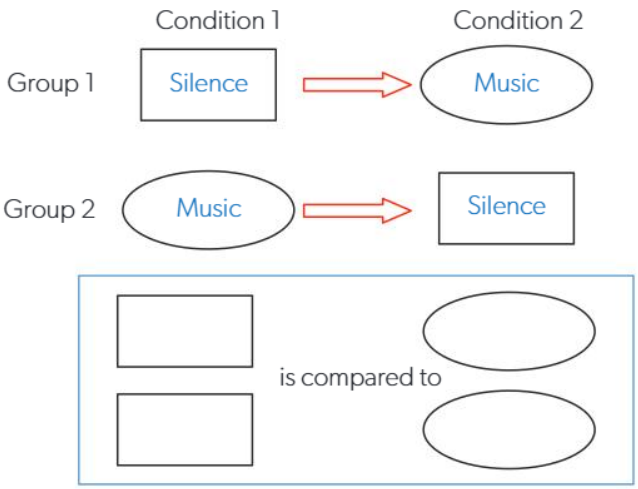
The variable that is controlled (memory abilities, in this example) is called the matching variable. Matched pairs designs are preferred when:

- the researcher finds it particularly important that the groups are equivalent in a specific variable
- the sample size is not large, therefore there is a chance that random allocation into groups will not be sufficient to ensure group equivalence.

Repeated measures design is used when the goal is to compare conditions rather than groups of participants. The same group of participants is exposed to two (or more) conditions, and the conditions are compared. For example, imagine your aim is to investigate the effect of classical music on learning. You ask your participants to learn a list of words for 10 minutes in silence, then you ask the same participants to learn a different list of words for another 10 minutes, but this time with classical music playing in the background. You compare results from the first and the second trial.

The problem with repeated measures designs is that they are vulnerable to order effects: results may be different depending on which condition comes first (e.g., silence then classical music or classical music then silence). Order effects may appear due to various reasons, such as practice or fatigue.

To overcome order effects, researchers use **counterbalancing**. This involves using other groups of participants in which the order of the conditions is reversed. For our example, two groups could be used: one given the sequence "silence then music" and one given the sequence "music then silence". It is important to note that comparison will still be made between conditions, not between groups. Data from group 1, condition 1 will be collated with data from group 2, condition 2, and vice versa. These two collated data sets will be compared.



▲ Figure 1.8 Repeated measures design with counterbalancing

Types of validity

The quality of experiments is characterized by their construct, internal, and external validity.

Construct validity characterizes the quality of operationalizations. As you know, the phenomenon under study is first defined theoretically as a construct and then expressed in terms of observable behaviour (operationalization). Operationalization makes empirical research possible. At the same time when results are interpreted, research findings are linked back to constructs. Moving from an operationalization to a construct is always a bit of a leap. Construct validity of an experiment is high if this leap is justified and if the operationalization provides sufficient coverage of the construct.

For example, in some research studies anxiety was measured by a "fidgetometer", a specially constructed chair that registers movements at various points and so calculates the amount of "fidgeting". Subjects would be invited to the laboratory and asked to wait in a chair, not suspecting that the experiment has already started. The rationale is that the more anxious you are, the more you fidget in the chair (Klorman, 1975).

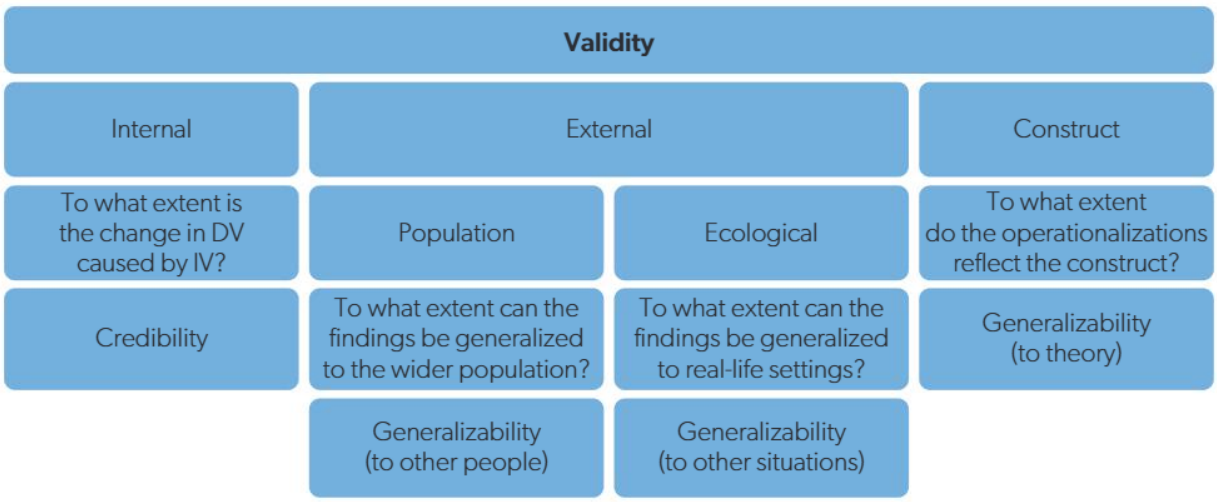
Are the readings of a fidgetometer a good operationalization of anxiety? On the one hand, it is an objective measure. On the other hand, fidgeting may be a symptom of something other than anxiety. Also, the relationship between anxiety and increased fidgeting first has to be demonstrated in empirical research.

Internal validity characterizes the methodological quality of the experiment. Internal validity is high when confounding variables have been controlled and we are quite certain that it was the change in the IV (not something else) that caused the change in the DV. In other words, internal validity links directly to bias: the less bias, the greater the internal validity of the experiment. Biases in the experiment (threats to internal validity) will be discussed in more detail later on.

External validity characterizes generalizability of findings in the experiment. There are two types of external validity: population validity and ecological validity.

1. **Population validity** refers to the extent to which findings can be generalized from the sample to the target population. Population validity is high when the sample is representative of the target population and an appropriate sampling technique is used.
2. **Ecological validity** refers to the extent to which findings can be generalized from the experiment to other settings or situations. It links to the artificiality of experimental conditions. For example, in memory experiments participants are often asked to memorize long lists of trigrams. To what extent can findings from such studies be applied to everyday learning situations?

There is an inverse relationship between internal validity and ecological validity. To avoid bias and control for confounding variables, you make the experimental procedures more standardized and artificial. This reduces ecological validity. Conversely, in an attempt to increase ecological validity you may allow more freedom in how people behave and what settings they choose, but this would mean that you are losing control over some potentially confounding variables.



▲ Figure 1.9 Validity of experiments and how it links to the concepts of generalizability and credibility

Activity

Leaf through this book, find a description of any experimental study and analyse its construct, internal, and external validity. If you feel that you do not have enough detail, you could find more information on the study online, or even by reading the original article.

Present the results of your analysis in class.

ATL Thinking, Communication, Self-management

Bias in experimental research: threats to internal validity

Bias in experimental research comes in the form of confounding factors that may influence the cause-and-effect relationship between the IV and the DV. This decreases internal validity. The following shows a description of some of the common sources of threats to internal validity, based on a classic text by Campbell (1957).

Selection bias

This threat to internal validity occurs if groups are not equivalent at the start of the experiment for some reason: apart from the planned IV-related difference, they differ in some other variable. As a result, we cannot be sure if the post-experiment differences between groups reflect the influence of the IV or this other variable. **Selection** bias occurs in independent measures and matched pairs designs in cases where group allocation was not completely random.

Maturation bias

In the course of the experiment participants go through natural developmental processes, such as fatigue or simply growth. For example, suppose you are piloting a psychological training programme to increase assertiveness in middle school students. You measure assertiveness at the start, conduct the training programme for several months, and measure assertiveness again. The resulting increase of assertiveness may be due to either the IV (the training) or simply to the fact that the middle school students grew up a little and naturally became more assertive. This is called **maturation** bias. The counteracting strategy would be using a control group (the same time period, the same measurements but no training sessions).

Testing effect

The **testing effect** is when the first measurement of the DV may affect the second (and subsequent) measurements. For example, suppose you are investigating the effectiveness of a video to reduce test anxiety in primary school children. For this, your participants take an ability test preceded by a self-report anxiety measure at time A. They then watch your specially designed video and repeat the procedure (test and self-report anxiety measure) at time B. The difference in anxiety between time A and time B may be the result of both the video and the fact that it is their second time taking the test—they are more familiar with the format and therefore may be naturally less anxious. A solution to this is to use a control group where you show a neutral video of the same duration.

Suppose you get the results in Table 1.4.

	Test anxiety (on a scale 0–100)	
Group	Before Test 1	Before Test 2
Experimental (specially designed video)	90	55
Control (neutral video)	90	70

▲ Table 1.4 Results of a hypothetical study illustrating the testing effect

Analysis of these results can reveal that a reduction of anxiety by 20 points is probably due to the testing effect. However, over and above that there is a 15-point anxiety effect of the specially designed video.

In repeated measures designs, testing effect is a special case of order effects, and counterbalancing is used to control for it.

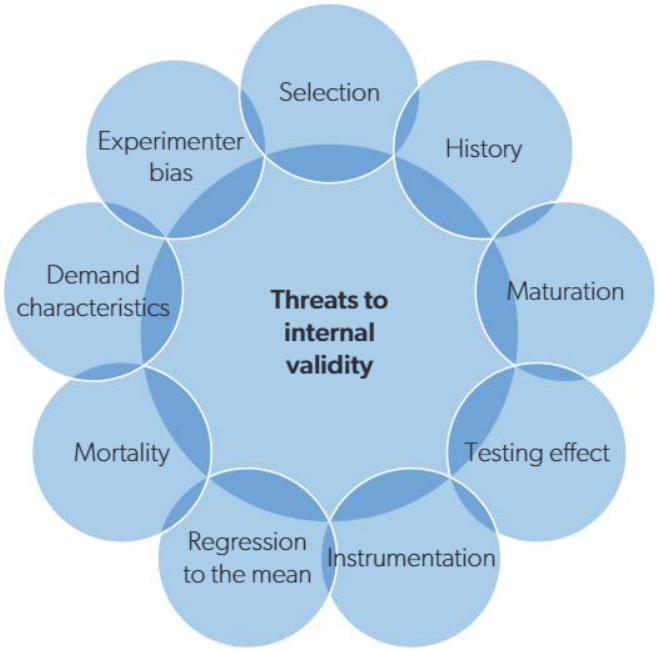
Activity

We have omitted some of the threats to internal validity that were discussed in Campbell's (1957) classic text because they are less common. If you are interested to know more, do some research into these other threats to internal validity and try to identify an example of where this bias may show up in psychological research.

In particular, check out the following threats to internal validity:

- history
- instrumentation
- regression to the mean.

ATL Research, Self-management



▲ Figure 1.10 Sources of threats to internal validity (after Campbell, 1957)

Experimental mortality (threat to internal validity) refers to the fact that some participants drop out during an experiment, which may become a problem if dropouts are not random. Suppose you are investigating the influence of emotion on ethical decision-making. For this you give your participants a number of scenarios of the type "Would you harm one person to save 1000?" In the control group the description of this "one person" is neutral, but in the experimental group this is someone they know personally, so there is more emotional involvement. You hypothesize that people will be less likely to be utilitarian in their decision-making when they are personally involved. (Note: this research would create distress among participants and so raises ethical issues—it is quite possible such a study would not be approved by an ethics committee.)

Suppose that several participants in the experimental group refuse to continue participation and drop out, more so than in the control group. Ethical issues aside, this also presents a methodological issue: even if the two groups were equivalent at the start of the experiment, they may be non-equivalent now. There appears a confounding variable (sensitivity) which is disproportionately represented in the two groups. There is no reliable way to counteract experimental mortality other than designing experimental conditions in such a way that participants would not feel the need to drop out.

Demand characteristics refer to a situation in which participants understand (or think that they understand) the purpose of the experiment and change their behaviour subconsciously to fit that interpretation. In other words, they behave in ways that they think the experimenter expects in the study.

This can happen for various reasons—for example, some participants may try to please the experimenter by being a "good subject", whereas others may intentionally behave in a way that will mess up what they believe is the hypothesis of the experiment.

To avoid demand characteristics, deception may be used to conceal the true purpose of the study. However, deception raises ethical issues. You can consider using post-experimental questionnaires to find out to what extent demand characteristics may have influenced the results (this strategy does not prevent demand characteristics but just estimates their impact). Note: in repeated measures designs demand characteristics are a larger threat because participants take part in more than one condition and so have greater opportunities to realize or guess the aim of the study.

Experimenter bias refers to situations in which the researcher unintentionally exerts an influence on the results of the study—as, for example, in the Clever Hans case discussed earlier. Existence of this bias was first rigorously supported by Rosenthal and Fode (1963). In their experiment rats were studied for their maze-running performance. Rats were split into two groups at random, but the laboratory assistants (psychology students) were told that one of the groups was "maze-bright" and the other one was "maze-dull" and that this difference in ability was genetic.

Laboratory assistants had to follow a standardized experimental procedure for which rats were tested on their performance in learning the maze task, but the results showed that the rats labelled "maze-dull" performed significantly worse than the ones labelled "maze-bright". It was concluded that the result was an artefact: it was caused by experimenter bias rather than any genuine differences between the groups of rats.

Post-experiment investigations revealed that experimenter bias was not intentional or conscious. The results were induced by subtle differences in the way laboratory assistants handled the rats. For example, without realizing, assistants handled "maze-bright" rats for slightly longer and so stress was more reduced for these rats than for "maze-dull" rats. A countermeasure against experimenter bias is using so-called **double-blind designs**, where information that could introduce bias is withheld both from the participants and from the people conducting the experiment. The study of Rosenthal and Fode would have been double-blind if the laboratory assistants had not been told which group of rats had which label.

Activity

Athabasca University in Canada has a great learning resource on threats to internal validity. One tutorial consists of two parts:

Part 1: the theoretical background and definitions.

Part 2: a practical exercise involving the analysis of 36 hypothetical experiments.

If you want to practise identifying potential sources of bias in experiments, carry out an internet search for "Athabasca University internal validity tutorial".

ATL Research, Thinking, Self-management

True experiments, quasi-experiments, and non-experiments

One of the key features of a **true experiment** is equivalence of groups. We assume that the groups we are comparing are equivalent in every respect except for the presence or absence of the experimental manipulation. To make this possible, we rely on random allocation of participants into groups.

Compare this to the opposite situation in which instead of allocating participants into conditions, we use pre-existing groups. For example, we compare male participants to female participants, Chinese to Canadian, teenagers to adults, people who have a pet cat to people who have a pet dog, people who play the guitar to people who do not. All these are examples of non-equivalent groups. Such groups are different from each other in multiple ways. If we discover that a measurement (e.g., memory) is different in the two groups, we cannot be sure what exactly this difference should be attributed to.

In fact, although such studies are based on a comparison between groups, we may think of them as correlational. Group membership may be coded as a variable with two possible values (e.g., 1 = male; 2 = female), and then we investigate the correlation between this group membership and the other variable (e.g., memory). As a result, we can conclude that the two variables are correlated (there is a tendency for one variable to change as the other changes), but we cannot say that one variable influences the other.

We can think of true experiments versus comparison studies with non-equivalent groups as two extremes on one continuum. True experiments use random group allocation, so we are certain that group membership is the only difference between otherwise equivalent groups, and therefore we can make cause–effect

inferences. In comparison studies there are pre-existing groups, no random allocation, groups are different in many ways and are not equivalent, and therefore we cannot make cause–effect inferences.



▲ Figure 1.11 A cut-off point: a research study is either experimental or not

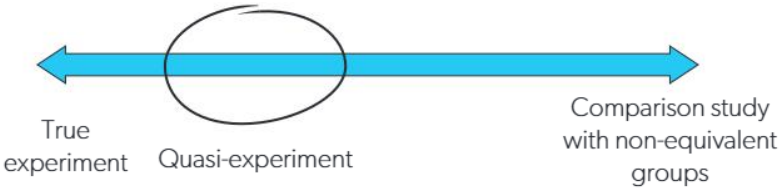
However, there is a grey area in between where we have studies with group allocation that was controlled, but not entirely random. On our continuum, such studies are closer to true experiments, but they lack some features of a true experiment.

For example, Sharot et al. (2007) compared two groups of participants who witnessed the 9/11 attacks in New York City in 2001: those who were close to the attack site in Downtown Manhattan when it happened, and those who were in Midtown, a few miles away. One may argue that the groups are quite equivalent: it was a matter of chance what part of the city the participant was in when the attack took place. On the other hand, it is still a fact that Sharot did not allocate participants into the Downtown and the Midtown condition randomly herself. It is still possible that there exist some differences that can affect the results of the study. For example, people who were in Downtown Manhattan at that time of the day could differ from the Midtown groups in terms of their socioeconomic status. How do we treat this study: is it an experiment or a **non-experiment**? Can we make the conclusion that the proximity to the site of the attack has influenced whatever differences between the two groups were observed in the study? (For a more detailed description of the study, see Unit 2.3: Biological factors in cognitive processes.)

Another example is the study of Maguire et al. (2000) who compared brain scans of London taxi drivers to brain scans of a control group of subjects, who were similar in age and other characteristics but did not drive a taxi. Some differences were found, most notably in the relative distribution of grey matter in the hippocampus. These differences were interpreted as evidence that the hippocampus is responsible for spatial memory. However, researchers did not randomly allocate participants into the group of taxi drivers and controls. They took pre-existing groups. Although unlikely, it is still conceivable that these groups were different in some other respects that remained unnoticed by the researchers. What if it is not the taxi driving experience, but these other unknown characteristics that influenced grey matter in the hippocampus?

There are two ways of looking at these studies residing in the grey area. There is no consensus among researchers as to which of the two should be preferred, leading to some debate in this area.

1. We could strictly draw a line between true experiments and non-experiments by saying that anything that does not have all the crucial features of a true experiment is a non-experiment. From this point of view, any study that did not use random allocation into groups or conditions will be considered a non-experiment. Cause–effect inferences from such studies will not be allowed.
2. We could use a softer approach and allow some of the studies from the "grey area" to be interpreted in terms of cause and effect, albeit with caution.



▲ Figure 1.12 The "grey area": some studies are designed as experiments, but do not implement enough control to be considered true experiments

It is within this softer approach that the term **quasi-experiment** was coined by Cook and Campbell (1979). "Quasi" is a prefix meaning "almost". Under the influence of behaviourism, psychology in the 20th century was dominated for a long time by the idea that a well-controlled true experiment conducted in the laboratory is the most scientific method and is superior to all other methods. However, as psychology kept developing, it became clear that perfectly controlled experiments are not always possible, especially if one wants to study participants in their natural environments. Cook and Campbell (1979) believed that by using the strict approach we are losing a lot of valuable information and dismissing some potentially valuable evidence. They recognized the "grey area". In using the softer approach, they recognize that quasi-experiments are limited and that as standalone studies they cannot be used to infer cause–effect relationships, but at the same time they claimed that combining evidence from quasi-experiments with other pieces of evidence does allow us to infer cause–effect with caution.

Exam tip

We accept this softer approach in the IB programme. We can therefore visualize a continuum with three points on it:

- true experiments
- quasi-experiments
- non-experiments.

While the boundary between true experiments and quasi-experiments is quite well-defined, the same cannot be said about the boundary between quasi-experiments and non-experiments. The closer we are to true experiments, the more confident we are that the study was well-controlled, the independent variable was manipulated, and the groups are equivalent—therefore, cause–effect inferences are possible. The closer we are to non-experiments, the more reasons we have to believe that experimental controls were not perfect, the independent variable was not manipulated by the researcher and the groups are not equivalent. Therefore, we can only make a generic conclusion that the variables are related to each other, but we cannot know if this is a relationship of cause and effect.

However, it is safe to say that a study is definitely a non-experiment when there is no treatment at all applied to either of the groups. It can be argued that in Maguire's study of London taxi drivers, for example, the "treatment" was acquiring taxi driving experience over a period of time (although this treatment was not caused by the experimenter). Similarly, in Sharot et al.'s study the "treatment" was the place where the participant happened to be during the 9/11 attack. On the other hand, when we compare Chinese participants to Canadian participants, for example, there is no treatment at all. Being Chinese and being Canadian are the participants' characteristics since birth—it was not something that was "applied" to them over a limited period of time.

Natural experiments

Natural experiments can be seen as a special type of a quasi-experiment. In natural experiments, the experimental manipulation occurs naturally and has not been caused by a researcher, although there are reasons to believe that group allocation (i.e., people who experienced the natural event versus those who did not) is fairly random.

An example of a natural experiment is the introduction of a governmental subsidy for families below the poverty line in an attempt to bring them out of poverty. Let us suppose that at a particular time, this subsidy was introduced in some states but not in others. We can then compare various parameters (e.g., cognitive and social development of children in these families) between the states that were and were not included in the subsidy programme.

Another example is Charlton et al. (2000) who studied how television affected children when it was first introduced on St. Helena island in the mid-1990s. They measured a number of variables related to prosocial and antisocial behaviour four months before television started and five years after its introduction. They found a decrease in some of the measures of prosocial behaviour.

Exam tip

Natural experiments could be categorized as quasi-experimental because, although the researcher did not have any control over the experimental treatment, the treatment itself existed and was well-defined, and there is usually an unaffected comparison group that we believe is equivalent to the experimental group in all or most respects.

Laboratory experiments and field experiments

Experiments are also divided into **field experiments** and **laboratory experiments**. This refers to the experimental setting in which the study is being conducted.

As the name suggests, laboratory experiments involve inviting participants to a specially designated venue where the study is conducted according to a standard script, in well-controlled conditions. A laboratory experiment has the following key characteristics:

- Potential confounding variables are controlled as much as possible.
- Experimental conditions are standardized, so that each participant goes through exactly the same procedure in the same circumstances.
- As a rule, participants know that they are taking part in a study.

Since laboratory experiments can be well controlled, they come with the advantage of higher internal validity. However, the cost of this is usually lower ecological validity. Well-controlled environments are also artificial, and the behaviour of people in an artificial situation is not necessarily the same as it would be in similar conditions in real life.

By contrast, field experiments are conducted in a real-life setting. The researcher manipulates the IV, but since participants are in their natural setting many extraneous variables cannot be controlled. The strength of field experiments is higher ecological validity as compared to experiments in a laboratory. The limitation is less control over potentially confounding variables so there is lower internal validity. An example of a field experiment is Piliavin, Rodin, and Piliavin's (1969) subway study in which the researchers pretended to collapse on a subway train and observed if other passengers would come to help. To manipulate the IV, some researchers were carrying a cane (the cane condition) while others were carrying a bottle (the drunk condition).

Chat with AI

You can use generative AI to help you work out methodological characteristics of a research study. However, remember that there exists disagreement in the field regarding some labels. For example, scientists could disagree if it would be more appropriate to consider a particular study a quasi-experiment or a non-experiment.

Labels are not the point. What you need to analyse are the particular considerations that increase or compromize credibility and generalizability of the study in a given context and for a given purpose.

Provide your favourite generative AI with details of the study (or you might even be able to upload the full research paper) and ask various specific questions, such as:

- Could you help me identify methodological characteristics of this study that increase or decrease its credibility/validity?
- How would you characterize the researcher's ability to generalize from this study to a wider group such as all students in the world?
- What aspects of the study do you notice that I should keep in mind when deciding if this study allows cause–effect inferences?
- Do you think we can say that the two groups used in the study were equivalent?
- What could be the reasons against categorizing this study as a true experiment?

Carefully designed questions and prompts will allow you to interact with AI in a way that will deepen your understanding of research methodology.

1.5 Correlational studies

Inquiry questions

- What does it mean for two variables to correlate with each other?
- What should be avoided when interpreting correlations?
- Can two correlating variables in fact be unrelated?
- Can correlations show curvilinear relationships?

What you will learn in this section

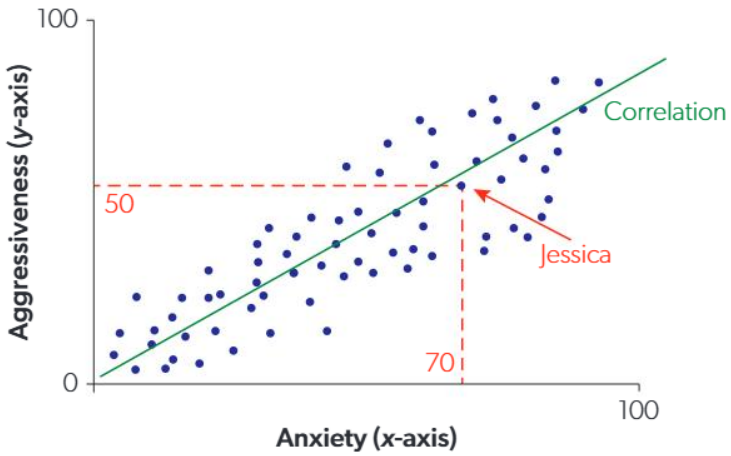
Key learning:

- What is a correlation? Effect size and statistical significance.
- Limitations of correlational studies: causation cannot be inferred, the third variable problem, curvilinear relationships, and spurious correlations.
- Sampling and generalizability in correlational studies.
- Credibility and bias in correlational studies.

Key terms: correlational studies, correlation coefficient, effect size, statistical significance, third variable problem, curvilinear relationships, spurious correlation

What is a correlation?

Correlational studies are different from experiments in that no variable is manipulated by the researcher, so causation cannot be inferred. Two or more variables are measured and the relationship between them is mathematically quantified.



▲ Figure 1.13 Scatterplot of anxiety and aggressiveness

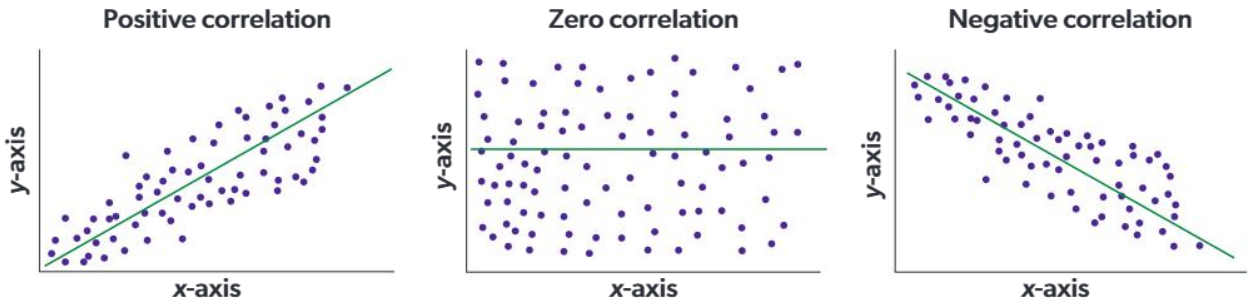
The way it is done can be illustrated graphically through scatterplots. Suppose you are interested in investigating if there is a relationship between anxiety and aggressiveness in a group of students. For this you recruit a sample of students and measure anxiety with a self-report questionnaire and aggressiveness through observation during breaks. You get two scores for each participant: anxiety and aggressiveness. Suppose both scores can take values from 0 to 100. The whole sample can be graphically represented with a scatterplot. See Figure 1.13.

Each dot on the scatterplot represents one person. The coordinates of each dot give you the scores obtained for each of the variables. For example, Jessica's score on anxiety is 70 (the *x*-axis coordinate) and her score on aggressiveness is 50 (the *y*-axis coordinate). The whole scatterplot looks like a "cloud" of

participants in the two-dimensional space of the two variables. A correlation is a measure of linear relationship between two variables. Graphically, a correlation is a straight line that best approximates this "cloud" in the scatterplot.

In Figure 1.13, the correlation is positive because there is a tendency: as X increases, Y increases. So if an individual got a high score on variable X, that person probably also got a high score on variable Y. This is where the name "correlation" comes from: the two variables "co-relate". Remember: correlation does not imply causation: we cannot say that X influences Y, nor can we say that Y influences X. All we know is that there is a link between them.

A **correlation coefficient** can vary from −1 to +1. The scatterplots in Figure 1.14 demonstrate some examples.



▲ Figure 1.14 Examples of correlations

A positive correlation demonstrates the tendency for one variable to increase as the other variable increases. A negative correlation demonstrates the inverse tendency: when one variable increases the other variable decreases. The steeper the line, the stronger the relationship. A perfect correlation of 1 (or −1) is a straight line with a slope of 45°: as one variable increases by one unit, the other variable increases (or decreases) by exactly one unit. A correlation close to zero is a flat line. It shows that there is no relationship between the two variables: the fact that a person scored high or low on variable X tells us nothing about their score on variable Y. Graphically, such scatterplots are more like a circle or a rectangle.

Effect size and statistical significance

The absolute value of the correlation coefficient (the number from −1 to 1) is called the **effect size**. How do you know if a correlation is small or large? There are widely accepted guidelines based on Cohen's (1988) suggestions to interpret the effect size of correlations in social sciences—see Table 1.5.

Correlation coefficient effect size (*r*)	Interpretation
Less than 0.10	Negligible
0.10–0.29	Small
0.30–0.49	Medium
0.50 and larger	Large

▲ Table 1.5 Effect sizes for correlation coefficients

The effect size is not the only parameter that is important when interpreting a correlation coefficient. Another parameter is the level of **statistical significance**. Statistical significance shows the likelihood that a correlation of this size has been obtained by chance. In other words, what is the probability that you will replicate the study with a different sample and the correlation will turn to zero? It depends on the sample size: with small samples you cannot be sure that an obtained correlation, even if it is relatively large, has not been obtained due to random chance. With large samples, correlation estimates are more reliable and you can be more confident that the correlation is not a product of random chance but a genuine reflection of a relationship between the two variables in the population. The probability that a correlation has been obtained due to random chance can be estimated. Again, there are conventional cut-off points when results are considered to be "statistically significant" or not—see Table 1.6.

The probability that the result is due to random chance	Notation	Interpretation
More than 5%	p = n.s.	Result is non-significant
Less than 5%	$p < 0.05$	Result is statistically significant (reliably different from zero)
Less than 1%	$p < 0.01$	Result is very significant
Less than 0.1%	$p < 0.001$	Result is highly significant

▲ Table 1.6 Thresholds of statistical significance

The conventional cut-off point for statistical significance is 5%. Whatever result you obtained, if the probability that this result is pure chance is <5%, we assume that the result is statistically significant, reliably different from zero and so would be replicated in at least 95 out of 100 independent samples drawn from the same population.

When interpreting correlations one needs to consider both the effect size and the level of statistical significance. If a correlation is statistically significant, it does not mean that it is large, because in large samples even small correlations can be significant (reliably different from zero). Therefore, scientists are looking for statistically significant correlations with large effect sizes.

Activity

Correlations are denoted by the letter r. Below are some examples of results of fictitious correlational studies. See if you can interpret them using your knowledge of Cohen's guidelines and levels of statistical significance:

1. $r = 0.14$, p = n.s.
2. $r = 0.10$, $p < 0.05$
3. $r = 0.34$, $p < 0.01$
4. $r = 0.61$, $p < 0.001$

ATL Research, Self-management

TOK

As you see, the nature of knowledge in psychology, just like the other social sciences, is probabilistic. We only know something with a degree of certainty and there is a possibility this knowledge is a product of chance.

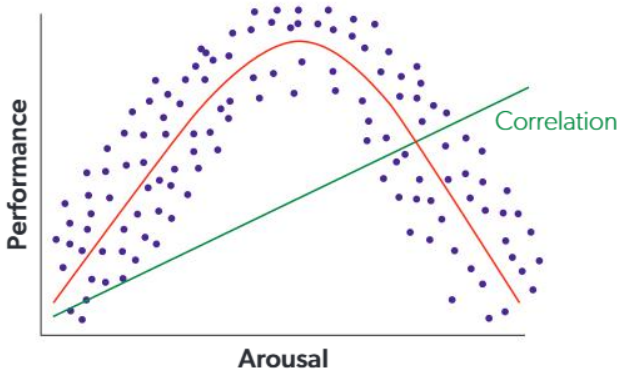
How does that compare to the nature of knowledge in other areas such as natural sciences (physics, chemistry, biology), history, or the arts?

What can we do to increase the degree of certainty in social sciences (for example, replication of studies)?

Limitations of correlational studies

Correlational studies have several major limitations.

- As already mentioned, correlations cannot be interpreted in terms of causation.
- "The **third variable problem**". There is always a possibility that a third variable exists that correlates both with X and Y and explains the correlation between them. For example, cities with a larger number of spa salons also tend to have more criminals. Is there a correlation between the number of criminals and the number of spa salons? Yes, but once you consider the third variable, the size of the city, this correlation becomes meaningless.
- **Curvilinear relationships**. Sometimes variables are linked non-linearly. For example, a famous Yerkes-Dodson law in industrial psychology states that there is a relationship between mental arousal (stress) and performance: performance increases as mental arousal increases, but only up to a point. When levels of stress surpass that point, performance begins to decrease. Optimal performance is observed when levels of mental arousal are average. This can be seen in the scatterplot in Figure 1.15. However, this relationship can only be captured by looking at the graph. Since correlation coefficients are linear, the best they could do is to find a straight line that fits best to the scatterplot. So, if we were using correlational methods to find a relationship between mental arousal and performance, we would probably end up obtaining a small to medium correlation coefficient. Psychological reality is complex and there are a lot of potentially curvilinear relationships between variables. However, correlational methods reduce these relationships to linear, easily quantifiable patterns.



▲ Figure 1.15 Mental arousal and performance

- **Spurious correlations**. When a research study involves calculating multiple correlations between multiple variables, there is a possibility that some of the statistically significant correlations would be the result of random chance. Remember: a statistically significant correlation is the one that is different from zero with the probability of 95%. There is still a 5% chance that the correlation is an artefact, and the relationship actually does not exist in reality. When we calculate 100 correlations and only pick the ones that turned out to be significant, this increases the chance that we have picked spurious correlations.

Sampling and generalizability in correlational studies

Sampling strategies in correlational research are the same as in experiments. First the target population is identified depending on the aims of the study and then a sample is drawn from the population using random, stratified, opportunity, or self-selected sampling.

Generalizability of findings in correlational research is directly linked with sampling and depends on the representativeness of the sample. Again, this is much like population validity in experiments.

Credibility and bias in correlational research

Bias in correlational research can occur on the level of variable measurement and on the level of interpretation of findings.

On the level of measurement of variables, various biases may occur and they are not specific to correlational research. For example, if observation is used to measure one of the variables, the researcher needs to be aware of all the biases inherent in observation. If questionnaires are used to measure variables, biases inherent in questionnaires become an issue. The list goes on.

On the level of interpretation of findings, the following considerations represent potential sources of bias.

- Curvilinear relationships between variables (see previous page). If this is suspected, researchers should generate and study scatterplots.
- "The third variable problem". Correlational research is more credible if the researcher considers potential "third variables" in advance and includes them in the research in order to explicitly study the links between X and Y and this third variable.
- Spurious correlations. To increase credibility, results of multiple comparisons should be interpreted with caution. Effect sizes need to be considered together with the level of statistical significance.

Discussion

Compare and contrast sampling, generalizability, credibility, and bias in correlational research with those in experimental research.

- In what aspects are the approaches different?
- In what aspects are they the same?
- Are there any aspects where the ideas are similar but the terminology differs?

ATL Research, Thinking

1.6 Qualitative research

Inquiry questions

- To what extent can findings from qualitative research be generalized?
- How can credibility of qualitative research studies be ensured?
- What are the differences and similarities in how qualitative and quantitative research approaches sampling, credibility, generalizability, and bias?

What you will learn in this section

Key learning:

- Credibility (trustworthiness) in qualitative research: triangulation, rapport, iterative questioning, reflexivity, credibility checks, "thick descriptions".
- Bias in qualitative research: participant bias and researcher bias.
- Sampling in qualitative research: quota sampling, purposive sampling, theoretical sampling, snowball sampling, convenience sampling.
- Generalizability in qualitative research: sample-to-population generalization, theoretical generalization, case-to-case generalization = transferability.
- Qualitative methods: observation, interviews and focus groups, case studies.

Key terms: trustworthiness, triangulation, establishing a rapport, iterative questioning, reflexivity, credibility checks, "thick descriptions", participant bias, acquiescence bias, social desirability bias, dominant respondent bias, sensitivity bias, researcher bias, confirmation bias, leading questions bias, question order bias, sampling bias, biased reporting, quota sampling, purposive sampling, theoretical sampling, snowball sampling, convenience sampling, sample-to-population generalization, theoretical generalization, case-to-case generalization, transferability, laboratory observation, naturalistic observation, overt observation, covert observation, participant observation, structured observation, unstructured observation, structured interview, semi-structured interview, unstructured interview, focus group, case study

Credibility in qualitative research

Credibility in qualitative research (also known as **trustworthiness**) is a counterpart of internal validity in the experimental method. As you have seen, internal validity is a measure of the extent to which the experiment tests what it is intended to test. We believe that the experiment is internally valid if we have controlled potential confounding variables and we are confident that it is the manipulation of IV, and not anything else, that causes a change in DV.

In a similar fashion, credibility in qualitative research is related to the question, "To what extent do the findings reflect reality?" If a true picture of the phenomenon under study is being presented, the study is credible.

The term "trustworthiness" is also used to denote credibility in qualitative research.

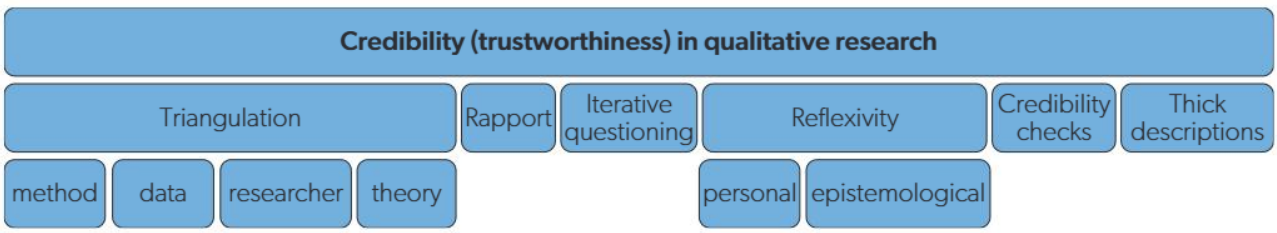
TOK

How do we know if the picture of a phenomenon presented in the findings from a qualitative study is "true"? If we had a way to know that, we wouldn't need a research study in the first place.

One of the popular definitions of knowledge is "justified true belief". A similar problem, however, arises with this definition: other than through "knowledge", we do not have a way of establishing if something is true. So, knowledge depends on truth but truth is a result of knowledge.

To solve this paradox, it has been suggested that we substitute "true" in this definition to "beyond reasonable doubt". So, to ensure that a qualitative research study is credible we need to demonstrate that its findings are "true beyond reasonable doubt".

How do you understand that? What do you think is "reasonable doubt" in this context?



▲ Figure 1.16 Some common methods to ensure credibility in qualitative research

To ensure that what is presented in the findings of a qualitative study is true, several types of measures can be taken.

Triangulation. This refers to a combination of different approaches to collecting and interpreting data. There are several types of triangulation:

- Method triangulation. If the same results are obtained using various methods (e.g., interviews and observations), credibility increases.
- Data triangulation. This refers to using data from a variety of sources. For example, if participants during an interview refer to certain documents, these documents may be studied in order to gain a clearer understanding of the participants' experiences.
- Researcher triangulation. As follows from the name, this refers to combining the observations/interpretations of different researchers. Undoubtedly, if two people see the same thing, this increases the credibility of their findings.
- Theory triangulation. This refers to using multiple perspectives or theories to interpret the data.

Establishing a rapport. Researchers should ensure that participants are being honest. It should be made clear to participants that there are no right or wrong answers. In general, a good rapport should be established with participants. They must alter their behaviour in the presence of the researcher as little as possible.

Iterative questioning. In many research projects, especially those involving sensitive data, there is a risk that participants will distort data either intentionally (lying) or unintentionally to try to create a certain impression on the researcher. Spotting ambiguous answers and returning to the topic later using a rephrased question might help researchers to gain a deeper insight into the sensitive phenomenon.

Reflexivity. Researchers should reflect on the possibility that their own biases might have interfered with the observations or interpretations. Arguably, due to the nature of qualitative research, a certain degree of subjectivity is unavoidable. However, researchers need to be able to identify the findings that might have been affected the most. There are two types of reflexivity:

1. Epistemological reflexivity—linked to knowledge of the strengths and limitations of the method used to collect data. Epistemology is the study of knowledge. For instance, "the following behaviours were observed … however, they should be interpreted with caution because participants were aware that they were being observed and hence might have modified their behaviour".
2. Personal reflexivity—linked to the personal beliefs and expectations of the researcher. For example, "I noticed that overcoming trauma was particularly emphasized in their conversations; however, since I myself have a history of overcoming childhood trauma, this observation could have been influenced by my personal beliefs and should be cross-checked by an independent interviewer".

Credibility checks. This refers to checking the accuracy of data by asking participants themselves to read transcripts of interviews or field notes of observations. Participants should confirm that the transcripts or notes are an accurate representation of what they said (meant) or did. This is often used in interviews, with the interviewees receiving the transcripts or notes and being asked to correct any inaccuracies or provide clarifications.

"Thick descriptions". This refers to explaining not just the observed behaviour itself, but also the context in which it occurred so that the description becomes meaningful to an outsider who never observed the phenomenon first-hand. It involves describing the phenomenon in sufficient detail that it can be understood holistically and in context. For example, imagine a stranger smiled at you. This behaviour out of context can be reported "thinly", just stating the fact, or it can be placed in a context (who, where, in what circumstances), making it meaningful. Thick descriptions are also referred to as "rich" descriptions—these terms are interchangeable.

Discussion

To what extent is ensuring credibility in qualitative research similar to the way internal validity is ensured in experimental research? What are the differences?

ATL Thinking, Research

Bias in qualitative research

In quantitative research, we deal with potential bias by trying to eliminate it completely or by keeping the potentially confounding variables constant in all comparison groups. In qualitative research this approach is not possible, and bias is actually an integral part of the research process because the researcher is a tool through which data is collected. Therefore, while some types of bias may be avoided, other types of bias are inevitable and need to be reflected on and accounted for.

Qualitative research may be associated both with **researcher** and **participant bias**. Let us now look at the major sources of bias.

Participant bias

- **Acquiescence bias** is a tendency to give positive answers whatever the question. Some people are acquiescent by nature, and in some others, acquiescence may be induced by the nature of the questions or the researcher's behaviour. To avoid this bias, researchers should be careful not to ask leading questions, making their questions open-ended, neutral, and focused on the opinions of the participant.
- **Social desirability bias** is a participant's tendency to respond or behave in a way that they think will make them liked or accepted. This may be done intentionally or unintentionally. Research into sensitive topics is especially vulnerable to social desirability. To reduce this bias, questions should be phrased in a non-judgemental way that suggests that any answer is acceptable. Another method that researchers use is to ask questions about a third person (e.g., "What do your friends think about … ?"). This helps participants to disengage from the sensitive topics and provide more honest answers.
- **Dominant respondent bias** occurs in a group interview setting when one of the participants influences the behaviour and responses of the others. Dominant respondents may "hijack" talking time or intimidate others by demonstrating their assertiveness or superior knowledge of the subject. Researchers should be trained to control dominant respondents and make sure that all participants are provided with equal opportunities to speak.
- **Sensitivity bias** is a tendency of participants to answer regular questions honestly, but distort their responses to questions on sensitive subjects. They may even give incorrect information to hide secrets. The solution to this problem is to build a good rapport with each participant and create trust between the participant and the researcher. To build trust, the researcher needs to increase the sensitivity of the questions gradually while being responsive to the participant's concerns.

Researcher bias

- **Confirmation bias** occurs when the researcher has a prior belief and uses the research in an unintentional attempt to confirm that belief. Confirmation bias may influence the way questions are worded, the small nuances in the researcher's non-verbal behaviour, and selectivity of attention while observing behaviour or interpreting the data. Reflexivity is the solution to confirmation bias. Confirmation bias is such a deeply grounded error in human information processing that it is largely unavoidable in qualitative research, in which data can only be collected "through" a human observer. Therefore, rather than avoiding it, researchers should be trained to recognize it and take it into account.
- **Leading questions bias** occurs when respondents in an interview are inclined to answer in a certain way because the wording of the question encourages them to do so. Even if an interview is carefully planned in advance, researchers often ask additional follow-up or clarification questions, and these may potentially cause distortions in the responses. Interviewers should be rigorously trained in asking open-ended, neutral questions that do not suggest a particular answer.
- **Question order bias** occurs when responses to one question influence the participant's responses to the following questions. This bias stems from the

human tendency to be consistent in our beliefs and actions. For example, if the first question in the interview asked whether you liked sports and you hesitated but said "yes", you would probably be inclined later to give more positive answers about your attitudes to gym membership. To minimize this bias, general questions should be asked before more specific ones, positive questions before negative ones, and behaviour questions before attitude questions.

- **Sampling bias** occurs when the sample is not adequate for the aims of the research. For example, the selection of people who are not "the best fit" in terms of the research purposes may be the result of convenience sampling.
- **Biased reporting** occurs when some findings of the study are not equally represented in the research report. For example, the researcher might choose to only briefly mention pieces of evidence that do not "fit". Reflexivity, integrity, and training of researchers are the means to counteract biased reporting.

In summary, some types of bias in qualitative research may be eliminated, while some others need to be recognized and considered. Reflexivity and triangulation are the two most important instruments that allow the researcher to reduce the influence of bias in qualitative research.

The presence of biases is directly linked to both credibility and generalizability of research findings.

Activity

Some of the sources of bias in experimental and qualitative research appear in Table 1.7. See if you can find any overlaps and then discuss these in class.

Experimental research	Qualitative research
• Selection • Maturation • Testing effect • Experimental mortality • Experimenter bias • Demand characteristics	• Acquiescence bias • Social desirability bias • Dominant respondent bias • Sensitivity bias • Confirmation bias • Leading questions bias • Question order bias • Sampling bias

▲ Table 1.7 Sources of bias in experimental and qualitative research

ATL Thinking, Communication

Sampling and generalizability in qualitative research

Generalization is a broad inference from particular observations. It is "an inference about the unobserved based on the observed" (Polit and Beck, 2010).

Traditionally generalizability has been the focus of debate between supporters of quantitative and qualitative methods. The main argument against generalizability in qualitative research is that samples are not statistically representative of the target population. As you know, representativeness in quantitative research is a necessary requirement for findings to be applied beyond the sample to the target population it represents.

A "weak" counter-argument is that qualitative methods do not aim to apply research findings to a wider population. In other words, the purpose of qualitative methods is the study of a particular sample but not the population it "represents". However, some scholars make a stronger argument and claim that generalizability to other groups of people is in fact achievable, to a certain extent, in qualitative research.

Some other scholars even argue that qualitative research is in fact more generalizable. They claim that rich data obtained in qualitative studies allows us to gain a deeper understanding of the phenomenon and so make more accurate inferences about its nature.

Sampling

In quantitative research, representativeness of the sample (and therefore the ability to generalize results to a wider population) is ensured through random sampling. In random sampling each member of the target population has an equal chance of being included in the sample. In other words, random sampling is probabilistic. However, sampling in qualitative research is non-probabilistic. These are the most commonly used types of sampling in qualitative research.

- **Quota sampling.** In quota sampling, it is decided prior to the start of research how many people to include in the sample and which characteristics they should have. This decision is driven by the research question. Using various recruitment strategies, researchers then recruit participants until the quotas are met. Quota sampling is similar to stratified sampling in quantitative research in that both the important participant characteristics and the necessary sample proportions are pre-defined.
- **Purposive sampling.** This is similar to quota sampling in the sense that the main characteristics of participants are defined in advance. Researchers then recruit participants who have these characteristics. However, the proportions and the sample size are not defined.
- **Theoretical sampling.** This is a special type of purposive sampling that stops when the point of data saturation is reached. Data saturation means that no new information is obtained from new participants added to the sample. Whether information is "new" or not is defined on the basis of a background theory: if no new evidence (or counterevidence) for the claims of the theory emerges, data saturation is reached. Generalization in this case is made from the data to the theory.
- **Snowball sampling.** In this approach, a small number of participants are invited and asked to invite other people they know who also are of interest for the purposes of the research. This approach is mostly used in pilot research studies (when there are insufficient resources to carefully select participants) or in research with groups of people who are very difficult to reach (e.g., youth gang members).

- **Convenience sampling.** The most superficial approach in which you just use the sample that is easily available or accessible (e.g., professors might conduct research with university students simply because it is time- and cost-efficient).

Types of generalizability

Firestone (1993) distinguished between three types of generalizability that provide a convenient framework for comparing quantitative and qualitative studies.

1. **Sample-to-population generalization.** The researcher starts by identifying the target population and then selecting a sample that is representative of this population. The best approach to achieve this is to use random sampling. The concept that is used to describe sample-to-population generalizability in experiments is "population validity" (part of "external validity"). Due to the non-probabilistic nature of samples in qualitative research, this type of generalization is difficult.
2. **Theoretical generalization.** Generalization is made from particular observations to a broader theory. In quantitative research theoretical generalization takes the form of construct validity: it is the leap we make from directly observable operationalizations to the unobservable construct. In qualitative research theoretical generalization is achieved through rigorous analysis and interpretation of research findings. We can generalize to a wider theory if data saturation was achieved, thick descriptions provided, analysis was in-depth and free of biases, and so on. Theory plays a much greater role in qualitative research than in quantitative.
3. **Case-to-case generalization,** also known as **transferability.** Generalization is made to a different group of people or a different setting or context. In qualitative research, transferability is the responsibility of both the researcher and the reader of the research report. The researcher's responsibility is to ensure that thick descriptions are provided so that the reader has sufficient information and details about the context of the study. The reader's responsibility is to decide whether or not the context described in the report is similar to a new situation (Polit and Beck, 2010). A rough and pretty distant equivalent of transferability in quantitative research would probably be "ecological validity" (another part of "external validity"). Refer to Table 1.2 on page 15.

Activity

Compare the sampling techniques used in experiments and in qualitative research studies. Use any kind of visual representation to demonstrate the results of this comparison and present it in class.

How are the three types of generalizability approached in experiments and qualitative research studies? Which of these do you think are better achieved in qualitative research as compared to experimental research?

ATL Research, Communication

Qualitative methods: observation

There are several common reasons for choosing the method of observation.

- The focus of the research is on how people act and interact in a natural setting. Most other research methods are artificial in the sense that they place the participant in a specially constructed research context.
- The researcher believes that meaningful knowledge cannot be generated without observation—for example, because it cannot be articulated. Suppose you want to gain an insight into the behaviour of your classmates during a fire drill at your school. It would probably be more meaningful to observe an actual fire drill than to conduct an interview and analyse verbal responses.

- Observation allows the researcher to become deeply immersed in the studied phenomenon, sometimes even becoming part of it. Arguably, this is a strength because they can gain almost first-hand experiences.

Observation is "experiential" and the researcher is strongly involved in the process of data generation. All generated data is the product of the researcher's selective attention and interpretations. This makes reflexivity especially important.

There are several types of observation, and the particular type chosen will have broad implications in terms of credibility, reflexivity, generalizability, and ethics.

- **Laboratory** versus **naturalistic observation.** Naturalistic observation is carried out in naturally occurring settings—that is, a place that has not been arranged for the purposes of the study. Sometimes naturalistic observation would be the only choice. For example, if you wanted to study inter-group discrimination and violence, it would be unethical to encourage violence in a research setting. A drawback is that it may be time-consuming because the behaviour of interest may only occur at certain times or infrequently.
- Observation may be overt or covert. **Overt observation** occurs when participants are aware of the fact that they are being observed. The ethics of this approach are clearly a strength as participants give their informed consent, but there are methodological limitations—biases related to participant expectations. In contrast, in **covert observation** the researcher does not inform the members of the group about the reasons for their presence. An advantage of covert observation is gaining access to groups that would not normally agree to participate in research (e.g., socially isolated or violent groups). Another strength is the avoidance of participant bias—subjects do not know they are being observed, so they behave naturally. The ethics here are a disadvantage. Participants do not give their consent to take part in the study. One way to avoid this issue is to debrief participants after the observation session and ask for their consent prior to using the data for research purposes.
- **Participant** and **non-participant observation.** In participant observation, the observer becomes part of the observed group. For example, many anthropologists spend time living among members of an indigenous society in order to study their culture "from the inside". The advantage of participant observations is that they allow the researcher to gain first-hand experiences with the phenomenon of interest, gaining valuable insights. However, the drawbacks include the risk that the observer will lose objectivity as they become more involved with the group of studied individuals. This may happen because the researcher begins to identify themselves with the group.
- **Structured** versus **unstructured observation.** In structured observation, information is recorded systematically and in a standardized way. For example, structured observation may be conducted with a checklist of behaviours of interest where the observer is required to note the occurrence of these specific behaviours in pre-defined time intervals. Unstructured observations do not have a pre-defined structure and observers simply register whatever behaviours they find noteworthy.

Activity

Suppose your aim is to study ways in which sports teams build up and maintain their team spirit, and observation is your method. What type of observation would you use and why?

Describe how you would set up your research procedure both in terms of preparation and the actual observation process.

ATL Research, Thinking, Self-management

Qualitative methods: interviews and focus groups

In-depth interviews are one of the most popular qualitative research methods. An interview may be the only way to get an insight into the nature of subjective experiences and interpretations. Since attitudes, values, patterns of interpretation, and other subjective phenomena are unobservable, the most straightforward way to study them is to rely on the participants' verbal reports.

Interviews are a very personal form of research because there is direct contact between the interviewer and the interviewee. At the same time, interviews can, and often do, touch upon sensitive topics, such as coping with a terminal illness, experiencing phobias, or daily routines related to internet addiction.

Interviewing techniques are driven by the goal of learning as much as possible about the interviewee's opinions and experiences. The interviewer is the main research instrument. Tiny nuances in verbal and non-verbal behaviour of the interviewer may affect the interviewee's responses. For example, it is common in everyday conversations to ask leading questions, but interviewers must avoid doing it. This is why interviewers receive intensive training.

Interview data comes in the form of an audio or video recording which is subsequently converted to an interview transcript. Transcripts are later coded and analysed in line with the aims of the research.

There are three types of interview, depending on how fixed the list and sequence of the questions is.

1. **Structured interviews** include a fixed list of questions that need to be asked in a set order. It is most useful when the research project involves multiple interviewers and it is essential that they all conduct the sessions in a similar way.
2. **Semi-structured interviews** do not specify an order or a particular set of questions. They are somewhat like a checklist: the researcher knows that certain questions must be asked, but beyond that they can ask follow-up questions to get clarifications. If it better fits the natural flow of the conversation, the researcher can change the question order.
3. **Unstructured interviews** are mostly participant-driven, and every next question is determined by the interviewee's answer to the previous one. Of course, the researcher still has to keep in mind the overall purpose of the research and stay focused on exploring a particular topic. However, two different interviewees may end up getting very different questions.

The **focus group** is a special type of semi-structured interview that is conducted simultaneously with a group of 6–10 people. The key factor is that participants are encouraged to interact with each other and the interviewer to serve as a facilitator. The focus group facilitator can observe group dynamics and make use of it by directing group members' interaction so that they stay focused on the research topic.

The advantages of a focus group include the following:

- It is a quick way to get information from several participants at the same time.
- It creates a more natural and comfortable environment than a face-to-face interview, ensuring less participant bias.

Discussion

Suppose you are interested in studying the reasons why teenagers join criminal groups. You used snowball sampling techniques to recruit 10 participants. Would you use a structured, semi-structured, or unstructured interview? Why? Give reasons for your answer.

What do you think are the factors that need to be considered in conducting an interview with teenage gang members?

ATL Thinking, Communication

- It is easier to respond to sensitive questions when you are in a group.
- Multiple perspectives are discussed so a more holistic understanding of the topic is achieved.

However, there are several "new" limitations that come as a cost for including group dynamics in the research process.

- If one of the participants is especially dominant, this may distort the responses of the other participants (e.g., if they feel a need to conform).
- It is more difficult to preserve anonymity and confidentiality.
- Focus groups are especially demanding in terms of sampling and creating interview transcripts.

Qualitative methods: case studies

A **case study** is an in-depth investigation of an individual or a group. You might say that this is not a proper definition because other research methods can also be defined this way, and you would be right. In fact, case studies can involve a variety of other methods (observations, interviews, and so on), anything that deepens our understanding of an individual or a group of interest. There are several reasons why case studies are referred to as a separate research method, even though they are actually a combination of other methods.

- The individual or group that is the object of a case study is unique in some way. As a result, the purpose is to gain a deep understanding of this particular individual or group.
- Sampling is not an issue: you are interested in a particular case, not the population the case "represents".
- There is less focus on generalizability. Findings do get generalized, but this is a by-product of the in-depth description and explanation of the case (case-to-case and theoretical generalization).
- The case is studied thoroughly, using a combination of different methods, and often longitudinally. This is why we defined a case study as an "in-depth investigation".

What are the reasons for choosing a case study as the preferred method?

First, case studies are useful to investigate phenomena that could not be studied otherwise. For example, it is a group that is hard to get access to and you may only get a chance to study one individual (think about studying the personality of a patient who has lost one of their brain hemispheres as a result of surgery).

Second, case studies can contradict established theories and help develop new theories. Why is this a good thing? For example, universal theories of memory in cognitive psychology can be tested by studying individuals with unusual or unique memory abilities. If in these individuals memory proves to function differently, then the universal theory of memory is not as universal as we thought. So, "boundary" cases are interesting, and since they are quite rare, we want to study them thoroughly.

1.7 Surveys and questionnaires

Inquiry questions

- How do questionnaires "work"?
- How can we establish that results obtained with a survey or a questionnaire are valid and reliable?
- What are the main parameters of quality of measurement in a questionnaire?

What you will learn in this section

Key learning:

- What are surveys and questionnaires?
- How are questionnaires created?
- Elements of psychometrics.
- How are surveys different from questionnaires?
- What are the common biases in surveys and questionnaires?
- How do we construct good survey/questionnaire items?

Key terms: questionnaire, survey, psychometrics, reliability of a questionnaire, validity of a questionnaire, quantitative descriptive study, social desirability bias, acquiescence bias, extreme responding

Exam tip

Surveys and questionnaires are different, but not all academics interpret these differences in the same way. What some researchers would call a survey others call a questionnaire, and vice versa.

It is advisable that you develop your own understanding of this difference and clearly delineate between the two. Do not be afraid to get it wrong: it is stated in the IB Guide that for exam purposes the two terms are used interchangeably.

What are surveys and questionnaires?

There are different approaches to defining **surveys** and **questionnaires**, and to describing the difference between them. To avoid confusion and to prevent students from losing marks for bureaucratic reasons, it was decided in the IB Psychology course to use these terms interchangeably. However, surveys and questionnaires are indeed different. Here we will suggest one approach to defining them, without claiming that this is the only possible approach.

The key difference between a survey and a questionnaire is what is considered a construct and what is considered an operationalization. A questionnaire measures an underlying construct that is not explicitly present in the questions. A survey makes no distinction between constructs and operationalization: participants' responses are taken at face value.

Let us start with a closer look at questionnaires and how they are created. This will allow us to clearly bring out the contrast between questionnaires and surveys later.

How are questionnaires created?

To better understand how questionnaires work, let us consider an example. Suppose you are creating a questionnaire to measure friendliness. You have a

theory that friendliness consists of three components: (1) friendly attitude, (2) interpersonal trust, and (3) readiness to help. None of these three phenomena can be directly observed.

What we do instead is treat friendliness and its three components as underlying constructs and write questions that serve as specific observable indicators of these constructs. For example, we may write the following statements for participants to indicate their agreement with:

Scale 1—"Friendly attitude":

1. On weekends I try to go out and meet new people.
2. It is easy for me to start a conversation with a stranger.
3. There are a few people I know that I cannot stand (scored in reverse).
4. People around me think of me as an approachable person.

Scale 2—"Interpersonal trust"

1. I trust that my friends will keep my secrets.
2. In the past I let my friends borrow valuable things or money from me.
3. In important matters I always ask for a written confirmation (scored in reverse).
4. I could trust my friend with the password to my social media if I needed to.

Scale 3—"Readiness to help"

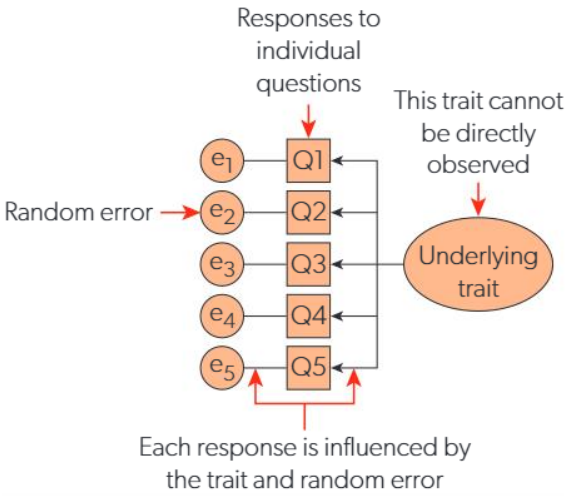
1. If someone I know needs something, I am usually among the first people to offer help.
2. People frequently abuse kindness (scored in reverse).
3. I donated to a charity at least once in the last year.
4. People who ask for help on the street are very frequently scammers (scored in reverse).

Suppose we have used the Likert scale from 1 to 5 to record participants' answers to each question:

- 5 = Absolutely agree
- 4 = Somewhat agree
- 3 = Neither agree nor disagree
- 2 = Somewhat disagree
- 1 = Absolutely disagree

Suppose that the three constructs that we hypothesized (friendly attitude, interpersonal trust, and readiness to help) actually exist. This means that there exists a certain stable trait that influences the behaviour of a person in a variety of situations. Although we cannot observe this trait directly, we can observe the behaviour of the person in various situations and make an inference about the trait.

Each questionnaire item is one such observation. The idea is that, if the person has that trait, they will respond to each of the questions in a way that is predicted by the trait. For example, if "friendly attitude" really exists as a trait of personality, and if a person really has this trait, then they should agree with items 1, 2, and 4 and disagree with item 3 on the friendly attitude list.



▲ Figure 1.17 Measurement of underlying traits in a questionnaire

Actually, the person's response to each question is determined by two factors: their true score (i.e., the true level of the trait that is being measured) and random error. For example, take the first item in the first scale: "On weekends I try to go out and meet new people". The person's response to this question will depend on their level of friendly attitude (true score). The higher it is, the more likely the person will be to say "absolutely agree". However, it is also influenced by many other factors, many of which can act in unpredictable ways. For example, the respondent may be a person who works on weekends, or generally seeks contact with people but gets so tired during the week that they prefer to rest at home on weekends. All these factors are believed to be "random error". Basing our measurement of the construct (trait) on one questionnaire item only would be unreliable.

However, when we use multiple indicators (questions) to assess the same construct (trait), we believe that random error cancels out and reliability of measurement is increased. If you have ever completed a personality questionnaire, you must have noticed (and been annoyed with) how multiple questions ask about similar things. Now you know why.

Exam tip

Paper 2 Section A is based on class practicals. Class practicals are opportunities throughout the course to conduct your own small-scale research investigation using different methods.

The IB recommends the following class practicals for each of the four contexts:

Context	Type of class practical	Minimum sample size
Health and well-being	Interview (structured, semi-structured, unstructured, or focus group)	One participant in an interview or three to eight participants in one focus group
Human development	Observation (naturalistic or controlled, overt or covert, participant or non-participant)	One participant
Human relationships	Survey/questionnaire (used interchangeably)	10 participants
Learning and cognition	Experiment (true, quasi-, or natural)	Five participants

▲ Table 1.8 IB recommended class practicals

Approach your class practical with the mindset of getting a deeper insight into the internal workings of a particular research method. In exam questions in Paper 2 Section A, you will need to make references to your class practicals—for example, describing the study and explaining its aim.

Psychometrics

A special discipline called **psychometrics** includes a range of mathematical techniques used to evaluate the validity and reliability of a questionnaire. A questionnaire is valid if it measures what it is supposed to measure (and not something else). A questionnaire is reliable if the measurement is consistent—for example, if the same questionnaire completed on several occasions yields the same results.

Here are some examples of metrics and indicators used in psychometrics to estimate the **reliability of a questionnaire**:

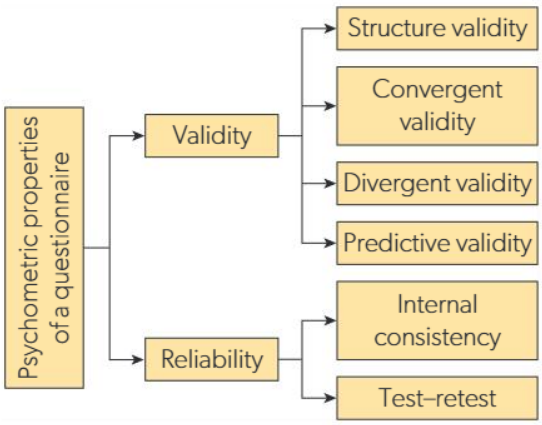
- One of the most basic characteristics of reliability of a questionnaire in psychometrics is known as internal consistency (of a scale). The most widely used measure of internal consistency is the so-called Cronbach's alpha. This metric depends on two parameters: the average correlation between items in a scale and the number of items. A scale with a large number of items in it where all items strongly correlate with each other is a reliable scale. Typically Cronbach's alpha of 0.7 is considered to be the acceptable cut-off, and Cronbach's alpha of 0.9 is considered to be very high.
- If the same sample of participants is given the same questionnaire again after a period of time, the results of the two assessments will not be the same. Two factors contribute to the discrepancy: the possibility that the trait itself has changed between the two measurements and the possibility that the scale is unreliable and yields different results each time you use it. Test–retest reliability is the correlation between results of the same questionnaire taken by a sample of participants on two different occasions. Typically the time period is taken to be six months as we have reasons to believe that relatively stable traits (such as personality qualities) are not likely to change within such a short time frame. There are acceptable thresholds in psychometrics below which the questionnaire is considered to be unreliable.

Here are some examples of metrics and indicators used in psychometrics to estimate the **validity of a questionnaire:**

- If the underlying traits really exist and it is these traits that the questionnaire actually measures, then we should observe the following pattern of correlations: (1) all items in one scale should be strongly or at least moderately interrelated; (2) at the same time, items from one scale should not be correlated with items from other scales. In other words, items in a questionnaire should comprise relatively independent groups of indicators. This is known as structure validity and is very important in questionnaires that include several scales. To estimate structure validity, researchers examine the matrix of correlations between each pair of items in a questionnaire. For this purpose, they also employ a statistical method known as factor analysis.
- If the questionnaire actually measures what it is supposed to measure, and not something else, then this should be reflected in the pattern of correlations between results of your questionnaire and other measures.

For example, it would be logical to expect that a measure of intelligence should be correlated with academic achievement and intellectual curiosity, but should not be correlated as much, say, with agreeableness and friendliness. When the questionnaire correlates with other measures that it is theoretically related to, this is called convergent validity. When it is not correlated with other measures that it is not supposed to be related to, this is called divergent validity.

- Another powerful measure of the psychometric quality of a questionnaire is predictive validity. It is the extent to which results of a questionnaire are predictive of some objectively registered parameter recorded in the future. For example, suppose you have created a questionnaire to measure students' inclination to cheat (academic dishonesty). If it actually measures what you say it does, then there must be a correlation between the results of your questionnaire and the number of academic dishonesty incidents recorded for a particular student in the next two years.



▲ Figure 1.18 Psychometric properties of a questionnaire

Exam tip

It is helpful to understand that there are various measures in psychometrics to estimate the validity and reliability of a questionnaire, but actually applying psychometrics is beyond the requirements of IB Psychology.

When you work on your class practical using a survey/questionnaire, it is advisable that you start with a research question, define what it is you want to measure, and create a survey or a questionnaire to measure it. Do not worry about numerically estimating the validity and reliability of a questionnaire. However, put effort into creating high-quality questions and interpreting the results of the study.

How are surveys different from questionnaires?

A survey is similar to a questionnaire in that it is also a data collection tool where measurement is based on responses of participants themselves. But the difference is that in a questionnaire you have an underlying construct (a scale) and

individual questions are indicators used to estimate the score on this construct, whereas in a survey questions represent measures in their own right and there is no underlying construct. For example, a survey about your use of social media may ask you questions like these:

1. On average how many hours per day do you spend on social media?
2. What are the social media platforms that you use most often?
3. How often do you post your own content on social media?

Responses to each of these questions will be taken as independent measures. There is no scale and no calculation of a total score per group of items.

Exam tip

The IB DP Psychology Guide also lists surveys/questionnaires as a separate research method, along with experiments, correlational studies, interviews, observations, and case studies.

We need to clearly distinguish between two scenarios:

1. When a survey or a questionnaire is used as a data collection tool as part of another research method. For example, you might conduct an experiment in which you manipulate an independent variable and measure the dependent variable with the help of a questionnaire.
2. When a survey/questionnaire is used as a research method in its own right. In this case, we are dealing with a **quantitative descriptive study**. This is different from experimental and correlational studies because we do not correlate any variables, we merely describe what percentage of participants provided specific answers to each of the questions.

What are the common biases in surveys and questionnaires?

Surveys and questionnaires are based on self-reporting, and this is probably the major weakness. The way people respond to questions does not always reflect the reality of their behaviour or experiences, especially when the topic is sensitive. Here are just a few of the well-known response biases that may affect the quality of data collected by surveys and questionnaires.

Social desirability bias. This is the tendency of respondents to respond in a way that they think is socially acceptable, rather than honestly. Social desirability is especially a problem when the topic of the research study is controversial—for example, excessive use of social media. Furthermore, different norms of what is "socially acceptable" exist in different social groups, which means that social desirability bias will affect the groups to varying extents and the validity of comparison between groups will be compromised.

There are a few measures that can be taken to prevent or minimize social desirability bias. Some examples include: using a forced-choice format, formulating neutral questions, asking a close person instead, and using a specially designed hidden social desirability scale.

Activity

Find out more about how the following techniques can help researchers deal with social desirability bias:

- Using a forced-choice response format
- Using a specially designed hidden social desirability scale.

Exchange what you have learned with a partner.

ATL Communication, Self-management



▲ Figure 1.19 Filling out a survey

Acquiescence bias is a participant's tendency to agree with all statements ("yes-saying"). It occurs because we are used to social situations in which our default response is endorsement or agreement.

A way to minimize the influence of this bias is to make the questionnaire balanced in terms of directly scored and inversely scored items. For example, the statement "I keep calm in most situations no matter how stressful" is a directly scored item to measure self control. Participants who say "yes" will gain a higher score. In contrast, the statement "I may lose control easily when something goes wrong" is an inversely scored item: participants gain a higher score if they say "no". If your scale has an approximately equal number of directly scored and inversely scored items, it will not prevent acquiescence bias, but will minimize its effects: participants who agree to every statement will not be able to gain the highest scores. Note: such a strategy is possible in a questionnaire (where an underlying construct is measured by several specific items) but not in a survey (where each item is taken as a measure in its own right).

Extreme responding is a bias that causes participants to select only the most extreme choices, such as "absolutely agree" and "absolutely disagree" on the Likert scale. This bias is especially detrimental to the validity of research in cross-cultural settings because people in some cultures are more likely to give extreme responses than others. The opposite of this bias—when participants only select neutral options and avoid providing extreme opinions—also exists. There is no foolproof way to counteract these biases except perhaps instructing participants directly that there are no right or wrong answers and asking them to try to use the entire scale. We should also consider background research on cultural specifics in response biases.

How to construct good survey/questionnaire items

Following some basic rules of formulating survey/questionnaire items may help avoid or prevent many potential response biases.

- Formulate statements in such a way that people with different traits, opinions, or beliefs will be likely to give different answers. Such statements are called "discriminative": they discriminate between people with a high true score on the trait and people with a low true score.
- Formulate statements in a way that is as unambiguous as possible, and refer to objective behaviour where applicable. For example, suppose you are constructing a questionnaire measuring anxiety. The statement "I am often anxious about things" is not great because: (1) different people will interpret "often" differently; (2) "anxious" refers to a mental state rather than behaviour, and people may differ in how they label various mental and emotional states; (3) "about things" is also very vague, leaving too much room for interpretation. Consider instead using a statement like this: "Last time I had to make a speech in public my palms were sweaty".
- Make sure that one item works towards measuring one aspect of one construct. This means avoiding complex sentences with conjunctions. For example, the item "I feel insulted and abandoned when I am excluded from my circle of friends" refers to feelings of insult and abandonment in one sentence. It would be better to simplify such statements or split them in two.
- Use clear and easy-to-understand language that will be accessible to respondents of all groups and educational levels. Make the statements as short as possible.

Activity

There are a variety of scales and response formats that can be used in surveys and questionnaires. Different formats are suitable for different purposes. Some examples include:

- Likert scale
- Forced-choice
- Multiple response
- Most–least
- Dichotomous (yes–no)
- Ranking
- Sentence completion.

Search online (or use your favourite generative AI) to find out more about these response formats. Discuss, for each of them, what purpose or research scenario would be the most suitable for choosing this response format over others.

ATL Research, Thinking, Self-management

1.8 Ethics in psychological research

Inquiry questions

- Since psychology is a study of living beings, what ethical issues does it raise?
- In psychology how can we decide what is ethical and what is not?

What you will learn in this section

Key learning:

- Ethical considerations in conducting a study include informed consent, protection from harm, anonymity and confidentiality, withdrawal from participation, deception, debriefing, cost–benefit analysis in ambiguous cases, and ethics committees.
- Ethical considerations in reporting results include anonymity, data fabrication, sharing research data for verification, handling of sensitive personal information, and social implications of reporting scientific results.

Key terms: informed consent, protection from harm, anonymity, confidentiality, withdrawal from participation, deception, debriefing, anonymity, data fabrication, sharing research data for verification, social implications of reporting scientific results

Overview

Ethics is an integral part of psychological research because it is research conducted with living beings (humans and animals). This is one of the things that distinguishes the human sciences from the natural sciences—ethically, the study of human beings is not the same as the study of material objects.

All around the world the activities of psychologists are regulated by codes of ethics. These codes outline the ethical principles and procedures to be followed in all aspects of a psychologist's professional activities: counselling, testing, and research. If a psychologist breaches the code, their professional licence may be discontinued. Codes of ethics have been developed by international as well as national psychological associations. There is a lot of overlap in their content as the ethical considerations in psychology are almost universal.

Since IB Psychology is an academic subject (involving no counselling), we will only focus on ethical considerations related to research. We will also break these ethical considerations into two large groups:

1. ethical considerations in conducting a study
2. ethical considerations in reporting results.

Activity

Explore the Code of Ethics on the website of the American Psychological Association (APA) and the Code of Human Research Ethics by the British Psychological Society (BPS).

Compare the two codes and make a poster for your classroom highlighting the main similarities and differences.

ATL Thinking, Social, Self-management

Ethical considerations in conducting a study

The following list outlines the main ethical considerations to be addressed when conducting a research study in psychology.

- **Informed consent.** Participation in a study must be voluntary, and participants must fully understand the nature of their involvement, including the aims of the study, what tasks they will be exposed to, and how the data will be used. Researchers should provide as much information as possible and in the clearest possible way, hence the name "informed" consent. If the participant is a young person (a minor), consent should be obtained from their parents or legal guardians.
- **Protection from harm.** At all times during the study, participants must be protected from physical and mental harm. This includes possible negative long-term consequences of participating in a research study.
- **Anonymity** and **confidentiality.** These two terms are often used interchangeably, but they refer to slightly different things. Participation in a research study is confidential if there is someone (e.g., the researcher) who can connect the results of the study to the identity of a particular participant, but the terms of the agreement prevent this person from sharing the data with anyone. Participation in a study is anonymous if no one can trace the results back to a participant's identity because no personal details have been provided. An example of anonymity would be filling out an online survey without providing your name.
- **Withdrawal from participation.** It must be made explicitly clear to participants that, since their participation is voluntary, they are free to withdraw from the study at any time. Researchers must not prevent participants from withdrawing or try to convince them to stay.
- **Deception.** In many cases, the true aims of the study cannot be revealed to the participants because it would change their behaviour (e.g., due to social desirability). So a degree of deception needs to be used. However, it must be kept to the necessary minimum.
- **Debriefing.** After the study, participants must be fully informed about its nature, its true aims, how the data will be used and stored. They must be given an opportunity to review their results and withdraw the data if they want to. If deception was used, it must be revealed. Care must be taken to protect participants from any possible harm including long-term effects such as recurring uncomfortable thoughts. In some cases, psychological help must be offered to monitor the psychological state of the participant for some time after the study (e.g., in sleep deprivation studies).

Frequently, ethical decisions prior to conducting a study are not easy, and a cost–benefit analysis needs to be conducted. For example, for their behaviour to be more natural in the study, sometimes participants should not know the true aim of the study. Studies of such phenomena as obedience, conformity, compliance, violence, and prejudice can rarely be designed so that they are harmless to the participants. So can we make the decision to relax some of the ethical standards for a particular study?

Activity

Ethical considerations in conducting a study may be combined in the following acrostic:

- can (consent)
- do (debriefing)
- cannot (confidentiality)
- do (deception)
- with (withdrawal)
- participants (protection from harm)

Try making such acrostics of your own with other lists in this chapter: threats to internal validity, types of bias in qualitative research, and so on.

ATL Self-management

Professional bodies of psychologists have ethics committees that resolve ambiguous issues and approve research proposals. Research proposals with a full description of the aims, procedures, and anticipated results are submitted to the committee and reviewed. In some cases, when research is potentially useful, ethically ambiguous research studies may get the "green light". Researchers will then need to be extra careful in making sure that participant harm is minimized and long-term follow-up after the study is provided. Failure to cooperate with an ethics committee is itself a violation of ethics.

Activity

The Little Albert experiment was carried out by Watson and Rayner (1920). They observed that a baby's fearful reaction to loud noises was an innate, automatic response. When they hear a loud noise, little children always display behavioural signs of fear (e.g., tears).

Their participant was a nine-month-old infant from a hospital who was referred to as "Albert". During the baseline test, Albert was exposed to a white rat, a rabbit, masks with hair, cotton, wool, and other objects. Albert showed no fear in response to these objects. During the experiment, a white laboratory rat was placed in front of Albert and he played with it. Every time the baby touched the rat, researchers hit a suspended steel bar behind his back with a hammer, producing a very loud sound. Naturally, the baby cried and showed fear. After pairing these two stimuli several times, the steel bar was taken away and Albert was only presented with the rat. In line with the theory of classical conditioning, Albert would show signs of distress, cry, and crawl away. So, the researcher "succeeded" in forming a fear of a rat in a baby.

In further trials, it was revealed that fear in Albert was actually generalized to other furry objects. He would show distress, cry, and crawl away at the sight of a rabbit, a furry dog, and even a Santa Claus mask with a beard. To make things worse, Albert left the hospital (taken away by his mother who did not leave any contact details) shortly after the experiment, and although Watson had planned to carry out desensitization, he never had the opportunity.

How would you go about conducting the study in a more ethically appropriate way?

What questions would you ask Watson and Rayner if you were sitting on the ethics committee reviewing their research proposal?

ATL Communication, Self-management

Ethical considerations in reporting results

In reporting the results of a research study through publication, the following ethical considerations should be addressed (among others).

- **Anonymity.** The identity of participants must be protected and there should be no way to identify them from the published results. This is especially sensitive in case studies of unique individuals (e.g., a feral child or a person with an unusual brain damage).
- **Data fabrication.** Publication of research results must be true to what was actually obtained in the study and presented in an unbiased way. If an error is found in already published results, measures should be taken to correct it.
- **Sharing research data for verification.** It is the responsibility of researchers to securely store primary data obtained in their research. Rules of good science dictate that any research study should be transparent for replication: procedures should be described in detail so that they are easily replicable by an independent researcher. Even without actual replication it should be possible to cross-check a researcher's analysis and conclusions based on the

same data set. Any request from an independent researcher to share raw data should be satisfied, provided both parties use the data ethically, responsibly, and within the limits of agreed-upon purposes.

- **Social implications of reporting scientific results.** Researchers must keep in mind how their conclusions are formulated in publications. They need to be mindful about the potential effects these conclusions may have on the scientific community and society in general. This is especially pertinent when results are published in popular journals that target a wider audience including non-scientists. Blunt conclusions like "intelligence is inherited" do not reflect the full complexity of the situation but can be taken at face value by non-scientists. If they become popular, such beliefs are resistant to change. They may affect public policies—for example, in the field of education.

Let us consider an example. There is much controversy about the work of Cyril Burt, a British psychologist who became famous for his contributions to intelligence testing. In his studies, he found that the IQ scores of identical twins reared apart were much more similar than those of non-identical twins reared together. He concluded that genetic inheritance in intelligence plays a much greater role than environmental factors (such as education).

Burt's research was very influential in forming educational policies in the country. For example, the belief that intelligence is fixed and hereditary led to the practice of using standardized tests to measure intelligence in school children and allocate them to schools based on the results.

After his death in 1971, the British Psychological Society found him guilty of publishing a series of fraudulent articles and fabricating data to support the theory that intelligence is inherited. The case was built on several details that were considered to be highly suspicious.

- There was a very unlikely coincidence of the same correlation coefficient between IQ of monozygotic twins (exactly 0.771) in two different studies.
- Some factors that should theoretically influence intelligence (such as mental illness or childhood influences) were suspiciously unimportant in Burt's data sets. This is almost a statistical impossibility.
- Identical twins reared apart is an extremely rare sample; there were only three other studies at that time using this kind of sample and none of them had more than 20 pairs of twins as participants. Burt's two studies involved 42 and 53 pairs of identical twins reared apart.
- Burt's two female collaborators who worked for him collecting and processing data could not be found. Their contact with Burt could not be traced and it was even suspected that these people never existed!

Data sets and publications that are not transparent and raise questions regarding their credibility are in themselves an ethical concern, even if they are not falsified intentionally. This is especially true for settings in which research findings are used to inform social policies (e.g., educational policies).



▲ Figure 1.20 Cyril Burt

1.9 Research with animals

Inquiry questions

- Can animal studies provide an insight into human behaviour?
- Is psychological experimentation with animals ethical?

What you will learn in this section

Key learning:

- What is an animal model?
- The value of animal research in psychology.
- The purposes of animal research.
- The assumption of similarity between animal and human brains.
- Ethical considerations in animal research.

Key terms: animal model

What is an animal model?

The most popular species to be used in psychological research are rats, mice, pigeons, cats, rabbits, hamsters, dogs, chimpanzees, and baboons. From 2017–18 in the USA alone, the number of animals used in experimentation was 780,000. This only includes mammals covered by the Animal Welfare Act (AWA). Rats and mice are not covered by this Act and statistics on these species is not readily available, but it is estimated that they comprise over 99% of animals used in research. In that same year in the USA an estimated 44.5 million mice and rats underwent potentially painful experiments (Carbone, 2021).

SAQ Animal research / animal models

Animal research/animal models is one of the content points that can appear in questions in Papers 1A and 1B. You will need to be able to explain it and support your explanation with one example.

An **animal model** is a concept that refers to using animal research to test a certain cause–effect hypothesis about a certain human behaviour. This term is often used somewhat loosely. For example, you can see mentions of "mouse models" to denote the general idea that mice can be used to model human behaviour. However, to fully identify an animal model, it is important to include the following information:

- What animal is being used?
- What behaviour is being modelled?
- What causal factor is being investigated (or what hypothesis is being tested)?

For example:

- For "a mouse stress model of depression" read: using mice to investigate the idea that exposure to stress increases the risk of developing depression in humans.
- For "a rhesus monkey separation model of cognitive delay" read: using rhesus monkeys to investigate if separation from attachment figures will cause delays in cognitive development, and generalizing these findings to humans.

As a rule, the purpose behind research with animals is to gain a deeper understanding of human behaviour. This rests on the assumption that the structure and the chemistry of the brain is, in some cases, very similar in humans and animals. However, this assumption cannot be said to be fully justified at all times.

There will be several prominent examples of animal studies in this book. You can use any one of them to support your answer. One of the examples—the study of Merzenich et al. (1984)—is also provided below.

Example: Merzenich et al. (1984)



▲ Figure 1.21 An owl monkey

Merzenich et al. (1984) investigated cortical remapping of digits (fingers) in owl monkeys.

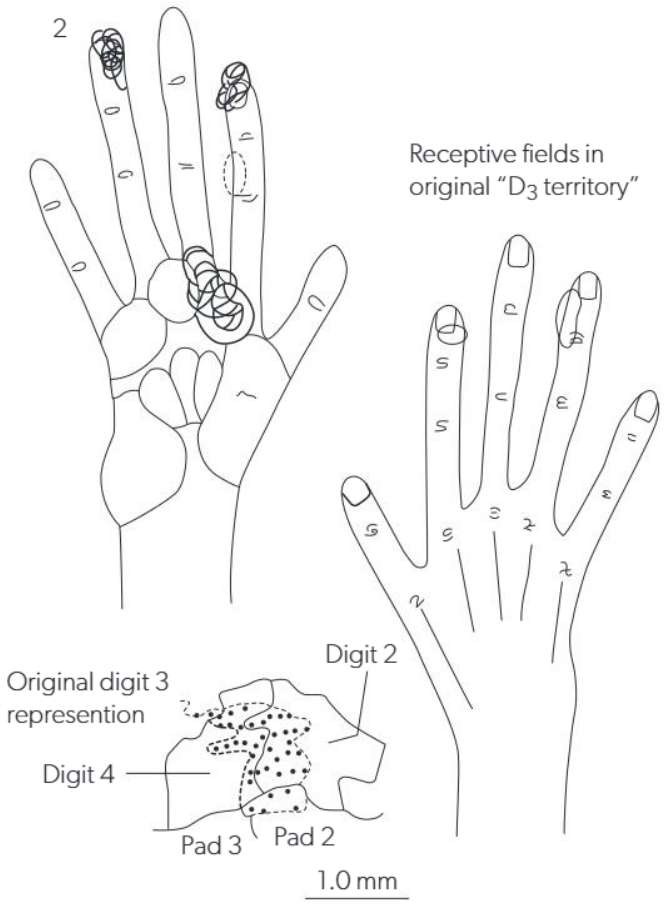
Participants in the study were eight adult owl monkeys. It was an experiment using the repeated measures design. First, researchers mapped the sensory inputs from all the hand digits (fingers) in the cortex. To do this, electrodes were attached to the part of the cortex known to be responsible for sensations from the hand, then different fingers were stimulated. It was noted which of the electrodes responded to the stimulation.

Following this, one or several fingers on the monkey's hand were amputated. A remapping was done 62 days after the amputation to see how the cortex adapted to the injury.

Results of the first mapping revealed that there were five distinct areas in the cortex, each responsible for one digit. Adjacent fingers were represented by adjacent areas in the cortex. Post amputation, the now unused area of the sensory cortex was occupied by adjacent intact fingers. For example, if digit 3 had been amputated, the cortical areas for digits 2 and 4 spread out and "consumed" the cortical area previously responsible for digit 3.

From this, researchers concluded that the sensory cortex of adult owl monkeys can adapt to injury by cortical remapping.

Such cortical remapping is an example of neuroplasticity—the ability of the brain to reshape itself in response to environmental conditions. Neuroplasticity in response to structural damage is observed in the human brain as well, but a direct experiment would not be possible for ethical reasons. Therefore, human research in this area is limited to case studies of people with injury. Cause-and-effect inferences cannot be made from case studies. Research with animal models helps test cause–effect hypotheses and in this sense provides further insight into human behaviour.

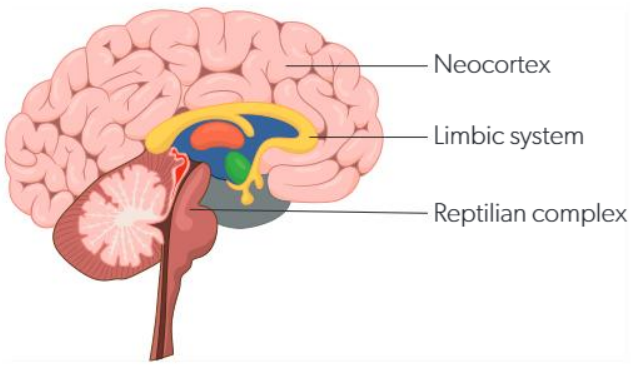


▲ Figure 1.22 Cortical remapping following digit amputation in adult owl monkeys (Merzenich et al., 1984, p. 595). Digit 3 (the middle finger) was removed, and the original digit 3 representation of the brain became redistributed to adjacent digits 2 and 4.

Advantages and disadvantages of using animal research

As a rule, animals are studied in psychology because it is believed that they can provide an insight into human behaviour. Using animal research to inform our understanding of human behaviour relies on the assumption that animal and human brains are similar.

There is a popular theory of the "triune" brain proposed by MacLean (1990). This theory divides the human brain into three parts: reptilian complex, paleomammalian complex (the limbic system), and neocortex. The idea is that the deeper brain structures can be found in animals as well. The further down you go inside the brain, the further back you see in evolution. For example, the reptilian complex that you have in your brain should resemble the full brain of a reptile.



▲ Figure 1.23 Triune brain

Some of the advantages of working with animal models:

- Humans and animals are identical in many ways, both in terms of brain structure and genetically.
- Studies with animal models do produce results: useful models of human behaviour and life-saving treatments have been developed based on animal experimentation. For example, insulin was discovered in an experiment in which dogs had their pancreas removed.
- Animal studies allow researchers to embrace the full lifespan. While human subjects often outlive researchers themselves, laboratory mice live two to three years and this presents an opportunity to see their behaviour across their lifespan and even across generations. This is especially helpful in genetic research.
- Animal research may be highly controlled. For example, the "knockout" technique has been developed to selectively switch off one of the genes in the DNA sequence. All other things being equal, this technique provides great insight into the function of individual genes. The ability to better control confounding variables means higher internal validity of experiments.
- Animal subjects are relatively inexpensive and easily accessible, easy to handle and manage.

Activity

Every time you encounter an animal study in this book, evaluate it by using the advantages and disadvantages listed earlier. The value of animal models links directly to whether or not animal research can provide insight into human behaviour.

ATL Research, Self-management

Some of the disadvantages of working with animal models:

- Animals and humans are never exactly the same, and we can never know the extent of the difference. This means that animal research, if successful, still needs to be replicated with humans in order to be sure that findings are generalizable.
- Animals are tested in strictly controlled laboratory environments, so arguably they may be under stress. As a result, their reactions to experimental manipulations may not be quite the same as in their natural environments: there may be an issue with ecological validity.
- Although humans and animals are similar in many ways, they are still essentially different. For example, over 85 vaccines for HIV worked well in primates but all of them have failed in humans (Bailey, 2008). On the contrary, some results that are negative in animals can actually turn out to be positive in humans. For example, aspirin proved dangerous for animals, but it is now one of the most widely used medications for humans.

Activity

Review the section of the PETA (People for the Ethical Treatment of Animals) website devoted to animals used for experimentation. Have a class debate for and against using animals for experimentation in psychology. Randomly split into two groups. One group will present the argument "for", the other "against". Both groups should take time to prepare their arguments and possible rebuttals for the opposing group's claims.

ATL Communication, Social

Ethical considerations in animal research

Ethical considerations in animal research are regulated by professional bodies such as the American Psychological Association (APA). Most ethical considerations revolve around making justified research choices, carefully monitoring potential pain inflicted on the animals, and getting approval from independent review teams.

The APA guidelines regulate every step of the process in an animal research study. Some of the major guidelines are summarized below (American Psychological Association, 2012).

- Justified choices:
  - Any animal study should be clearly justified with a scientific purpose. It should either increase our knowledge or benefit humans or other animals.
  - The chosen species must be the best choice for the research purpose.
  - The minimum required number of animals must be used.
- Inflicting pain:
  - It has to be assumed that whatever procedures cause pain in humans would cause pain in animals, too.
  - Researchers conducting the study must be familiar with the species-specific characteristics of normal behaviour so that they will be able to tell when the animal is stressed or unhealthy.
  - Whenever possible, laboratory procedures must be designed in a way that minimizes animal discomfort.
  - Whenever reasonable, researchers must first test the painful stimuli to be used with animals on themselves.
  - If a research animal is observed to be in distress or chronic pain and this is not necessary for the purposes of the study, it should be euthanized.
  - Animals reared in the laboratory must not be released into the wild.

- Obtaining approval:
  - All animal research proposals must be submitted to the ethics committee prior to conducting the study.

The British Psychological Society (BPS) published a policy on the use of animals in psychology, which is based on the so-called "three Rs" (British Psychological Society, 2012):

- Replacement (animals should only be used when no alternative exists).
- Reduction (the minimal necessary number of animals must be used).
- Refinement (it must be ensured that experimental procedures cause minimal necessary distress in the animals).

Think back to the study of Merzenich et al. (1984). The research used a very invasive experimental manipulation which had a damaging and irreversible effect on the animals' lives (amputation of fingers). The technique used to measure the cortical response also involved inserting electrodes into the brain. Such research proposals must be carefully scrutinized by the ethics committee and only approved if potential gains outweigh the costs. Experimenters must make sure that animals are properly anaesthetized during the procedure and taken care of after the end of the study for the duration of their lives.

TOK

Here is a thought experiment that has been used as an argument to say that animal experiments are ethically justified.

Imagine you see a small van with 500 mice in it rolling slowly towards the edge of a cliff. There is no driver in the van. In its way there's a stroller with a human baby in it. There are two possible outcomes:

1. You push the stroller away and let the van roll slowly off the cliff, killing the mice.
2. You do nothing and the stroller prevents the van from going over the cliff, but this kills the human baby.

Defenders of animal experimentation say that everyone would choose the first option because human life is more valuable than the lives of 500 mice. They also say that using animal studies to develop potentially life-saving medicine is equivalent to this thought experiment.

What do you think about this argument?

1.10 Descriptive statistics

Inquiry questions

- How can we numerically describe a variable?
- How do we interpret the measures of central tendency and dispersion?

What you will learn in this section

Key learning:

- Levels of measurement: nominal, ordinal, interval, ratio.
- Normality of distribution.
- Measures of central tendency: mean, median, mode.
- Measures of dispersion: standard deviation, semi-interquartile range.

Key terms: nominal-level variables, ordinal-level variables, interval-level variables, ratio-level variables, normal distribution, mean, median, mode, standard deviation, semi-interquartile range, quartiles

As the name suggests, the purpose of descriptive statistics is to describe variables. To describe a variable means to determine the level of measurement, normality of its distribution, and calculate a measure of central tendency and a measure of dispersion.

Levels of measurement

There are four levels of measurement, or types of variables.

Nominal-level variables cannot be quantified. They represent a set of labels that cannot be placed in either descending or ascending order. Examples of nominal variables include: car brands, gender, nationality, and music genre preferences. Some people prefer jazz and some listen to rock, but you cannot rank people by their music preferences from the lowest to the highest. Since nominal variables cannot be quantified, procedures such as addition, subtraction, or multiplication cannot be applied to them. This affects the range of statistical tests that is possible.

Ordinal-level variables can be ranked from the lowest to the highest, but the intervals between ranks may be unequal. An example would be the results of a car race: competitors come first, second, third, and so on. But the difference between the person who comes first and the person who finishes second is not necessarily the same as the difference between the second and the third finishers.

Interval-level variables are like ordinal variables, but the intervals are assumed to be equal. Importantly, these variables may have a zero value, but zero does not mean an "absence" of the parameter. An example would be an ordinary thermometer that measures temperature in centigrade. You can say that the

difference between 35°C and 37°C is the same as the difference between 31°C and 33°C. You also know that 0°C does not mean an "absence of temperature", it is just an arbitrary point on the scale. Due to these properties of interval variables, addition and subtraction are allowed but multiplication and division are not. For example, it is correct to say that 12°C + 5°C = 17°C, but it is incorrect to say that 20 °C is twice as much as 10°C because if you move the zero point, that would no longer be true.

Finally, **ratio-level variables** are like interval variables, but the zero is fixed and meaningful. It means the absence of something. For example, the number of words correctly recalled from a list may be considered ratio data because if you recalled no words, it simply means that you have not recalled anything. Annual salary, age, and weight are all examples of ratio-level data. Building on our example above, if you measure temperature in degrees Kelvin (where 0° corresponds to minus –273°C and is known as the "absolute zero", the lowest temperature possible), then your data is measured on the ratio level. The full range of mathematical operations can be applied to ratio-level data, including division and multiplication.

Exam tip

There exists some disagreement as to what level of measurement the Likert scale belongs to. Some say ordinal, some say interval.

It would be reasonable to assume that when we are dealing with a single Likert-scale question (e.g., in a survey), this is an ordinal scale. This is because there is no guarantee that the psychological distance between "absolutely disagree" and "disagree" is the same as the distance between "disagree" and "neither agree nor disagree".

However, when we combine multiple Likert-scale questions into a scale (like we normally do in a questionnaire) and calculate the total score per scale, this scale is usually considered to be interval-level.

(1)	(2)	(3)	(4)	(5)
Absolutely disagree	Disagree	Neither agree nor disagree	Agree	Absolutely agree

▲ Table 1.9 The range of responses on a Likert scale

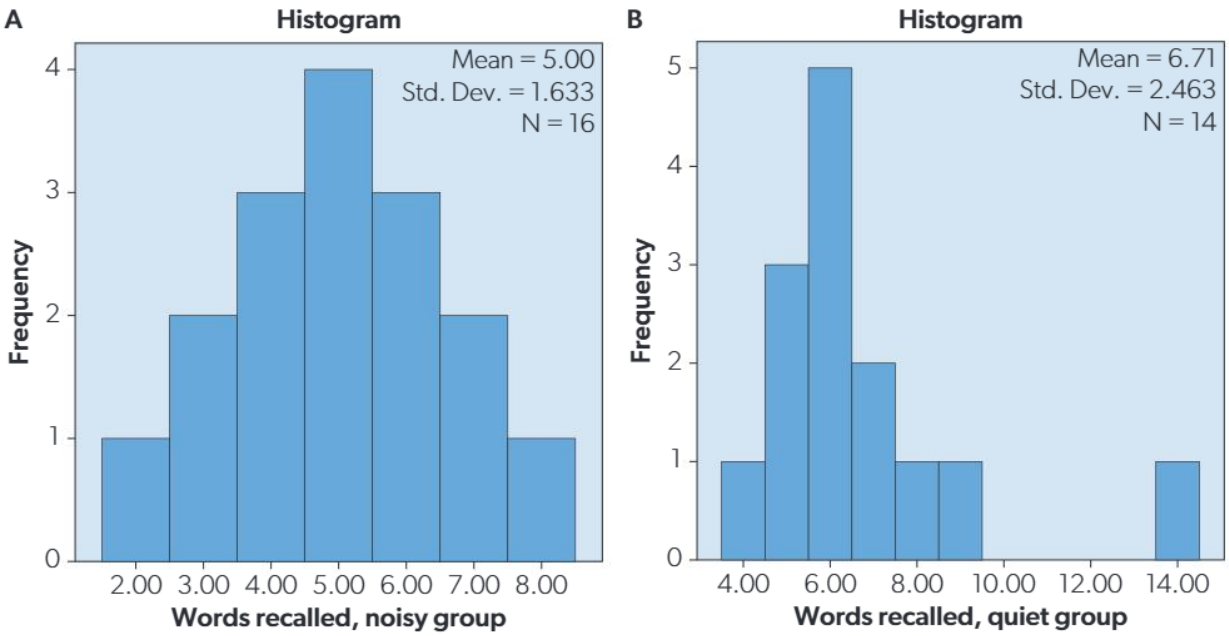
Normality of distribution

Distribution is the range of values of a variable according to their frequency. Suppose you gave two groups of people a list of words under two conditions—quiet and noisy—and asked them to recall the words later, registering the number of words that they recalled correctly. Table 1.10 could be the numerical results of the investigation.

Experimental condition (noisy)		Control condition (quiet)	
Participant	Number of words correctly recalled	Participant	Number of words correctly recalled
1	5	**1**	14
2	4	**2**	5
3	6	**3**	9
4	3	**4**	4
5	4	**5**	6
6	5	**6**	5
7	5	**7**	6
8	6	**8**	5
9	4	**9**	6
10	6	**10**	7
11	7	**11**	8
12	5	**12**	6
13	3	**13**	7
14	7	**14**	6
15	8		
16	2		

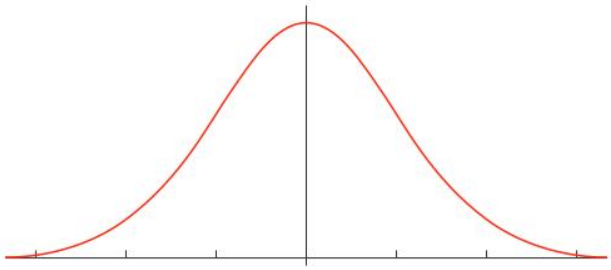
▲ Table 1.10 Data from a hypothetical experiment

If we plot the frequency distribution of this data, Figure 1.24 is what we will get:



▲ Figure 1.24 Distribution of words recalled correctly, noisy condition (A) and quiet condition (B)

Normal distribution is a special type of distribution that looks like a bell. This is why it is sometimes called the "bell curve". Figure 1.25 is a visual representation of an ideal normal distribution.



▲ Figure 1.25 Normal distribution

Normal distribution occurs naturally in many situations. For example, IQ scores, height of people, age, and blood pressure are all be normally distributed. Most people would be "average", and when the scores are more extreme, fewer people in the population would have these extreme scores.

In reality, most data sets will deviate to some extent from the ideal normal distribution. How large is this deviation, however, and can we still assume that data is normally distributed even if minor deviations take place?

Rigorous tests exist to compare the empirical distribution (the one you obtained in a study) to the ideal normal distribution and conclude whether the deviation is small enough to assume normality. However, these tests fall outside the scope of this IB Psychology course; you just need to know the concept of normality of distribution and use the following general principles.

- Scan the raw data for outliers. An outlier is an observation that is extremely different from most or all other observations. For example, in our fictitious data set, the participant who scored 14 (participant 1 in the control condition) may be called an outlier because the score is so different from all other scores. If severe outliers are present, the distribution is probably not normal, especially with small sample sizes.
- Assess the frequency distribution visually. If its shape approximately resembles a bell (with average values being most frequent and extreme values less frequent), you may assume normality.

Measures of descriptive statistics: central tendency and dispersion

All measures of descriptive statistics can be divided into measures of central tendency and measures of dispersion. Together they help us summarize data to make sense of it.

Measures of central tendency

There are three measures of central tendency: the mean, the median, and the mode.

The **mean** is simply the average of all scores. Add all the data points and divide them by the number of observations (participants in the group). The mean is the most common measure of central tendency, but it may be biased if there are

some extreme outliers in the sample. The problem with outliers is that they can skew the mean so that it is no longer a meaningful indicator of central tendency. Be careful when applying the mean to data sets with extreme outliers and maybe use the median or the mode instead.

The **median** is the "middle" of a sorted list of numbers. To find the median, place the numbers in value order. For example, suppose we have the following values:

3, 5, 7, 12, 13, 14, 21, 23, 23, 23, 23, 29, 40, 56

There are 14 values so the "middle" should be between the seventh and the eighth value: we just find the average.

$$\frac{(21 + 23)}{2} = 22.$$

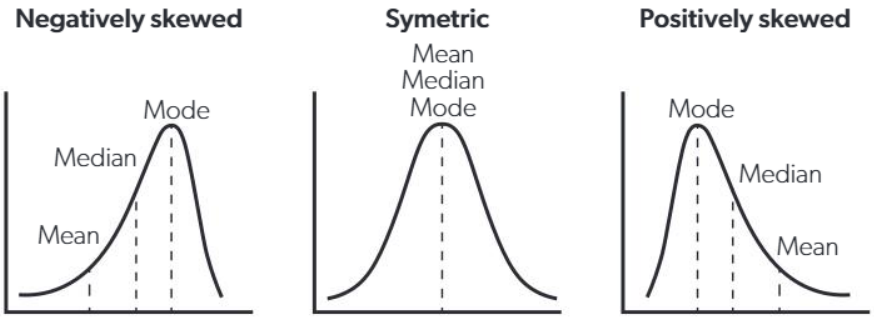
Therefore, the median of this data set is 22.

The median is used when there are strong deviations from normality of distribution, making the mean misleading.

The **mode** is the value that appears most often in the data set. In our hypothetical dataset of people recalling a list of words under the quiet condition, the mode is 6. There are cases when a dataset has more than one mode. For example, here is a binomial data set for which the modes are 3 and 6:

1, 3, 3, 3, 4, 4, 6, 6, 6, 9

In an ideally normal distribution the mean, the median, and the mode will coincide. As the distribution deviates from normality, especially if there are outliers, they may move apart.



▲ Figure 1.26 Mode, mean, and median in normal and non-normal distributions

Therefore, when the level of measurement is interval or ratio, your choice of the measure of central tendency should be guided by your assumptions about the normality of distribution in the data sets. When the level of measurement is ordinal, the mean cannot be used. When the level of measurement is nominal, you can only use the mode.

Measures of dispersion

To make sense of data, measures of central tendency are not enough because some data sets are more spread out around the centre than others. Two measures of dispersion are commonly used: the standard deviation and the semi-interquartile range.

Standard deviation is calculated using the following formula:

$$SD = \sqrt{\frac{\sum(x_i - \overline{X})^2}{N - 1}}$$

Let us now decipher the formula step by step.

- Calculate the mean of the dataset ($\overline{X}$).
- For each individual value (x_i), calculate its deviation from the mean ($x_i - \overline{X}$).
- Square the deviation, obtaining squared deviations $(x_i - \overline{X})^2$. This eliminates the signs (positive and negative).
- Calculate the sum (Σ) of all squared deviations and divide it by the number of observations (N) minus one. This gives you the "average squared deviation".
- Finally, calculate the square root of that value.

Standard deviation is an accurate measure of dispersion that considers all individual values and uses all information available in the dataset, making it the preferred choice when possible. A low standard deviation indicates that the data points tend to be closely grouped around the mean, whereas a high standard deviation indicates that the data points are further spread out around the mean.

However, since the formula uses the mean, the standard deviation cannot be used with nominal or ordinal data, or with datasets that severely violate normality of distribution.

Semi-interquartile range, unlike standard deviation, does not assume normality of distribution and can be used with ordinal-level data.

When you sort a list of numbers, you can find the so-called **quartiles**:

- Q2 (the middle quartile) is the median
- Q1 (the lower quartile) is the median of the numbers to the left of Q2
- Q3 (the upper quartile) is the median of the numbers to the right of Q2

Taking our example from before:



▲ Figure 1.27 Finding quartiles

- Q1 is the median of the dataset 3, 5, 7, 12, 13, 14, 21. Therefore, Q1 = 12
- Q2 = the median = 22
- Q3 is the median of the dataset 23, 23, 23, 23, 29, 40, 56. Therefore, Q3 = 23.

The semi-interquartile range is calculated simply as:

$$SIR = \frac{Q3 - Q1}{2}$$

In this example, the semi-interquartile range is equal to $\frac{23 - 12}{2} = 5.5$.

Similarly to the standard deviation, the higher the semi-interquartile range, the more the individual data points are spread out around the centre of the distribution.

Chat with AI

Consider using the following sample prompts with your favourite generative AI to deepen your understanding of measures of central tendency and measures of dispersion in data analysis:

- What does it mean for data interpretation when standard deviation is large? Could you explain with a simple example?
- Could you give an example of a dataset for which you would prefer semi-interquartile range to the standard deviation as a measure of dispersion? Please explain.
- Why is it important to know a measure of dispersion such as the standard deviation? Why is the mean not enough?

1.11 Plotting and graphing

Inquiry questions

- How do we interpret data presented in visualizations, such as scatterplots and bar graphs?
- What is the value in visualizing data?

What you will learn in this section

Key learning:

- Histograms and frequency distribution charts.
- Bar graphs.
- Box-and-whisker plots.
- Scatterplots.
- Other visualizations.

Key terms: histogram, frequency distribution chart, bar graph, box-and-whisker plot, scatterplot, skewness, kurtosis, outlier

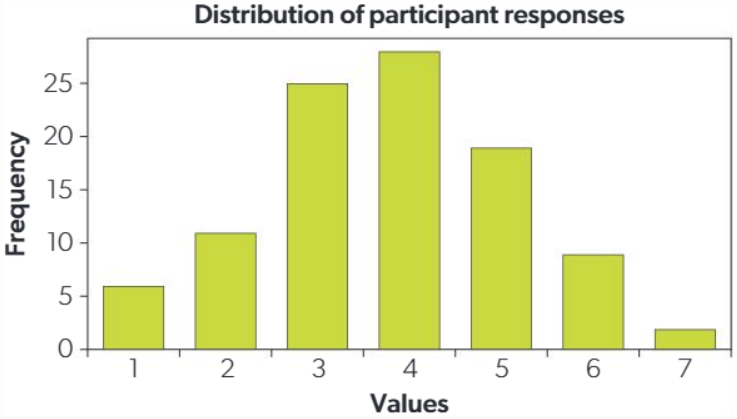
Overview

To better grasp the meaning of data, we can use visualizations. Visualizing data is a discipline in its own right, and rapidly developing. For example, if you make a career as a data journalist, you will be telling stories through original visualizations. Today's data visualization techniques go far beyond drawing a nice chart. They include, for example, creating interactive dashboards where the user can interact with the data, interrogating it by filtering, slicing, dragging, combining, and comparing. Animations are used to bring data to life and tell a story dynamically.

Here we will consider some basic data visualizations that are commonly used in psychological research papers. We will focus on how to interpret them and what to look out for in terms of potentially misrepresenting the findings.

Histograms and frequency distribution charts

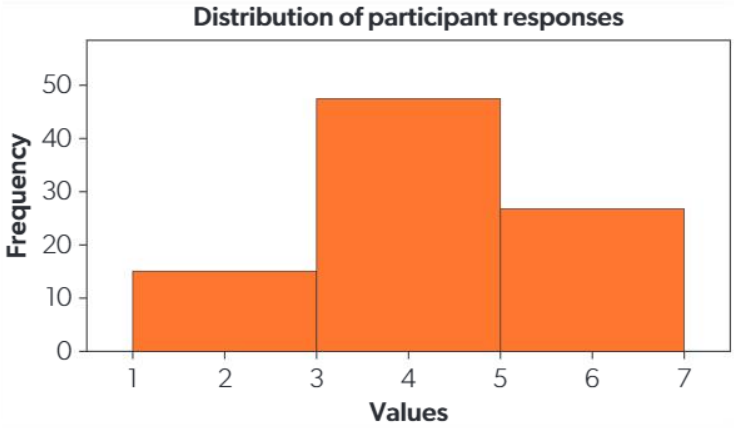
Histograms are usually used to show the distribution of a variable in the population. Typically along the *x*-axis (the horizontal axis) a histogram includes the possible scores that a variable can take, and along the *y*-axis (the vertical axis) the number of people who got that particular score. For example, in Figure 1.28 you can see the distribution of participants' responses to the survey item "I feel anxious when I have to speak in public" (responses were collected with a 7-point Likert scale ranging from "absolutely disagree" to "absolutely agree").



▲ Figure 1.28 Histogram: distinct bands

An advantage of a histogram is that it allows you to quickly capture both the central tendency and the dispersion of your variables. You can visually assess if the distribution is normal or deviates from normality, and you can also see **outliers.**

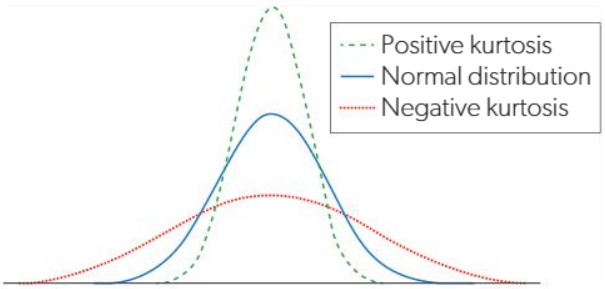
One thing to look out for is the bands. If someone wanted to intentionally misrepresent data using a histogram, it would be possible to some extent by manipulating the bands on the *x*-axis. For example, in Figure 1.29 we show the same data for the item "I feel anxious when I have to speak in public", but this time we grouped the possible values in larger bands of values. Visually, the two histograms may suggest different interpretations, although the underlying data is the same.



▲ Figure 1.29 Histogram: combined bands

There are two parameters that are used to describe the deviations of a given distribution from normality: skewness and kurtosis. **Skewness** is how slanted the distribution is towards left or right. Skewness has already been shown, in Figure 1.26 (Normal distributions on page 68). In a distribution with a positive skewness, many people score low values and there is a relatively long tail on the right. In a distribution with a negative skewness, the tendency is the opposite. Note that the graphs in Figure 1.26 and Figure 1.30 (opposite) technically are not histograms because they do not have discrete bars. We refer to them as "**frequency distribution charts**". But you can think of them as a histogram with a very large number of bars.

Kurtosis is the "pointiness" versus "flatness" of a distribution. A large kurtosis means that the distribution peaks too much as compared with what would be expected from a normally distributed variable. A negative value of kurtosis means that the distribution is too "flat", with not as many people getting particular scores as would be expected with a normal distribution.

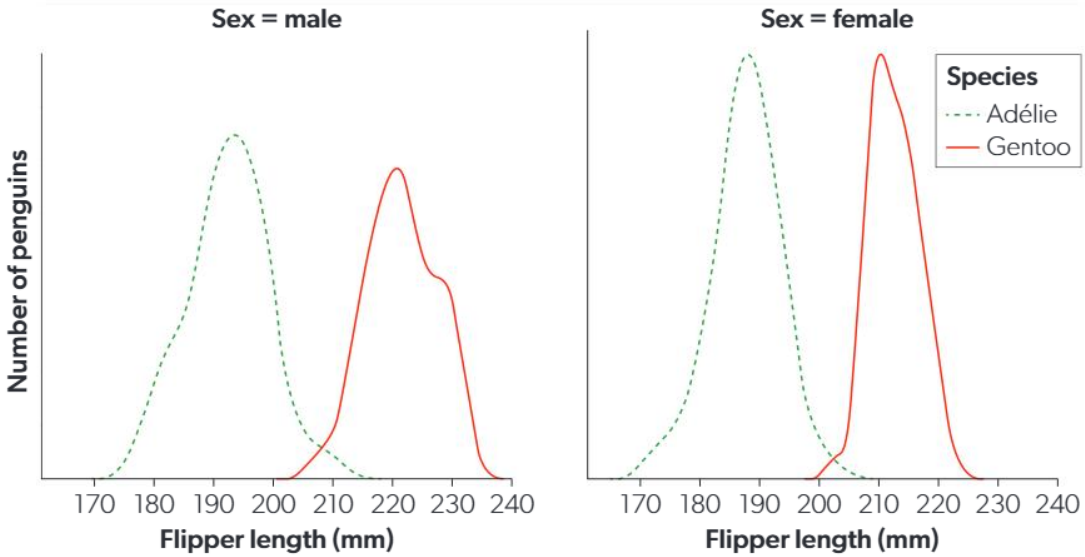


▲ Figure 1.30 Kurtosis

Both skewness and kurtosis can be quantified, although this goes beyond the scope of this book. There are statistical formulae that use skewness and kurtosis to make conclusions about whether or not the given distribution can be considered normal.

Note: if you have more than one variable measured on the same scale, it is possible to overlay more than one frequency distribution on one chart, in which case you could visually compare the distributions (see Figure 1.31).

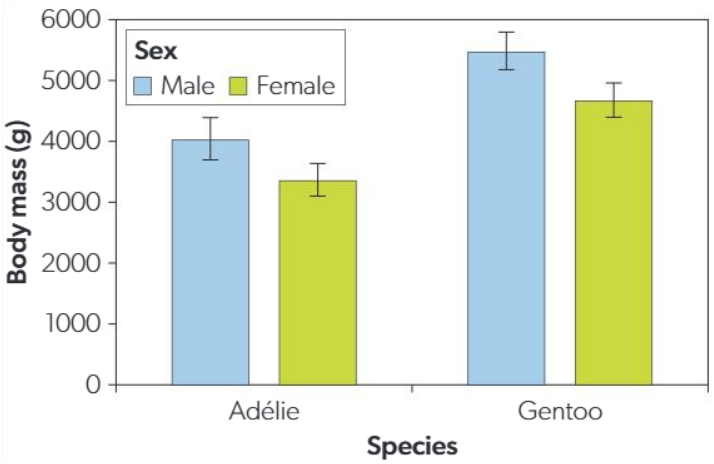
Note: to illustrate some of the ideas in this section, we are using a real-life dataset originating from penguin research conducted at Palmer Station in Antarctica (Gorman, Williams, and Fraser, 2014). It is a popular dataset among those who study data science. Although it is not related to psychology, the principles of data representation are the same.



▲ Figure 1.31 Frequency distribution charts: length of flippers (in millimetres) in two species of penguins—Adélie and Gentoo, male and female

Bar graphs

Bar graphs are commonly used to visualize results of experimental research that involves a comparison between groups or conditions (e.g., experimental versus control). Look at Figure 1.32—a typical bar graph:



▲ Figure 1.32 This bar graph compared the average body mass in two species of penguins—Adélie and Gentoo, male and female

Along the *x*-axis in Figure 1.32, we have variables and groups—the sex (male versus female) and the species (Adélie versus Gentoo) of the penguin.

Along the *y*-axis we have the value of the variable on which the groups are being compared—in this case, the penguin's body mass in grams.

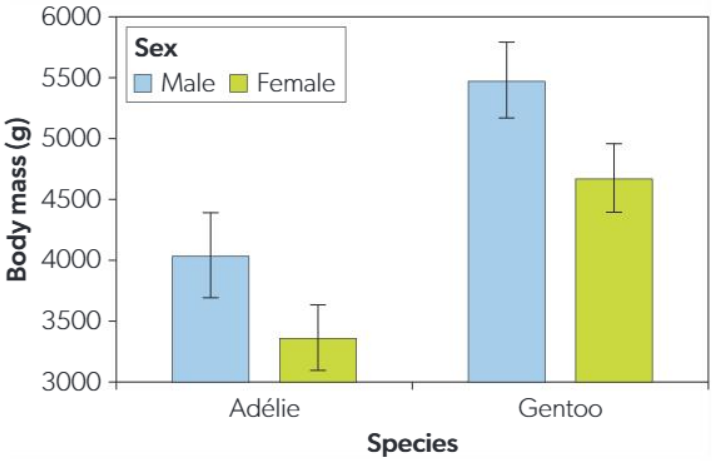
If more than one variable is used in the same bar graph (like this one), it is important that they are measured on the same scale. The height of the bar itself usually denotes the mean value and the error bar denotes the confidence interval (the spread around the mean). The following confidence intervals are usually used:

- 68% confidence: ±1 × standard deviation (or standard error)
- 95% confidence: ±1.96 × standard deviation (or standard error).

Note: it is good practice to include a legend or a note explaining what exactly the bar and the error bars denote. For example, the bar could denote the median rather than the mean and the error bars could denote standard deviations or even quartiles.

Examples of poor practice in constructing a bar graph include the following:

1. Not starting the *y*-axis with 0. Doing so can visually inflate the differences between conditions and make a small difference look large—see Figure 1.33.
2. Not including a legend or a note explaining what the elements of the graph denote (e.g., the error bars). This makes the graph ambiguous and not self-sufficient.
3. Combining two variables that used different measurement scales in one bar graph. This may make one of the variables appear "squashed" and downplay the difference between conditions in the variable.

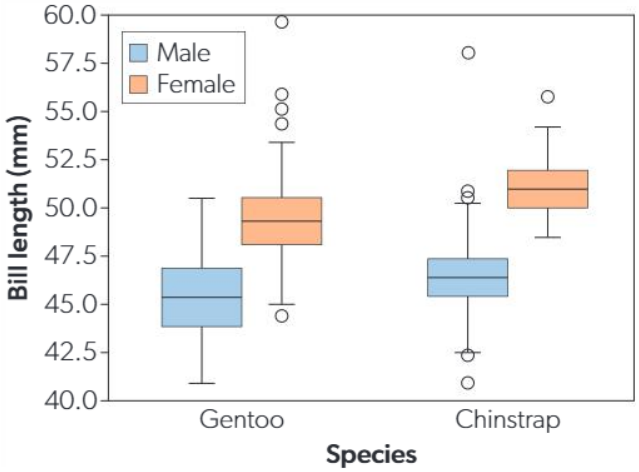


▲ Figure 1.33 This bar graph is based on the same data, but it does not start at 0

Note: from the bar graph alone we cannot judge whether or not the difference between groups or conditions is statistically significant. We need to conduct an inferential statistical test to make these types of judgement. Unit 1.12 covers inferential statistics. As a rule of thumb, you could see if the confidence intervals between the two groups or conditions are overlapping. If they are not overlapping, it is likely (but not certain) that the inferential statistical test will demonstrate a statistical significance.

Box-and-whisker plots

Box-and-whisker plots are used for a similar purpose—to visually compare two or more groups or conditions. They are used instead of bar graphs when the variables are not normally distributed—so there are reasons to believe that using the standard deviation to represent the spread of data would be misleading, or that there are outliers in the dataset that can distort these descriptive statistics.



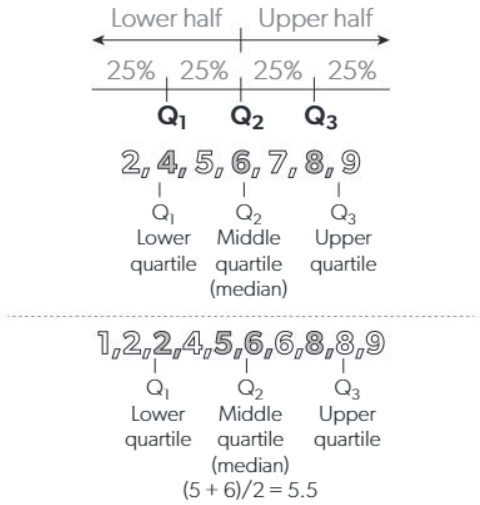
▲ Figure 1.34 Box-and-whisker plot: bill length of penguin, depending on their sex and species (Gentoo or Chinstrap)

A box-and-whisker plot usually includes the following elements:

1. The dot or line in the middle represents the central tendency in the distribution of a variable (since box-and-whisker plots are used when the variables are not normally distributed, this point typically represents the median).

2. The "box" around the central point shows the range of values within which lies 50% of the sample (25% to the left of the middle point and 25% to the right).
3. The whiskers represent the maximum and the minimum points, excluding outliers.

Box-and-whisker plots use quartiles. Together, quartiles divide the distribution into four equal parts each containing 25% of the sample.

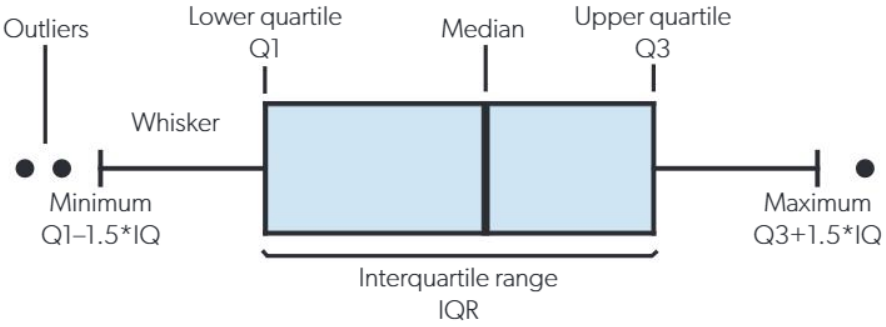


▲ Figure 1.35 Quartiles

The box in the box-and-whisker plot represents the range between Q1 and Q3. This range is also called the interquartile range (IQR). The whiskers are usually calculated as:

- Lower part: Q1 – 1.5 × IQR
- Upper part: Q3 + 1.5 × IQR.

Any values falling beyond the whiskers are considered outliers and represented as dots on the box-and-whisker plot.

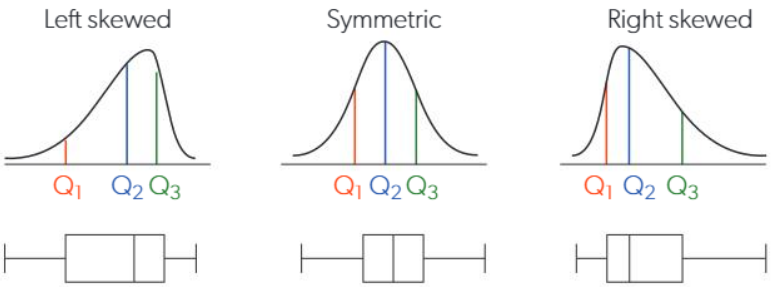


▲ Figure 1.36 Elements of a box-and-whisker plot explained

The plot is very helpful to get a sense of how the variable is distributed. For example:

- the position of the line within the box tells you in which direction the distribution is skewed
- the width of the box tells you how spread out the values are
- the outliers are informative too.

Importantly, measures of central tendency and dispersion in box-and-whisker plots are not distorted by outliers as much as they would be if we simply plotted a frequency distribution.

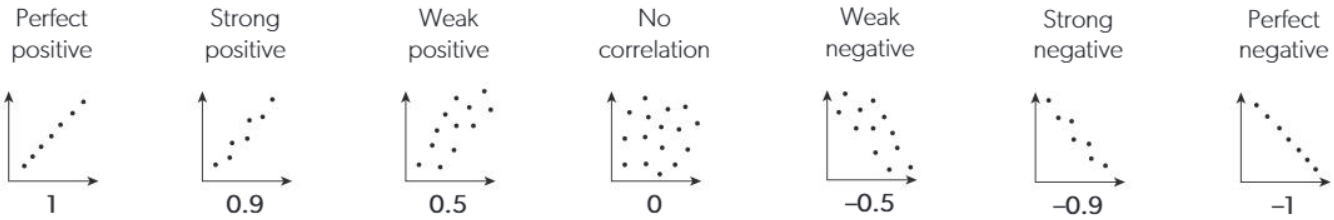


▲ Figure 1.37 Box-and-whisker plots and skewness of the distribution

Scatterplots

We have already considered **scatterplots** when we were discussing correlational research in psychology in Unit 1.5.

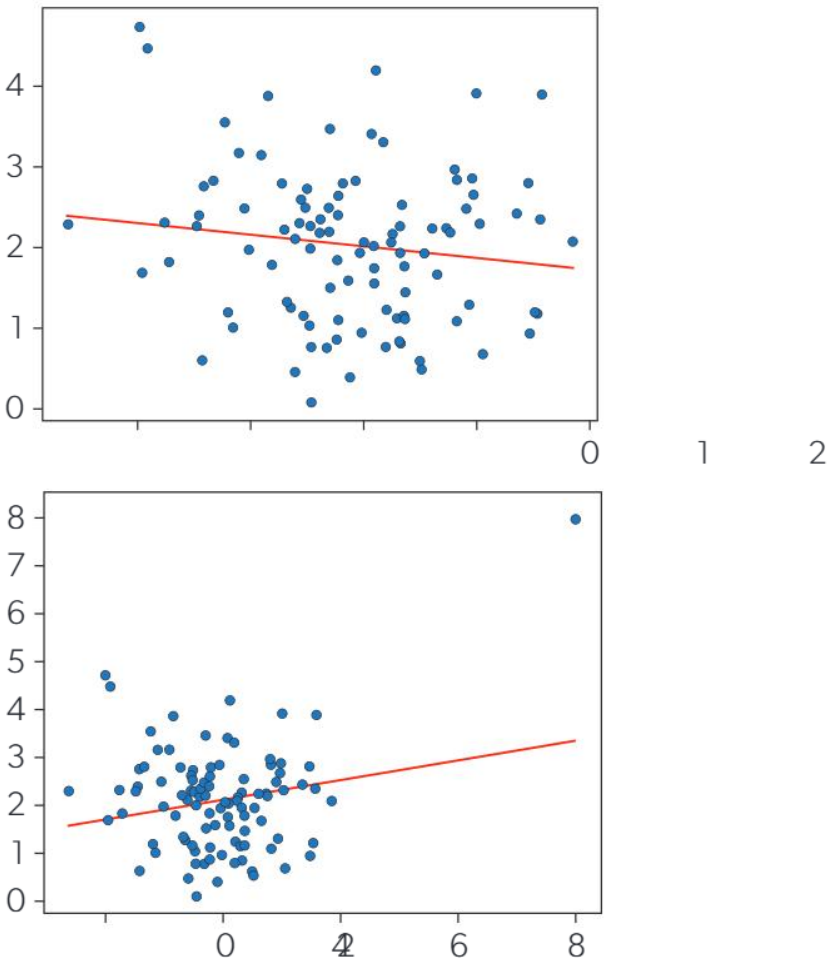
Scatterplots are useful to visualize the relationship between two variables. When there is a positive correlation, participants who score higher on one variable will also tend to score higher on the other. Visually, this will appear as a scatterplot that is squashed and stretched diagonally and can be best represented by a straight line with an ascending trend. In a perfect correlation ($r = 1.0$), every single data point will fall exactly on the trend line.



▲ Figure 1.38 Correlations and scatterplots

When there is a negative correlation, the data points will stretch from the top left to the bottom right: participants who have higher values on one variable also tend to have lower values on the other variable. When there is no relationship between the variables, the scatterplot will look like a swarm with no particular shape, and the linear trend would be best represented by a "flat" line running parallel to the *x*-axis.

One needs to bear in mind that scatterplots can be visually very affected by outliers. Consider the two images in Figure 1.39 on page 78. They are based on the same data. Only one extra data point was added to make the bottom scatterplot. It would be good practice to remove the outliers or make them less extreme by changing their values before constructing a scatterplot.



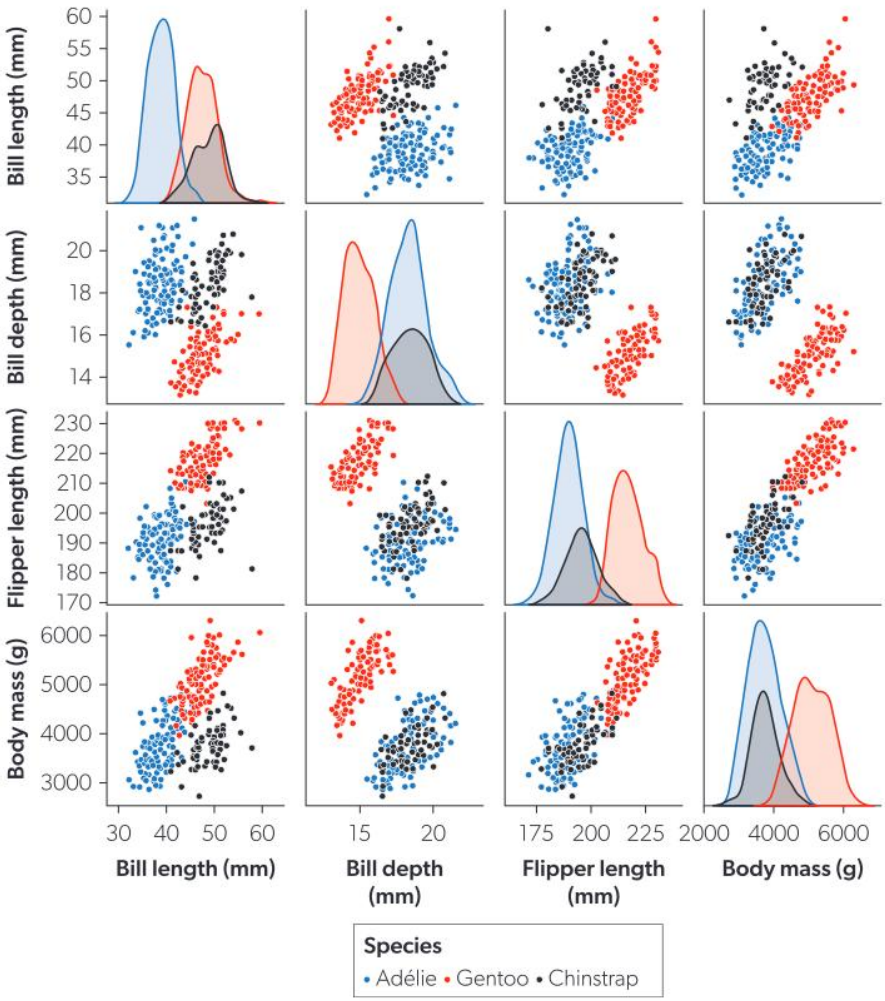
▲ Figure 1.39 An outlier (the data point in the top right corner of the second graph) affecting a scatterplot

However, researchers must remember to document all such data manipulations and explicitly report them in the publication. We cannot just change real-world data because it does not look nice on our chart. And, if we do, we must say so and clearly justify the decision.

Other visualizations

We have considered the most basic data visualizations that are commonly used in psychological research. A large variety of other visualization techniques exist. Some visualizations are created by using the simple elements, but combining them to represent information about multiple groups of participants or multiple interrelated variables at the same time. Unfortunately, in most cases we are limited by two dimensions and have to create something that can be shown on a flat sheet of paper or screen. This limits the possibilities of visualization techniques.

Figure 1.40 is an example of a complex visualization combining several simple techniques. It represents four parameters (depth of bill, length of bill, length of flipper, body mass) in three species of penguins (Adélie, Gentoo, Chinstrap). It uses scatterplots to show pairwise relationships between variables (grouped by species) and frequency distribution charts for each of the four variables (also grouped by species).



▲ Figure 1.40 Depth of bill, length of bill, length of flipper, and body mass in three species of penguins: a categorized chart

1.12 Inferential statistics

Inquiry questions

- What is the difference between descriptive and inferential statistics?
- How do you read and interpret the results of inferential statistical tests?
- What should be considered to fully understand a research result reported in statistical terms?
- What is the computational meaning of some of the basic inferential tests?

What you will learn in this section

Key learning:

- What are inferential statistics?
- The null hypothesis and the alternative hypothesis.
- Type I and Type II error rates.
- Parametric and non-parametric tests.
- Effect size and statistical significance.
- Which inferential test is appropriate in a given situation?
- Measures of difference: the unrelated *t*-test.
- Measures of difference: Mann–Whitney *U* test.
- Measures of relationship: Pearson's correlation coefficient.
- Measures of relationship: Spearman's correlation coefficient.
- Reporting and interpreting the results of inferential statistics.

Key terms: null hypothesis (H_0), alternative hypothesis (H_1), Type I error, Type II error, parametric test, non-parametric test, effect size, Pearson's *r*, Cohen's *d*, statistical significance, unrelated *t*-test, Mann–Whitney *U* test, Pearson's correlation coefficient, Spearman's rank correlation coefficient

What are inferential statistics?

Unlike descriptive statistics, which only allow us to describe data (as the name suggests), inferential statistical tests make it possible to test hypotheses and make conclusions (inferences).

Broadly and for our purposes, all inferential statistics can be divided into two groups: measures of difference and measures of relationship. Measures of difference answer the question "Is there a difference between these groups or these conditions?" Measures of relationship answer the question "Is there a relationship between these variables?"

The null hypothesis and the alternative hypothesis

When researchers conduct a statistical analysis, they formulate two hypotheses—the **null hypothesis (H_0)** and the **alternative hypothesis (H_1)**. The null hypothesis (H_0) represents the idea of no difference or no effect, suggesting that any observed differences or relationships in the data are due to random chance or sampling variability. On the other hand, the alternative hypothesis (H_1) proposes that there is a true effect or relationship in the population.

The researchers conduct inferential tests of H_0, trying to reject it. If they reject H_0 with high statistical probability, they accept H_1. If they fail to reject H_0, they accept it. In other words, by default H_0 is correct, and the scientist needs to find reliable evidence that it is false to claim that they have discovered something.

Suppose a researcher conducts a study to investigate whether a new teaching method improves students' academic performance compared to a traditional teaching method. The null hypothesis in this case would state that there is no difference in academic performance between the two teaching methods. The researcher assumes that any observed differences in the sample data are due to random chance and not a genuine effect of the teaching method. The alternative hypothesis in this case would suggest that the observed differences in the sample data are not due to chance alone but rather reflect a true effect of the teaching method.

Note: the alternative hypothesis may be directional (one-tailed) and non-directional (two-tailed). When we compare two groups, a one-tailed hypothesis predicts which of the groups will score higher, whereas a two-tailed hypothesis simply predicts that the groups will score differently. The corresponding null hypothesis will be different in each case—see Table 1.11.

Type of alternative hypothesis (H_1)	Example	Null hypothesis (H_0)
Non-directional (two-tailed)	There is a difference in anxiety levels between bank tellers and truck drivers	There is no difference in anxiety levels between bank tellers and truck drivers
Directional (one-tailed)	The anxiety level of truck drivers is higher than that of bank tellers	The anxiety level of truck drivers is not higher than that of bank tellers

▲ Table 1.11 Formulating the null hypothesis

Type I and Type II error rates

When making decisions based on hypothesis testing, it is important to know the concepts of Type I error and Type II error.

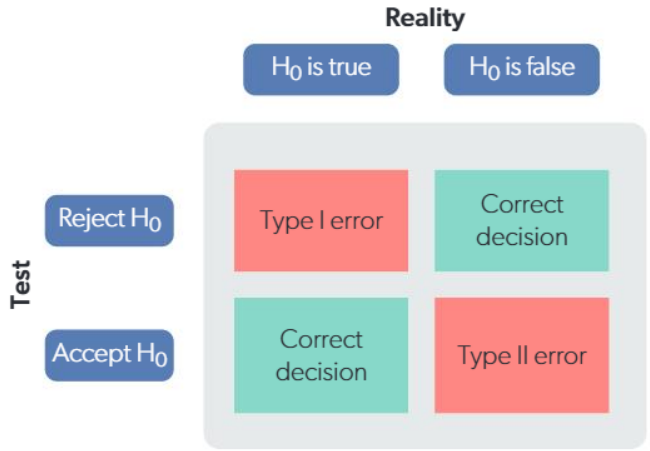
Type I error is also known as a false positive. It occurs when we reject a null hypothesis that is actually true. For example, imagine a pharmaceutical company conducting a clinical trial for a new medication meant to lower blood pressure.

TOK

This logic of hypothesis testing links to the fundamental principle of scientific methodology known as falsifiability (or the falsification principle). Originally suggested by Karl Popper in 1959, this principle currently defines how we draw the line between science and non-science. According to this principle, we can never prove that a theory is correct because no matter how much supporting evidence we find, there will always exist a chance that it is wrong and we just haven't seen the counter-evidence yet. Therefore, it makes no sense to try to prove a theory. On the other hand, even one piece of counter-evidence is sufficient to reject a theory with confidence. Counter evidence is more important than evidence, and we should search for it, trying to reject our theory as best we can. If we try to reject a theory and fail, our confidence in this theory grows.

The null hypothesis (H_0) states that the medication has no effect, while the alternative hypothesis (H_1) asserts that the medication does have an effect. In this scenario, a Type I error would occur if the company rejects the null hypothesis and claims that the medication is effective when, in reality, it is not. This could lead to false conclusions and could potentially harm patients if the medication is widely adopted based on this erroneous claim. So we make a Type I error when we find a difference or a relationship that in reality is not there.

Type II error is also known as a false negative. It occurs when we fail to reject a null hypothesis that is actually false. It represents the probability of missing a true effect or relationship. In our example the pharmaceutical company would commit a Type II error if the medication is actually effective, but the company fails to reject the null hypothesis (H_0), therefore failing to identify an effect or relationship that actually exists. This may result in a missed opportunity to provide a beneficial treatment to patients.



▲ Figure 1.41 Type I and Type II errors

Type I error rate is denoted as α (alpha). This is also typically referred to as the "level of **statistical significance"** of a research finding. The accepted cut-off of $p < 0.05$ means that the probability of Type I error is less than 5%. We cannot eliminate this probability, but we can estimate it, and 5% has been traditionally used as the acceptable cut-off level, although more rigorous levels are also used in research studies: $p < 0.01$ and even $p < 0.001$ (the latter represents a 0.1% probability of Type I error).

Type II error rate is denoted as β (beta). It is also directly linked to the concept of statistical power: 1-Type II error rate. Statistical power is the probability of correctly rejecting a false null hypothesis.

Interestingly, the typically accepted cut-off for Type II error rate is 20% (in other words, statistical power of at least 80%). This means that we accept a 20% probability of not finding a result when one exists, but we accept a much stricter probability of 5% of finding a result when none exists. It makes sense: concluding that a medication is effective when in fact it is not may have quite dramatic consequences. Failing to see that a medication is effective (when it is) is somewhat less consequential.

Understanding and controlling Type I and Type II error rates allows researchers to make informed conclusions and draw accurate inferences from their data.

Parametric and non-parametric tests

Some inferential tests assume that the data is normally distributed while some tests do not make such an assumption. The former are called **parametric tests**, the latter are called **non-parametric**. The name "parametric" derives from the fact that such tests in their formulas use the "parameters" of mean and standard deviation. As you know, the mean only serves as a meaningful indicator of central tendency when the distribution is normal because the mean is heavily affected by skewness and outliers. The same is true for the standard deviation.

Parametric tests can only be used with ratio-level and sometimes interval-level data. This is understandable because the mean and the standard deviation are not meaningful with ordinal data and impossible to calculate with nominal data.

Effect size and statistical significance

When reporting and interpreting results of any inferential statistical test, it is important to mention the following two parameters: effect size and statistical significance. A report that contains only one of these parameters is incomplete and cannot be meaningfully interpreted.

Effect size is a measure of magnitude: how big the difference is or how strong the relationship is. There are several measures of effect size, the two most commonly used being **Pearson's *r*** (Pearson's effect size) and **Cohen's *d***. Where exactly to set cut-off boundaries between a "small" effect size and a "large" effect size is a difficult question, but here are some commonly accepted boundaries (Cohen, 1988):

- Cohen's *d*:
  - 0.2 = small effect size
  - 0.5 = moderate effect size
  - 0.8 = large effect
- Pearson's (*r*):
  - 0.1 = small effect size
  - 0.3 = medium effect size
  - 0.5 = large effect.

Statistical significance is an estimate of the probability that the result was obtained due to random change. In other words, statistical significance is the probability that Type I error has occurred (alpha). The following three levels are usually accepted as cut-off boundaries for interpretation:

1. $p < 0.05$ = statistically significant result
2. $p < 0.01$ = very significant
3. $p < 0.001$ = highly significant

It is important to know that the level of significance is affected by sample size, but the effect size is not. For example, the same correlation coefficient will be more statistically significant in a larger sample and less statistically significant in a smaller sample. This is because in a smaller sample it is more probable to obtain a correlation by chance. Conversely, in very large samples even small correlations may be shown as statistically significant: although they are small, they are unlikely to have been obtained by chance.

This is why it is important to interpret these two parameters together.

Which inferential test is appropriate in a given situation?

Your choice of the inferential test will depend on:

- the experimental design
- the level of measurement
- your assumptions regarding normality of distribution of the DV.

Table 1.12 is an overview of some of the appropriate inferential tests. Note: all these tests are only suitable for situations when the IV has two levels (i.e., you are comparing either two groups or investigating the relationship between two variables).

Measures of difference					
		Unrelated groups (e.g., independent measures design)		Related groups (e.g., repeated measures or matched pairs design)	
Level of measurement	Nominal or ordinal	Chi-squared (χ^2) test		McNemar's test (for dichotomous data)	
		Mann–Whitney *U* test		Wilcoxon signed-rank test	
	Interval or ratio	**Non-normal distribution:** Mann–Whitney *U* test	**Normal distribution:** Unrelated *t*-test	**Non-normal distribution:** Wilcoxon signed-rank test	**Normal distribution:** Related *t*-test
Measures of relationship					
		Non-normal distribution:		Normal distribution:	
Level of measurement	Nominal	Phi correlation		–	
	Ordinal	Spearman's rank correlation		–	
	Interval or ratio	Spearman's rank correlation		Pearson's correlation	

▲ Table 1.12 **Some common univariate inferential tests**

In order to demonstrate how inferential tests work, we will now consider several examples of specific tests. Note that students are not required to know the calculations for exams. You will only be required to interpret the results of statistical tests as presented in the stimulus material. However, understanding some calculation examples will give you a deeper insight into data interpretation.

Measures of difference: the unrelated *t*-test

The *t*-test is a parametric statistical procedure that can be used to test the difference between two groups or conditions. There are several types of *t*-test:

- The **unrelated *t*-test** is used for independent measures designs where a comparison is being made between two groups of participants.

- The related *t*-test is used for repeated measures designs where a comparison is being made between two testing occasions in the same sample of participants.

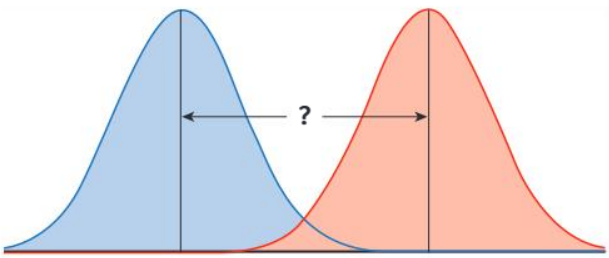
Here is the formula used to calculate the unrelated *t*-test.

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- $\bar{x}_1$, $\bar{x}_2$ = the mean of the first and the second group
- σ_1, σ_2 = standard deviation of the first and the second group
- n_1, n_2 = sample sizes of the two groups.

Looking at this formula, we see that the value of *t* is affected by the three parameters in the following ways:

1. *t* is larger when the observed difference between the means of the two independent groups is larger
2. *t* is larger when the standard deviations in both of the groups are smaller
3. *t* is larger when the sample sizes are larger.



▲ Figure 1.42 Testing the difference between two independent groups

In other words, if we have two large independent samples that show little variance within themselves, but are at the same time considerably different from each other, then the *t* statistic is large. It shows us that the difference between groups is considerably larger than the difference between any two participants within any of the groups: inter-group variance outweighs intra-group variance.

Once we have the *t* statistic, we can calculate its effect size and statistical significance.

Activity

Visualize a *t*-test.

Take a collection of small objects of two types—for example, red marbles and green marbles. They will represent participants belonging to two independent groups. Let the surface of your table represent a scale from 0 to 100. Cluster all red marbles on the right side of the table, so that the mean is approximately 70 and they are not spread too much around the mean (perhaps in the range 60–80). Cluster all green marbles on the left side with the mean approximately 30 and the spread between 20 and 40. Looking at these results, are you confident that the two groups are different?

Now try moving the marbles within the groups further away from each other and discuss how this affects your confidence.

For example, what happens if:

- there is a much larger spread within groups, for example 0–60 and 40–100?
- the spread is mostly the same, but each group has a couple of outliers—for example, some green marbles "scored" 90 and a couple of red marbles "scored" below 20?

ATL Thinking, Communication

To calculate the effect size, a special formula is used to convert it into Cohen's *d*. Remember: there are some conventionally accepted thresholds in the interpretation of Cohen's *d*:

- 0.2 = small effect size
- 0.5 = moderate effect size
- 0.8 = large effect.

To estimate statistical significance, somewhat more complex procedures are used, but the end result of this is an estimate of the probability that Type I error has been made (alpha).

Chat with AI

Generative AI and large language models are very good at explaining statistics and calculations. Use them extensively as your personal tutor to become proficient at data interpretation. Here are some suggested prompt starters:

- I am a psychology student learning about the unrelated *t*-test but I don't fully understand the formula. Could you explain it step by step with a simple example?
- Here are the results of correlational analysis reported in a psychology paper [paste the paragraph reporting the results]. Could you help me understand it—what does it mean? [Remember to ask specific follow-up questions after the AI's first response.]
- What exactly is an effect size and why do we need to calculate it after the *t*-test has already been calculated?

Measures of difference: Mann–Whitney *U* test

For comparison, let us consider a non-parametric test that would be used in the same scenario (investigating the difference between two independent groups), but with variables that are ordinal or non-normally distributed.

The **Mann–Whitney *U* test** does not rely on any assumptions regarding the distribution of the variable, making it non-parametric. Therefore, the test uses neither the mean nor the standard deviation in the formula. Instead, it reduces the level of measurement to ordinal and uses ranks.

For an example, let's take the hypothetical dataset presented in Table 1.10 (see page 66). The dataset was obtained in an experiment where one group of people had to remember a list of words in a noisy environment (the noisy condition) and another group had to remember the same list of words in a quiet environment (the quiet condition).

First, all data is pooled and ranked. This procedure reduces the level of measurement from (in this case) ratio level to ordinal. There are four steps:

- Step 1: grouping. We order all values, regardless of the group, from the smallest to the largest, placing all "ties" in the same lines.
- Step 2: ordering. We assign these values ranks, from the lowest to the highest, ignoring the ties.
- Step 3: ranking. We replace the ranks for the ties with their respective means, so that equal values get the same ranks.
- Step 4: we calculate the mean rank and the sum of ranks in each group.

Step 1 (grouping)		Step 2 (ordering)		Step 3 (ranking)	
Control condition (quiet)	Experimental condition (noisy)	Control condition (quiet)	Experimental condition (noisy)	Control condition (quiet)	Experimental condition (noisy)
	2		1		1
	3		2		2.5
	3		3		2.5
4	4	4	5	5.5	5.5
	4		6		5.5
	4		7		5.5
5	5	8	9	11	11
5	5	10	11	11	11
5	5	12	13	11	11
	5		14		11
6	6	15	16	18.5	18.5
6	6	17	18	18.5	18.5
6	6	19	20	18.5	18.5
6		21		18.5	
6		22		18.5	
7	7	23	24	24.5	24.5
7	7	25	26	24.5	24.5
8	8	27	28	27.5	27.5
9		29		29	
14		30		30	
Step 4:			**Mean rank**	19.04	12.41
			Sum of ranks	266.5	198.5

▲ Table 1.13 Calculations of Mann–Whitney *U* for the hypothetical experiment

Now we have the values, they can be used in a special formula to calculate the test statistic *U*. The idea is that if the two groups really are different, then the mean ranks should be different too. The larger the difference, the larger the Mann–Whitney *U* coefficient will be.

The test statistic *U* in itself is meaningless for interpretation. But it can be used to estimate the effect size and statistical significance.

The most commonly used effect size for Mann–Whitney *U* test is denoted as *r* (rank-biserial correlation coefficient). This coefficient can be interpreted in line with the standard effect size *r* (see page 83):

- 0.1 = small effect size
- 0.3 = moderate effect size
- 0.5 = large effect.

For our data presented in Table 1.13 (on page 87) the following is the result: Mann–Whitney $U = 62.5$, r (effect size) = 0.44, statistically significant for a one-tailed hypothesis at the level $p < 0.05$.

This means that we should reject H_0—that there is no difference between the two groups—and accept H_1—that the quiet condition and the noisy condition are significantly different. From the mean ranks we can see that the number of words recalled by participants in the quiet condition was higher.

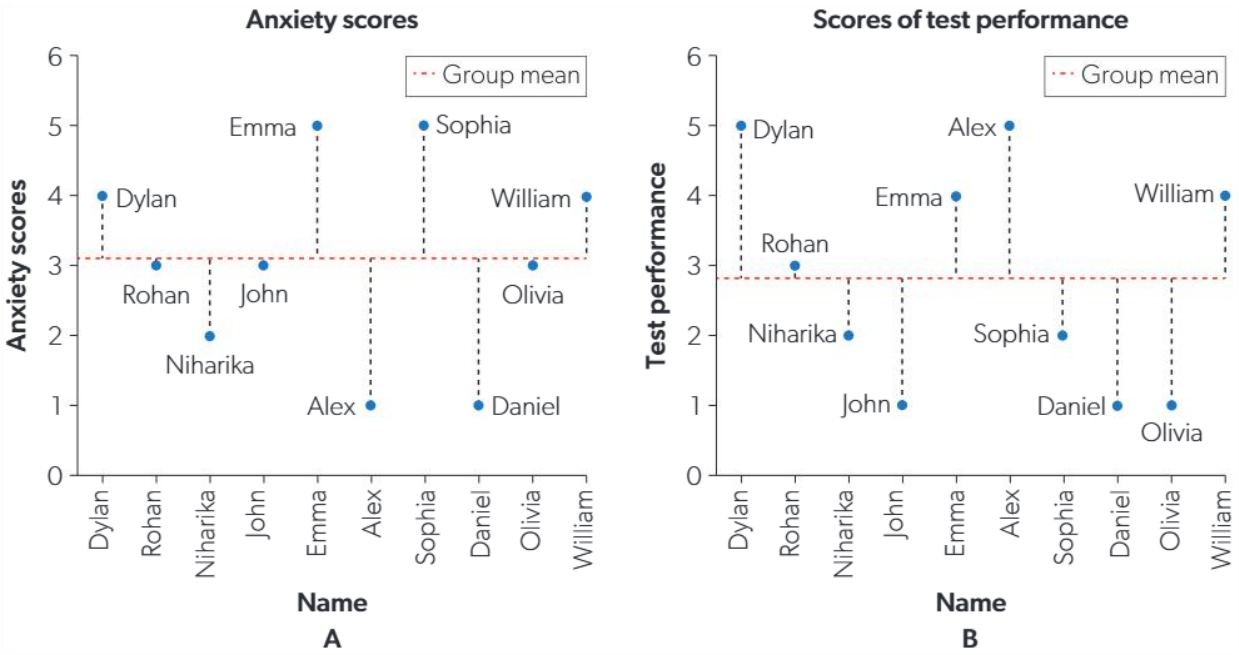
Measures of relationship: Pearson's correlation coefficient

You already know what correlational studies are and how they are different from experiments. To understand how a correlation "works" on the level of calculations, we will look at a simple hypothetical example in which a sample of participants ($n = 10$) was given two measures—anxiety and test performance—both on a scale from 1 to 5 (see Figure 1.43).

For example, looking at panel A in Figure 1.43 we see that Dylan scored 4 on anxiety, Rohan scored 3 and Niharika scored 2. The mean value for all participants is 3.1. This value is shown as the horizontal line on the plot. We will call anxiety our variable x. Then the value of each participant on anxiety can be denoted as x_i (x_1 = the value of the first participant; x_2 = the value of the second participant; … x_{10} = the value of the 10th participant).

We can now calculate each participant's deviation from the group mean. On the plot, we will denote this deviation as a line from the dot to the mean line. An individual participant's deviation from the mean is calculated as $(x_i - \overline{X})$.

For example, Niharika's deviation from the mean is negative: $(x_3 - \overline{X}) = 2 - 3.1 = -1.1$.



▲ Figure 1.43 Hypothetical example: 10 participants scoring on two variables

We can perform the same operations with the second variable (y)—scores of test performance.

The formula of **Pearson's correlation coefficient** r is such that it will be large if participants' deviations on both variables are "coordinated". This would be the case if people who score higher than the mean on one variable also tend to score higher than the mean on the other variable, and those who score below the mean on one variable also tend to score below the mean on the other variable.

Values of r lie within the boundaries from –1 to +1.

Pearson's correlation coefficient r does not require a separate calculation of effect size because it is already an effect size in itself. We can apply the same interpretation boundaries as described above. Obviously, for each r we can also calculate the level of statistical significance (it will also depend on the sample size).

Measures of relationship: Spearman's rank correlation coefficient

One of the popular existing non-parametric equivalents of Pearson's correlation is the **Spearman's rank correlation coefficient**, or Spearman's ρ ("rho").

Similar to other non-parametric statistics, Spearman's correlation deals with the problems of non-normality by reducing the level of measurement to ordinal and ranking the raw data. When data is ranked, the influence of outliers is reduced. For example, the following distribution:

0, 1, 35, 37, 40, 45, 51, 52, 394

will be ranked as:

1, 2, 3, 4, 5, 6, 7, 8, 9

The order of values is preserved, but the distances between data points are ignored.

Remember: when there are ties in the dataset, we assign mean ranks. So:

0, 3, 5, 5, 9, 12, 12, 15

becomes:

1, 2, 3.5, 3.5, 5, 6.5, 6.5, 8.

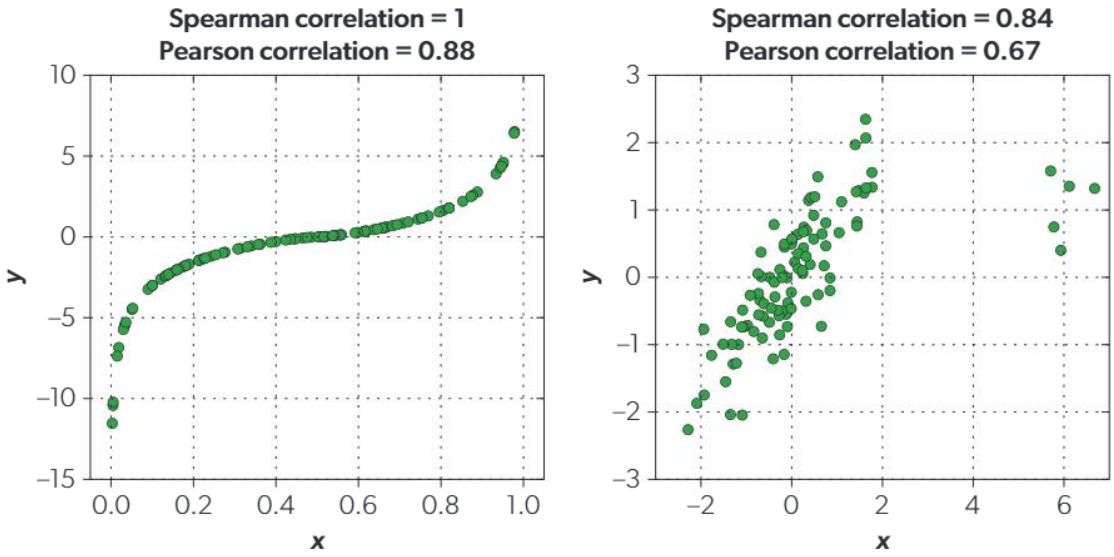
In fact, Spearman's correlation can be calculated by applying Pearson's formula to ranked data.

Some consequences of using ranked values instead of raw values can be seen in Figure 1.44 on the following page. The figure on the left shows a scatterplot that yields a perfect correlation of 1 when Spearman's formula is used, but only 0.88 when Pearson's coefficient is applied. The figure on the right shows that Spearman's correlation is not as affected by outliers.

Chat with AI

Some further suggested prompt starters for you to try:

- Could you give me some specific examples from psychology of when it is preferable to use Spearman's correlation rather than Pearson's? How about the other way around?
- How do you calculate the effect size for Spearman's correlation and how do you interpret it? Keep it simple please.



▲ Figure 1.44 Spearman's correlation and Pearson's correlation in different scenarios

The coefficient does not have to be converted into anything to calculate the effect size—it is an effect size in its own right, interpreted using the same guidelines as those applied to Pearson's *r*.

Reporting and interpreting the results of inferential statistics

When researchers conduct inferential statistical tests to investigate their hypotheses, they must report on the results in a way that is complete and transparent. The report should include descriptive measures (such as means), a measure of effect size, and a measure of statistical significance. It is also good practice to provide an interpretation of the result explicitly. In this section, we will consider four hypothetical examples of research studies that report on the results of one of the inferential tests discussed above.

Note

The sample reports below were created using generative AI and slightly edited. Some of them contain mistakes and inaccuracies. We are intentionally leaving the reports unchanged to illustrate how generative AI can sometimes be incorrect.

Unrelated *t*-test

Sample report:

An unrelated *t*-test was conducted to examine the effect of a mindfulness intervention on levels of anxiety. Participants were randomly assigned to either a mindfulness group ($n = 30$) or a control group ($n = 30$). After the intervention, participants in the mindfulness group showed significantly lower levels of anxiety (mean = 23.50, SD = 4.10) compared to the control group (mean = 26.75,

SD = 3.60), $t(58) = -2.34$, $p = 0.022$, Cohen's $d = 0.70$. The effect size was medium, indicating a moderate reduction in anxiety levels among participants in the mindfulness group compared to the control group.

Comment:

This report contains all the necessary information including descriptive details (means and standard deviations), statistical significance, and effect size. The number 58 in brackets next to the *t* statistic shows the number of degrees of freedom. It is conventional to report this number like this for the unrelated *t*-test. Degrees of freedom are linked to the sample size. In the unrelated *t*-test, they are calculated as the total sample size minus two. The level of statistical significance here is $p < 0.05$—that is, the result can be described as statistically significant, but not "highly significant". This means that the probability that a group difference like this is obtained by pure chance and in reality is equal to zero is <5% (but >1%). Although the effect size is labelled as moderate, it can actually be considered large by the standards that we described earlier. Note: with a large effect size the difference is only moderately statistically significant. This is because statistical significance is sensitive to sample size. The fact that the *t* statistic is negative is in itself meaningless because it depends on which of the groups we consider the "first group" and enter into the equation first. To understand the direction of difference, we need to look at the mean values. This is how we know that the mindfulness group scored lower on anxiety than the control group.

Mann–Whitney *U* test

Sample report:

To examine the differences in job satisfaction between employees in two different departments, a Mann–Whitney *U* test was conducted. The results revealed a significant difference in job satisfaction scores between the two departments ($U = 225$, $p = 0.02$, $r = 0.29$). The effect size was small, indicating a modest difference in job satisfaction levels between the departments. Employees in Department A (median = 72) reported higher job satisfaction compared to employees in Department B (median = 68).

Comment:

This paragraph is only a snippet from a larger report and, if the report follows the best practice, it also includes a rationale for the choice of the Mann–Whitney *U* test. It could be the case that job satisfaction was measured on the ordinal level, or that the dataset contained outliers or other signs of deviation from normality. The report contains the required elements: descriptive statistics for the two groups (medians), statistical significance, and effect size. Note that *r* in this case is the effect size—the rank-biserial correlation. Its interpretation is similar to the standard effect size *r*, so the value of 0.29 can indeed be considered a small-to-medium effect size. The level of statistical significance is $p < 0.05$, but the authors decided to report it precisely, which is acceptable. We are dealing with a group difference that is statistically significant, but the effect size is small—this is a sign that the sample was large. Small differences in large samples can be statistically significant (i.e., reliably different from zero). It would be better if the report explicitly contained an indication of the sample size for both groups.

The *U* statistic in itself does not tell us which of the departments (A or B) had higher job satisfaction—for this, we need to look at the descriptive statistics. The report chose to present the medians (72 and 68). However, we know it would be more appropriate to present mean ranks instead.

Pearson's correlation coefficient

Sample report:

Pearson's correlation coefficient was used to assess the relationship between hours of sleep and academic performance among college students ($n = 120$). The analysis revealed a significant negative correlation between hours of sleep and academic performance ($r = –0.32$, $p < 0.05$). The effect size was moderate, indicating that as the number of hours of sleep decreased, academic performance tended to decline. These findings suggest that sleep plays a role in college students' academic success.

Comment:

Pearson's correlation coefficient is based on certain assumptions such as normality of distribution. These assumptions should be checked prior to conducting the analysis, with the results of this check included in the report. The report actually contains a mistake in data interpretation: a negative correlation means that with the increase in one variable the other variable decreases. In this case, a negative correlation means that the more students sleep, the worse their academic performance, and the less they sleep the better they study (remember: this is a hypothetical example). The conclusion should say instead: as the number of hours of sleep decreased, academic performance tended to improve. The effect size is indeed moderate and the effect is statistically significant.

Spearman's correlation coefficient

Sample report:

In order to investigate the relationship between self-esteem and depression levels in a non-normal distribution, Spearman's correlation coefficient was utilized. The results showed a strong negative correlation between self-esteem and depression (ρ ("rho") $= –0.62$, $p < 0.001$). The effect size was large, indicating that individuals with higher levels of self-esteem tended to have lower levels of depression. These findings highlight the importance of self-esteem in mental well-being and suggest that interventions targeting self-esteem may have potential benefits in reducing depression symptoms.

Comment:

If the distribution of data collected in the study deviates from normality, Spearman's correlation is indeed an appropriate choice of test. With depression, it is expected that the distribution of depressive symptoms in a population will not necessarily be normal, but it is somewhat strange that a measure of self-esteem deviates from normality. An explanation of this would be expected in the report prior to describing the results of the study. The correlation obtained in the study is large (effect size of ρ is interpreted similarly to Pearson's *r*). The rest is also highly statistically significant with the probability of obtaining such a correlation due to

random chance of <0.1%. It would be good practice to include information about the sample size and descriptive parameters (median and interquartile range) of each variable—these are missing. Additionally, the report contains an inaccuracy in interpretation: while it is fine to interpret results of a correlational study as suggesting that self-esteem "plays a role" in mental well-being, it is not justified to conclude that targeting self-esteem will reduce depressive symptoms. Such an interpretation implies a cause–effect relationship between the two variables, for which we have no evidence in this study.



Activity

Generate more sample reports using AI and write a short analysis such as the one presented here.

Exchange your work with a partner and give each other feedback.

Ask your teacher to conduct an overall check.

ATL Research, Communication, Self-management

1.13 Analysing qualitative data

Inquiry questions

- How can we analyse qualitative data such as texts?
- Is there a clear-cut distinction between qualitative and quantitative analysis in real-life research?

What you will learn in this section

Key learning:

- Quantitative analysis of qualitative data: content analysis.
- Qualitative approaches to analysis: thematic analysis (also known as inductive content analysis), grounded theory.

Key terms: content analysis, thematic analysis (inductive content analysis), grounded theory

Overview

Qualitative data comes in the form of texts. These could include interview transcripts, responses to open-ended questions on a survey, field notes from an observation, think-aloud protocols, and written diaries.

Analysis of such data rests on principles that are very different from statistical processing of numerical information. However, the distinction between qualitative analysis and quantitative analysis is not absolutely clear, and it is common for research studies to use elements of both.

Here we will consider three main groups of methods: content analysis, thematic analysis (also known as inductive content analysis), and grounded theory. Note: the first method (content analysis) is actually a quantitative method of analysis applied to qualitative data. It is listed here to provide a comparison point to better understand the other two qualitative methods of data analysis.

Content analysis

Content analysis follows the same standard of hypothesis testing as experimental and other quantitative research. Categories and indicators are defined in advance based on theory. Researchers then count the frequency of words or word combinations related to each category. For example, suppose we study participants' perceptions of a health prevention programme that they were involved in. We could ask them an open-ended question such as "In general what did you think about the programme?" and collect responses. Then, in each

response, we could count the number of words that have a positive emotional colouring (connotation) and the number of words with a negative connotation, based on pre-existing lists of words. Variations of this procedure are known as sentiment analysis.

Note: in content analysis:

1. we are analysing elements of language (e.g., words and combinations of words) rather than the meaning behind them
2. there is no interpretation on the part of the researcher
3. the categories of analysis are defined in advance
4. the result of analysis is numerical data that can be further subjected to statistical analysis, including inferential statistical tests
5. this method of data analysis is quantitative.

Thematic analysis

Thematic analysis (also known as **inductive content analysis**) is the most basic method of analysing qualitative data. In thematic analysis, the researcher reads the texts (e.g., interview transcripts) and identifies the themes that are being mentioned. The researcher tries to stay as close to the text as possible. However, identifying a theme still requires an element of researcher's interpretation. For example, suppose the participant said the following, "I felt understood when the therapist expressed understanding and encouraged me to continue talking" and the researcher coded this statement as "supportiveness of the therapist".

There is a procedure that researchers are expected to follow. Here are the main steps outlined in the procedure (Elo and Kungäs, 2008):

1. Writing the transcript. There are two types of transcript: verbatim or post-modern. Verbatim transcripts are word-for-word accounts of everything the participant said. Post-modern transcripts include notes about the intonation, gestures, and other non-verbal elements in the participant's behaviour.
2. Reading the raw material several times and identifying initial themes. This is done iteratively. Researchers start with low-level themes, trying to stay as close to the text as possible. When the first reading is done, a set of initial themes is identified and may be written on the margins. The second reading is done and the themes are confirmed (and revised); new themes may also be added. This is done several times. Sometimes independent coders are used to check the credibility of deriving low-level themes from the text.
3. Low-level themes are grouped into a smaller number of high-level themes. This grouping involves an element of interpretation on the part of the researcher: they need to decide if X, Y, and Z belong to category A. As a credibility check, other researchers may be involved in the process so that results of grouping can be compared across researchers. The result of this stage of analysis is a manageable set of high-level meaningful units that summarize the transcript.

TOK

What is the difference between induction and deduction? If you do not remember, look it up. What are some of the specific examples of inductive reasoning and deductive reasoning that you have come across so far in the IB Psychology course?

4. A summary table of themes is prepared. The table lists all the high-level emergent themes, all the lower-level themes within them, and supporting quotations from the raw transcript. The structure of themes can also be revised slightly at this point to account for parts of the transcript that are still unexplained. Data saturation is reached when subsequent readings of the transcript do not lead to identifying any new themes.
5. Finally, conclusions are formulated based on the summary table. As a credibility check, participants may be shown the results of the analysis and asked to confirm the emergent themes as well as the derived interpretations.

The resulting analysis may be accompanied by "memos" that explain to the reader how and why certain analysis decisions were made, increasing the "thickness" of descriptions (which, as you know, increases credibility).

Inductive content analysis can also be applied to observational data. In this case, the raw material for analysis comes in the form of field notes describing a participant's behaviour rather than interview transcripts.

The result of thematic analysis is a list of "themes" that emerge from the data and capture participants' responses in a holistic way. Each theme is usually accompanied by examples of what participants actually said (quotations).

Sometimes thematic analysis incorporates elements of quantitative analysis. This occurs when researchers count the number of times a particular theme was mentioned in participant responses. However, the focus of the research is still the themes themselves, and frequency information is usually used as supplementary in order to help researchers make decisions about including, excluding, or rewording the themes.

Exam tip

Remember: qualitative and quantitative research are based on different paradigms and pursue different purposes. It would not be appropriate to use evaluation categories of one of these approaches to judge the quality of the other. For example, be careful with criticizing sample-to-population generalizability of qualitative research: this kind of generalizability is not the point.

Grounded theory

Grounded theory is a complex methodology of qualitative research. It is actually not limited to the analysis of qualitative data—it also dictates how the data should be collected in the first place. As the name suggests, grounded theory is a theory that is constructed based on data, "grounded" in evidence. This approach was suggested as a result of critique of the standard scientific method where the researcher first "constructs" a theory, then tests it against data.

Research under the framework of grounded theory would use the following approach:

1. Data is gathered and analysed by methods such as thematic analysis, resulting in the identification of emerging themes.
2. The researcher tries to formulate a generic theory that explains the observations, remaining as close as possible to the data.
3. The theory then dictates what other data to look for. For example, it could be that a category emerged, but it is not clear because participants did not speak about it too much. Then we want to gather more data to clarify that category. We might want to sample some participants who are likely to bring it up.

Activity

As part of your class practicals, you will perform both an interview and an observation. You will get some text data as a result of this: either interview transcripts or observation notes. Practise applying the steps of thematic analysis (outlined above) to your data. There is nothing as powerful in learning about data analysis and interpretation as trying it on your own dataset.

ATL Research, Self-management

4. Additional sampling is conducted to gather more data, in accordance with the previous step. This is the case of using theoretical sampling—the type of purposive sampling that is driven by theoretical analysis and the desire to clarify certain aspects of the emerging theory.
5. Additional data gathering and analysis continues until the point of saturation. It is said that data saturation has been reached if adding new data (e.g., interviewing new participants) does not contribute any more to the refinement of the theoretical categories. This is when we can conclude that the grounded theory is now strong enough to account for all possibilities.



2 Learning and cognition

Introduction

Generally speaking, learning is a change in behaviour that occurs as a result of experience. The opposite of learned behaviour is innate behaviour. Different species will have different proportions of learned and innate behaviours, but more intelligent species typically have a larger proportion of learned behaviours.

Regarding innate behaviour, it is amazing how many behaviours that appear purposeful are actually mindless patterns that follow a genetic script.

A special kind of innate behaviour is a fixed action pattern. A fixed action pattern is a complex sequence of actions that gets triggered by a specific stimulus.

Researcher M.W. Fox (1974), who studied the behaviour of turkeys, observed that maternal instincts in these creatures were triggered by a specific "cheeping" sound made by their babies. It is the "cheep–cheep" sound that launches a complex sequence of nurturing actions. If the chicks produce this sound and the mother hears it, she will take care of them and feed them. If the chick stops producing the sound or the mother does not hear it for some reason, she is likely to ignore the chicks and even, in some cases, kill them.

Another well-known example of a fixed action pattern is the greylag goose. If an egg rolls out of its nest, the goose will push the egg back into the nest with its bill, in a fixed sequence of movements. This behaviour is triggered by the sight of an egg outside of the nest. However, if you remove the egg that has rolled out, the goose will still perform the same actions, as if pushing an imaginary egg. If you quickly replace the egg with a golf ball or even a larger egg-shaped object like a volleyball, the goose will push the ball into its nest. Moreover, when given a choice between the egg and a volleyball, they will actually leave the egg and bring back the volleyball (presumably because it is larger)—this is published in a classic study by Konrad Lorenz and Niko Tinbergen (1970).

Innate behaviours are triggered by some environmental stimulus (also known as a releaser): they are rigid and stereotypical, and they are genetically hard-wired and automatic. When organisms are born, they are already capable of innate behaviour and there is no need for training.

The ability to learn new behaviours provides a huge evolutionary advantage. It allows an organism to rely less on the rigid genetically pre-programmed patterns that may no longer be suitable in a new environment.

In contrast to innate behaviours, learned behaviours involve changes due to life experiences. When born, the organism is not capable of these behaviours.

Abbott and Wong (2008) experimented with planaria (*Dugesia tigrina*), otherwise known as freshwater flatworms. These creatures are the favourites of many researchers because their nervous system is simple and well-defined, allowing one an opportunity to explore how the simplest forms of behaviour arise from the simplest forms of brains. There is a knot-like structure at the end of the flatworm's nervous system which is often considered to be a very primitive brain. There is not much you can do with a flatworm in an experiment, but you could train it to go through a Y-shaped maze. This is simply a corridor that ends with a fork: the

flatworm can either go right or go left. So what happens if you give the flatworm a treat every time it chooses to go left and punish it with an electric shock every time it chooses to go right? An intelligent creature should quickly learn that going left is the correct choice. How intelligent is the freshwater flatworm?

Results showed that flatworms can indeed learn from negative reinforcements (electric shocks), but their memory is short-lived. If a flatworm has a natural tendency to go right, it will go right slightly less frequently once it starts receiving an electric shock for doing so (but it will still go right on some occasions!). However, if you pause the experiment for a couple of days, the flatworms will quickly forget the lesson and revert back to their original preferences.

This shows how powerful innate behaviours are in the life of a flatworm, although some form of learning also occurs.

Learned behaviour is obviously the main focus of psychology. The key types of learned behaviour are:

1. Classical conditioning. If you have heard about Pavlov's dogs, then you have heard about classical conditioning. We will consider classical conditioning in detail in this chapter.
2. Operant conditioning. This is also sometimes referred to as "trial-and-error learning", although the terms are not exactly the same. Again, we will consider operant conditioning in detail in this chapter.
3. Cognitive learning. This requires constructing some kind of internal representation of the environment. In conditioning, learning is believed to happen independent of the "mind": it is an automatic association between environmental stimuli and behavioural reactions. Cognitive learning considers all sorts of intermediate variables that act in the mind: memory, thinking, self-regulation.

Cognitive learning describes the highest forms of learning and, of course, the majority of topics in this chapter will be related to cognitive forms of learning.

Here is an overview of the topics in this chapter as presented in the IB DP Psychology Guide:

Thinking and learning	Cognitive processes
• Cognitive biases • Schema theory • Conditioning (classical and operant) • Dual processing model • Social learning theory	• Biological factors in cognitive processes • Cognitive models • Cultural factors in cognitive processes • Environmental influences on cognitive processes • Potential for improving a cognitive process

Note that in this book, the order of topics within the chapter will be changed to ensure a logical sequence of learning.

2.1 Classical and operant conditioning

Inquiry questions

- Is behaviour driven by its consequences?
- Can pigeons be superstitious?
- Is classical conditioning or operant conditioning a better description of intentional behaviour?

What you will learn in this section

Key learning:

- Behaviourism suggests that the focus of research in psychology should be on observable behaviour rather than unobservable constructs. It dismisses cognitive variables as a "black box" that is not worthy of scientific investigation.
- Classical conditioning follows the principle "stimulus–reaction". Repeated pairing of a conditional stimulus with an unconditional stimulus leads to the formation of an association and a conditioned response. Ivan Pavlov discovered a number of laws of classical conditioning, such as generalization of stimulus.
- One might say that in operant conditioning causality is reversed, with reactions first and stimuli to follow. The starting point in operant conditioning is always operant variability—naturally occurring spontaneous variations of behaviour.
- Thorndike's law of effect states that behaviours are a result of their consequences. Types of consequences in behaviourism include punishment and reinforcement, either positive or negative. Principles of conditioning have been applied to explain almost every aspect of human behaviour.
- Objective and reliable measurements are very important. Behaviourists focused on observable and objectively registrable reactions, often using experimental "chambers" for controlled conditions.
- Conditioning provides a tool for controlling behaviour. This raises multiple concerns about how such a tool may be used, and the ethics surrounding this issue.

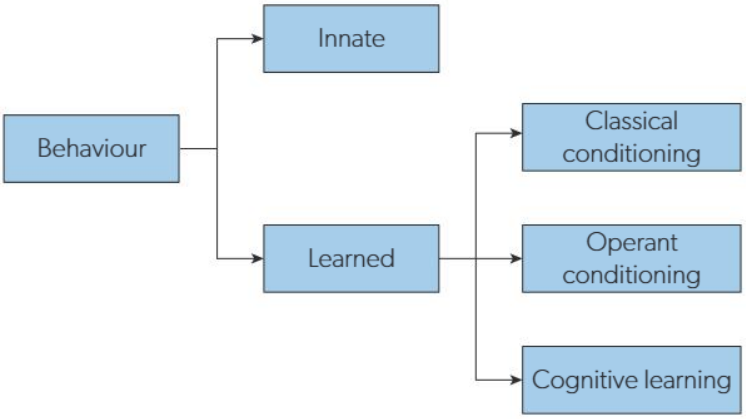
Key terms: behaviourism, classical conditioning, operant conditioning, cognitive revolution, "black box", unconditioned stimulus/response, conditioned stimulus/response, generalization of stimulus, law of effect, punishment and reinforcement (positive and negative), extinction, operant conditioning chamber, radical behaviourism, operant variability

In a wider context

Conditioning is a type of learning. Learning is a change in behaviour that occurs as a result of experience. Learned behaviour is the opposite of innate behaviour.

Innate behaviours are triggered and regulated by a genetic programme. They are rigid and cannot be changed. An organism is born with a repertoire of innate behaviours already hard-wired into its brain.

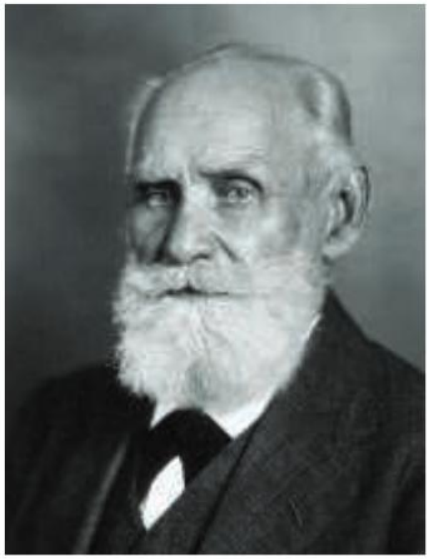
Learned behaviours are acquired during an organism's lifetime. They make the organism better adapted to its environment. There are various types of learned behaviours, such as classical conditioning, operant conditioning, and cognitive learning.



▲ Figure 2.1 Types of behaviour in relation to learning

The history of behaviourism

To truly understand the significance of conditioning in psychology, it is necessary to acknowledge the history of this idea. **Classical conditioning** was discovered by Ivan Pavlov in 1897. Operant conditioning originated in the works of Edward Thorndike and John Watson at the start of the 20th century, but was extensively developed and came to be associated with the name of Burrhus Frederic Skinner (B.F. Skinner) who laid out his research programme in his book *The Behavior of Organisms: An Experimental Analysis* published in 1938.



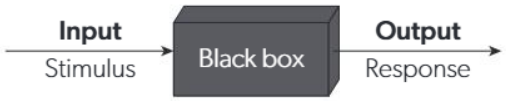
▲ Figure 2.2 Ivan Pavlov

At that time, psychology was dominated by speculative theories based on subjective methods. These included observing one's own mental processes (introspection) or a psychologist's subjective interpretation of the meaning of a patient's dream. This was in stark contrast to the rigorous objectivity of other sciences such as physics and chemistry.

As a solution to this problem, Skinner, Watson, and others started a movement known as **behaviourism**, which quickly became the most influential paradigm that dominated psychology for decades.

Behaviourism suggests that the focus of research in psychology should be on observable behaviour rather than unobservable constructs such as "emotion", "memory", or "motivation". If we cannot observe and measure something in an experiment, then it should be dismissed as something that is not worthy

of investigation. Behaviourists introduced the metaphor of a "**black box**" to describe this inability to observe or measure. A participant in an experiment is exposed to certain stimuli that can be observed and registered; the participant then demonstrates observable behaviour that can also be objectively registered. The goal of psychology, according to behaviourists, is to study associations between these two variables. There is probably something happening in the "mind" after the stimulus is received and before the behavioural response is produced, but we cannot look inside this "black box". Therefore, we should dismiss all speculations about it as useless non-scientific constructs.



▲ Figure 2.3 The "black box" metaphor

The great advantage of behaviourism was that it allowed psychology to firmly establish its status as a rigorous science.

The reign of behaviourism in psychology ended with the onset of the "**cognitive revolution**" in the 1950s: an explosion of studies focused on the contents of the "black box".

Classical conditioning

Ivan Pavlov was a Russian physiologist who conducted well-controlled experiments with dogs and published his first findings back in 1897.

SAQ Classical conditioning

There are a few terms that Pavlov introduced that are important to know in order to fully understand the mechanism of classical conditioning:

- An **unconditioned stimulus** (US) is an environmental stimulus that is biologically potent, for example, the taste of food. You do not have to teach a dog that "food is good": it is a biologically predetermined knowledge.
- A **conditioned stimulus** (CS) is an environmental stimulus that is neutral and that by itself does not produce any biologically predetermined reaction. For example, the sound of a bell ringing does not carry any information that would be biologically important to a dog.
- An **unconditioned response** (UR) is a reflex response to an unconditioned stimulus. For example, a dog will naturally salivate at the sight or smell of red meat. Red meat in this example is the unconditioned stimulus and salivation is the unconditioned response to this stimulus.

Pavlov's idea was that, if you repeatedly present the meat (US) and the bell (CS) several times, you will be able to remove the meat and the dog will salivate at the sound of the bell. Such a learned response was called a **conditioned response** (CR). Unlike the UR which is hard-wired into the brain as a biological reflex, the CR is acquired through experience.

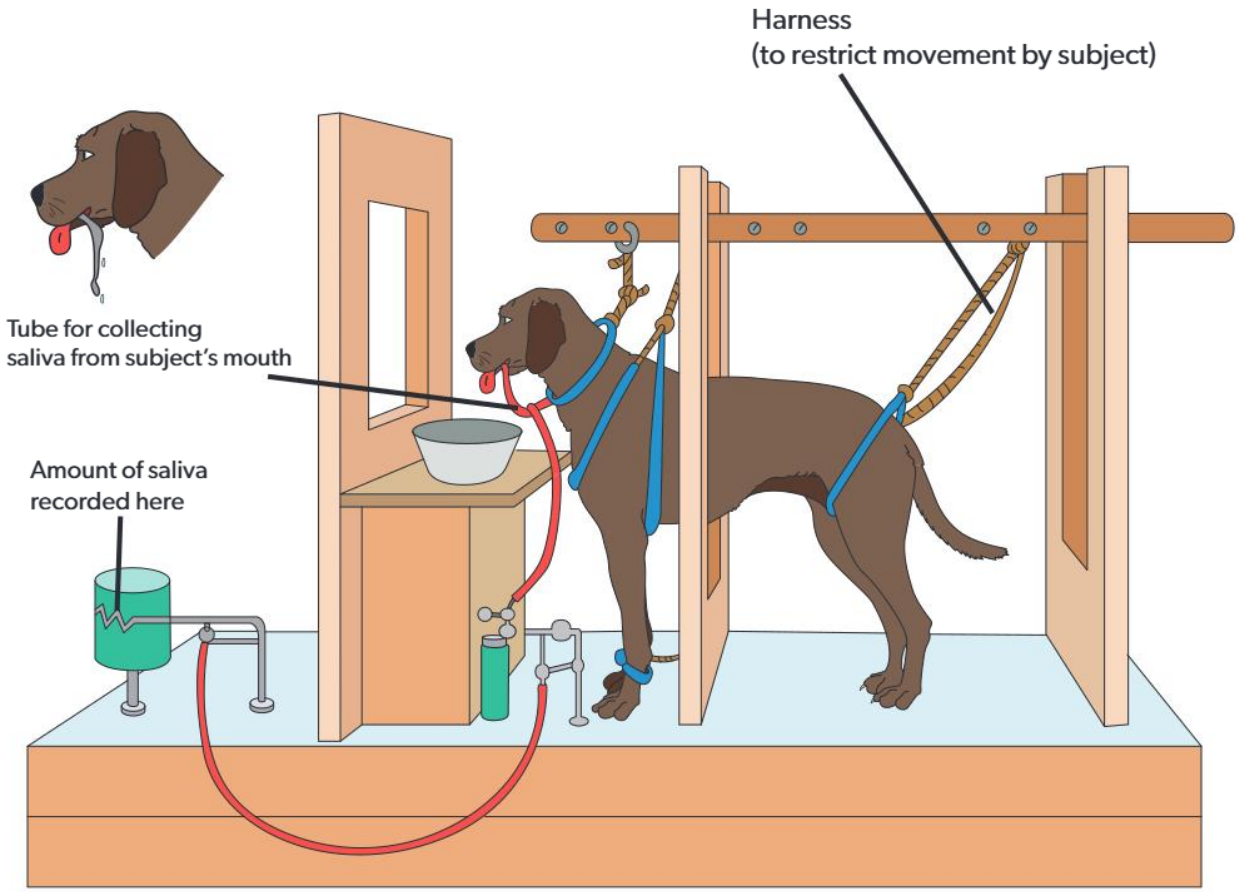
TOK

Physics envy is a term suggesting that many areas of knowledge outside of natural sciences are perceived as insufficiently objective, insufficiently scientific, insufficiently rigorous, or otherwise inferior. They want to be like physics.

To what extent do you think physics envy is justified? To what extent does it apply to psychology?

In his studies, Pavlov used dogs whose digestive fluids were redirected outside the body. The dogs were secured with harnesses so that they could not move. In front of them there was a surface on which a food bowl would be positioned, either paired or not with the sound of a metronome (or a bell). The dog's digestive tract was surgically modified so that saliva, instead of going down into the stomach, was redirected to a measurement device.

Pavlov experimented with how many times the two stimuli have to be paired to form an association. He discovered that in some forms of behaviour (such as fear conditioning) even one repetition may be enough. A classic study that is related to this and commonly brought up as one of the examples of unethical studies in psychology is "Little Albert" conducted by Watson and Rayner (1920) in which a nine-month-old boy was conditioned with frightening loud sounds to respond with fear and crying to objects such as a white rat or a hat.



▲ Figure 2.4 Pavlov's experimental setup

Activity

Find out more about the Little Albert study. Do you agree that the study was unethical? Was it justified at the time when it was conducted?

ATL Thinking, Self-management

Further research led to the discovery of several related phenomena, such as **generalization of stimulus**. This is when an organism, after being conditioned to respond with a certain behaviour to a certain conditional stimulus, starts responding with this behaviour to other similar stimuli. In Pavlov's research it was shown that, after dogs have been conditioned to salivate at the sound of a bell, they would also salivate at similar sounds such as a buzzer. In the study of Little Albert, the boy's fear of white rats also generalized to other stimuli, such as rabbits, other white fluffy objects, and even Santa's beard.

Operant conditioning

SAQ Operant conditioning

Operant conditioning is a learning process in which an association is made between a behaviour and its consequence (such as reward or punishment). When considered from the perspective of the learner, it is also referred to as "trial-and-error learning".

For example, consider a rat that learns its way through a maze. In trial-and-error learning, the rat demonstrates a particular behaviour (e.g., it chooses a specific turn in a maze) and the consequence follows (e.g., the rat finds a treat or gets an electric shock). Depending on the consequence, certain behaviours get reinforced (e.g., if the rat was shocked by electricity, it will not be as likely to choose the same turn in the maze next time). Gradually, repetition after repetition, what used to be random activities are directed into a behavioural pattern. The starting point of learning in operant conditioning is a pool of naturally varied behaviours, but then—through punishments and reinforcement—there is a "natural selection" process through which some behaviours become extinct while others become more prominent.

Compared to classical conditioning, operant conditioning may be said to operate in the "opposite direction". In classical conditioning, the stimulus occurs first and the behaviour second. In operant conditioning, this relationship is reversed: behaviour occurs first and then some form of stimulus (reinforcement) follows.

The idea of operant conditioning was born in the work of Edward Thorndike who formulated the **law of effect**: behaviours are a result of their consequences, either satisfying or discomforting. B.F. Skinner was the one who developed these ideas to completion, getting rid of the speculative constructs that describe the "black box".

All consequences of behaviour in the theory of operant conditioning are divided into **reinforcements** (that increase behaviour) and **punishments** (that decrease behaviour). Both can be **positive** or **negative**. Positive reinforcement is when a reward is given (e.g., a rat gets a food pellet after it presses a lever). Negative reinforcement is when an aversive stimulus is removed (e.g., a loud unpleasant noise stops when a rat presses a lever). Positive punishment is when a negative consequence is introduced (e.g., an electric shock). Negative punishment is when an attractive factor is removed (e.g., parents not allowing you to play video games after you've misbehaved).

When a previously conditioned behaviour is no longer reinforced or punished, **extinction** occurs. During extinction, behaviour gradually becomes less and less probable. Sometimes extinction may be a slow process and behaviour may persist for a long time even though it is no longer reinforced. Interestingly, even brief occasional re-introduction of reinforcement during the extinction period may considerably delay the process.

To carefully register all variables in his experiments, Skinner invented the "**operant conditioning chamber**" (now more commonly known as the Skinner box). It was an isolated environment for a small animal (like a rat or a pigeon).



▲ Figure 2.5 Skinner box

The chamber was equipped with a means of delivering reinforcements and punishments (such as a food-dispensing mechanism). It had transparent walls for observation and a "cumulative recorder" that produced a graphical record of the frequency of observed reactions. All experiments followed a well-defined "reinforcement schedule", which outlined what reinforcements should be delivered at what interval, frequency, and intensity.

Study of operant conditioning: superstition in pigeons

In 1948, Skinner carried out a study to investigate superstition in pigeons. His participants were eight pigeons. They were brought to an experimental cage for a few minutes each day. He kept the pigeons hungry (at 75% of their normal weight). The cage had a food-dispenser mechanism (a food hopper) attached which was engineered in such a way that it would swing into place at regular intervals.

Remember: the food was dispensed at regular intervals, and the key point here is that presentation of food occurred no matter what the bird was doing at that moment. Since there was no actual relationship between the pigeon's behaviour and the presentation of food, Skinner expected that whatever behaviour the pigeon happened to be demonstrating when the food hopper swung in place would be reinforced.

Skinner observed that in six out of eight cases, the responses were very clearly defined (so much so that independent observers agreed perfectly in their evaluations). This is how he described them:

"One bird was conditioned to turn counter-clockwise about the cage, making two or three turns between reinforcements. Another repeatedly thrust its head into one of the upper corners of the cage. A third developed a 'tossing' response, as if placing its head beneath an invisible bar and lifting it repeatedly. Two birds developed a pendulum motion of the head and body, in which the head was extended forward and swung from right to left with a sharp movement followed by a somewhat slower return... Another bird was conditioned to make incomplete pecking or brushing movements directed toward but not touching the floor." (Skinner, 1948a)



▲ Figure 2.6 Skinner's pigeons

Why superstition? Skinner explains that "the bird behaves as if there was a causal relation between its behavior and the presentation of food, although such a relation is lacking" (Skinner, 1948a). This is indeed the essence of superstition. Skinner compares the behaviour of his pigeons in this experiment to that of a card player engaging in various rituals to change their luck or a 10-pin bowler who, after releasing the ball down the alley, continues to twist and turn their arm and shoulder as if still controlling the ball.

Activity

Consider the following statements about the behaviour of a pigeon in an experiment: "The pigeon pecked at the disk several times but was distracted by a loud sound. It continued pecking after hesitating for two to three seconds". This kind of language would not be accepted by a behaviourist: it is full of inferences and assumptions about mental states.

A more appropriate statement would be: "In the reported period the pigeon pecked at the disk five times, stopped immediately after the presentation of the noise, remained motionless for three seconds, then continued pecking with the same frequency and intensity".

But of course, we find it hard not to use inferences because we are so used to them in daily life.

To illustrate this, watch the video created by F. Heider and M. Simmel in 1944 as part of their "experimental study of apparent behaviour" (Heider and Simmel, 1944). This is an animated film showing simple abstract shapes moving with no specific pattern. They asked participants to watch the film and describe what they saw.

Here is the beginning of a description one participant wrote: "A man has planned to meet a girl and the girl comes along with another man. The first man tells the second to go; the second tells the first, and he shakes his head. Then the two men have a fight, and the girl starts to go into the room to get out of the way and hesitates and finally goes in…" (Heider and Simmel, 1944).

Can you rewrite this description in a scientific way, focusing on pure behaviour, and avoiding inferences?

To find the video online, simply search for "Heider–Simmel animation".

ATL Research, Communication

Applications of operant conditioning

Skinner is often referred to as a "radical behaviourist", meaning that he took the idea of control to its extreme. **Radical behaviourism** is the belief that even the most complex human behaviour could be reduced to operant conditioning and therefore potentially controlled. A quote that is popularly assigned to him is "Give me a child and I'll shape him into anything". In 1948 he published a fictional novel, *Walden Two* (Skinner, 1948b), proposing a kind of utopian society where people are shaped into good citizens through behaviour modification (rewards and punishments).

Discussion

Would you consider using an "air crib" if you were a parent?

ATL Self-management

Skinner is also famous for his controversial invention of an "air crib" in 1944 (Joyce and Faye, 2010). It was a metal crib with a transparent safety-glass pane at the front. It allowed parents to regulate temperature and humidity inside, and clean air was filtered into the device from below. Skinner used this crib with his second daughter (Deborah) for the first two years of her life. He thought the crib was advantageous for both the baby and the parents. Keeping the child warm (which was vital in the harsh Minnesota climate where Skinner lived and worked) usually meant wrapping the baby in clothes and blankets. Skinner believed that this would have the negative consequence of constraining the child's movement and lead to overheating, and that his crib was the solution to the problem.

The crib was commercially produced, with an estimated 300 children raised in them. However, the device did not catch on. The American public's perception was negative. The newspaper article that provided the first coverage was entitled "Baby in a Box". This might have created an association with the more well-known "Skinner box" for pigeons. Skinner himself used the word "experiment" when describing his invention. Together, all of these factors repelled the public from using the invention, leading to a strongly voiced opinion that science and technology should not come in the way of the loving labour of a child's mother.

Conceptual analysis

Perspective

Classical and operant conditioning represent different perspectives on learned behaviour.

Classical conditioning operates on the principle "stimulus–reaction". We introduce a new unconditioned stimulus and condition the organism to develop an automatic, reflex-like response to it. Operant conditioning operates on the principle "behaviour–reinforcement". We wait for the desirable behaviour to appear by itself, then we reinforce it, thus increasing its probability in the future. One might say that in operant conditioning the process is reversed: first reaction, then stimulus. In fact, Skinner believed that operant conditioning is better than classical conditioning at explaining human behaviour because it is well suited for the study of *intentional* behaviour.

Causality

This brings up the problem of causality. What is the cause of behaviour? What does "intentional" mean from the perspective of operant conditioning?

An interesting idea in behaviourism is **operant variability**. As you know, reinforcement is introduced after a desirable behaviour appears. But why does the behaviour appear in the first place? Something must have caused it. Operant variability suggests that organisms just engage in a variety of explorative behaviours for no apparent reason (strangely enough for behaviourism, this idea is reminiscent of free will). For example, a pigeon left alone in a cage would move to different corners, peck at various objects in the cage, and perform other actions simply due to the natural variability of behaviour.

There does not have to be a teacher or an experimenter in operant behaviour: it unfolds naturally through trial-and-error exploration of one's environment. Quite simply, if you try something and the consequences are rewarding, you will probably try it again.

Bias

Experiments in classical and operant conditioning are probably the most well-controlled experiments ever conducted in psychology. It was important for behaviourists to perform experiments to the same standards that would be expected from an experiment in physics or chemistry. An example is the Skinner box which was invented to provide the researcher full control over all variables that could potentially affect the behaviour of the pigeon.

However, one key disadvantage of behaviourism is that it ignores cognitive variables. We might say that behaviourism created a biased understanding of behaviour as driven solely by external stimuli and reinforcements, as something that happens automatically in response to these stimuli.

Measurement

In today's psychology, we widely accept the idea of constructs and operationalizations. Constructs are hypothesized, directly unobservable phenomena, such as motivation, intelligence, attachment, emotions, and cognitive processes. Operationalizations are observable behaviours that we use to measure constructs.

The fact that we acknowledge the importance of clearly observable, well-defined operationalizations in the measurement of constructs can be traced back to the influence of behaviourists and their research on conditioning. However, behaviourists themselves rejected constructs completely. For them, it was the operationalization itself that was the focus of inquiry, and they saw no point in speculating about the invisible internal forces that might provide an explanation for observed behaviour.

Change and responsibility

Principles of conditioning give us a powerful weapon to potentially influence—or control, or manipulate the behaviour of others.

The story of Skinner's "air crib" highlights the complex nature of the interaction between scientific discovery and public perception. This relates to the concepts of change and responsibility. If we can do something, it does not mean we should do it. Apart from the already complex task of making valid scientific discoveries, researchers are responsible for choosing the way in which such discoveries will be communicated to the public.

2.2 Schema theory

Inquiry questions

- How is the way we process information influenced by our pre-existing beliefs?
- Why is it difficult for people to change their mind once they have formed an opinion?
- What are the pros and cons of having schemas?

What you will learn in this section

Key learning:

- Behaviourism failed to completely eliminate unobservable cognitive constructs from the study of behaviour.
- A cognitive schema is a generalized mental representation of some aspect of reality.
- "Cognitive schema" is a very broad explanatory principle that includes a range of more specific phenomena, such as stereotypes or mental maps.
- Schemas develop from experience through a process of interaction with the world. There are two processes: accommodation and assimilation. Generally speaking, schemas are resistant to change, and accommodation will only occur when assimilation fails.
- A schema is a construct, in the sense that it cannot be observed directly in behaviour. We infer the existence of a schema from observing the effects that it has on other processes, such as perception or memory.
- Cognitive schemas influence our behaviour and mental processes, but schemas themselves are influenced by experience. Therefore, it is not a unidirectional causality (A→B), but rather a complex interaction that dynamically develops over time.
- Cognitive schemas may be viewed as something that makes us biased, by creating distortions in our perception or memory. However, they also make it possible for us to efficiently process the ocean of information that we encounter.

Key terms: teleological behaviourism, schema, schematic processing, mental representation, stereotype, script, self-schema, accommodation, assimilation, construct, alternative explanation

In a wider context

A cognitive **schema** is a generalized **mental representation** of some aspect of reality. From a behaviourist perspective, a cognitive schema would be something happening in the "black box" because it is a type of mental representation.

Radical behaviourists failed to completely eliminate unobservable cognitive **constructs** from the scientific study of behaviour. This is because it was demonstrated that we cannot fully explain why animals (let alone humans) behave the way they do without considering such cognitive constructs.

Revolution in behaviourism: mental representations

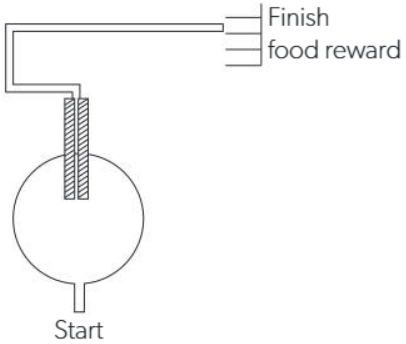
E.C. Tolman (1948) challenged the notion that operant conditioning can explain all observable learning behaviour. By extension, he also challenged the idea that we can dismiss the "black box" as something unworthy of scientific investigation. He did so by demonstrating that even rats learn in ways that cannot be fully explained by conditioning.

One of his studies, published in the article called "Cognitive maps in rats and men" (1948), involved training rats to run through a maze.

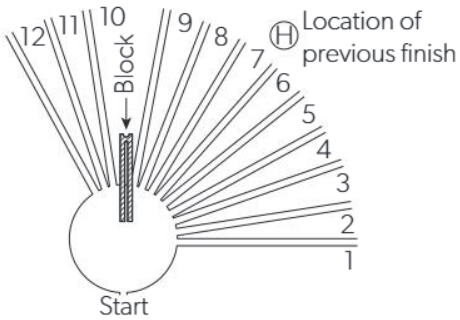
At Stage 1 rats were trained in an angular maze as shown in Figure 2.7.

In this maze, they had to run across a round table and into the entrance of a curved corridor that took several turns and finally led them to a food reward. The rats learned their way through the maze quickly (in 12 trials on average).

At Stage 2 Tolman suddenly changed the maze into a radial pattern like the one shown in Figure 2.8.



▲ Figure 2.7 Stage 1 of Tolman's study



▲ Figure 2.8 Stage 2 of Tolman's study

As you can see, the pathway that the rats had been conditioned to choose was now blocked. Instead, they had 12 alternative paths.

If operant conditioning is an accurate model of real-life learning, then rats should choose the paths closest to the original path, that is, paths 9 and 10. However, results of the experiment showed that the significant majority of rats chose path 6 in the new maze.

This finding contradicts operant conditioning. It suggests that rats were able to form some kind of internal mental representation of the first maze, so they somehow "knew" the direction in which the reward is likely to be found. Of course, a mental representation belongs in the "black box". If we deny the importance of the "black box" (like behaviourists did), we cannot explain why rats behave this way in the study.

Tolman's study was a precursor to the emergence of cognitive psychology. He created a new version of behaviourism that he labelled as "**teleological behaviourism**". Teleology means a study of purposes (which is the opposite to determinism, a study of causes). The term implied the existence of some kind of purposeful behaviour as well as the idea of "purpose" itself—which, of course, resides in the "black box"!

What is a schema?

SAQ Schema theory

A cognitive schema is a generalized mental representation of some aspect of reality. The idea of "mental representation" is best understood through examples.

- Rats in Tolman's (1948) experiment were believed to create a kind of internal map of the maze, which allowed them to acquire a sense of direction and find food more easily on subsequent trials.
- A **stereotype** is a mental representation. When you hear "firefighters", you probably have a generalized mental image forming in your head, and that mental image is your cognitive schema of that group of people.
- When you study a topic in one of your school subjects, you form a mental representation of that topic. Think of astrophysics or reproduction of cells or English grammar—you have your own schemas for all of those things.
- A person with anorexia has a distorted body image: they think they look overweight when in reality they are not. They have a mental representation of their own body that is not a good reflection of the real state of things.

Schema theory suggests that cognitive schemas influence the way we interpret information around the world. Schema creates a "lens" through which we perceive reality.

Schema theory also maintains that schemas develop from experience through a process of interaction with the world. There are two processes: **accommodation** (adjusting the existing schema to better fit the environment) and **assimilation** (adjusting the reality to better fit the existing schema). According to the theory, schemas are resistant to change, and accommodation will only occur when assimilation fails.

A schema is a construct, in the sense that it cannot be observed directly in behaviour. We infer the existence of a schema from observing the effects that it has on other processes, such as perception or memory.

Cognitive schemas may be viewed as something that makes us biased, by creating distortions in our perception and memory. However, they also have an important adaptive function: they make it possible for us to efficiently process the ocean of information that we encounter every day, saving us time and energy.

We will further explore some of these points through examples given in this section.

Different kinds of schemas

Cognitive schema is a very broad concept. It includes a range of specific phenomena. Let us consider some examples.

A **script** is a schema of events. For example, a mental representation of the typical course of actions involved in performing some activity, such as going to a restaurant or making coffee.

Bower, Black, and Turner (1979) conducted a study where they presented participants with a text and asked them to recall it 20 minutes later. The text described a typical sequence of events for which participants are normally expected to have a schema (a script). For example, here is one of the texts:

"*John was feeling bad today so he decided to go and see a family doctor. He checked in with the doctor's receptionist, and then looked through several medical magazines that were on the table by his chair. Finally the nurse came ... The doctor was very nice to him. He eventually prescribed some pills for John. Then John left the doctor's office and headed home*" (Bower, Black, and Turner, 1979, p. 190).

When participants recalled this text later, they filled in the details that were part of the script, but not part of the text. For example, arriving at the doctor's office, looking around at the posters while waiting, the doctor examining him, hearing the diagnosis, picking up the prescription, or making another appointment at reception before leaving. The script that pre-existed in their mind filled in the missing details when they were "remembering" them.

Self-schema is one's schema about oneself. In patients with clinical depression self-schema may be severely impaired. They may develop an unrealistically negative perception of themselves, catastrophizing beliefs about the future, and a sense of worthlessness. Their self-schema is said to be unrealistic because it has no real basis underlying it. For example, they may believe that they are failing at everything they do, but objectively that is not the case. Cognitive behavioural theory of depression developed by Aaron Beck aims to replace negative self-schema in such patients with a more realistic representation of themselves. Cognitive behavioural therapy has demonstrated significant results in the treatment of depression, surpassing the effectiveness of many other approaches to treatment (see Unit 4.4: Psychological treatment).

Development of schemas

Jean Piaget was the first person to seriously introduce the term "schema" into psychology. He believed that children, as they grow up and interact with the world, construct various schemas to help them understand it. According to Piaget, this occurs through two processes—assimilation and accommodation (Piaget and Inhelder, 1969).

Accommodation is adjusting the existing schema to enable a better fit with the environment. Imagine a child who sees for the first time an old rotary dial telephone (have you ever seen one?). The only time they have ever heard the word "phone" before was in relation to smartphones. So that is their schema of a "telephone". If accommodation occurs, the child will adjust their schema after learning this new information—perhaps the new schema may include any device that can perform the function of voice call.

Chat with AI

Use your favourite generative AI to better understand the concept of cognitive schema and the ideas of schema theory. Have a conversation with it using some of the following as possible fragments of your prompts:

- I am learning about schema theory and I learned that examples of cognitive schema include scripts and self-schema. Could you give some other examples?
- I don't quite understand why schema theory is a "theory". Isn't it just the notion of cognitive schema and the idea that a schema acts as a "lens" determining how we perceive things? What makes it possible for us to refer to this idea as a "theory"?
- This is what I have learned about cognitive schema and schema theory in psychology: [provide an outline and some key words]. Could you test my understanding by providing me with a real-life situation and asking a series of questions that require me to analyse this situation using my understanding of cognitive schema? After I answer your questions, please give me feedback and suggest ways in which I could improve my answers and make them more precise/accurate.

Activity

Can you identify examples from your own experience when you observed how a person's schema was resistant to change? This could be something to do with parent–child relationships or friendships. Work in a group of four and share one example each, briefly discussing afterwards how your example is related to the notions of cognitive schema and schematic processing.

ATL Communication, Social

Assimilation is the opposite process of fitting the environment into an existing schema. This may involve distorting information or ignoring some aspects to fit the existing schema without any need to modify it. For example, a person prejudiced against a certain group of people meets a very nice representative of that group and decides that they are a rare exception while still being prejudiced to the same extent against that group as a whole.

It appears that schemas are quite resistant to change. Once a schema is formed, we will try to maintain it by finding ways to fit the world into the existing schema of it. It is only when we fail to do so that we will resort to accommodation.

Usefulness of schemas

Schematic processing is the processing of information based on a schema. It provides an advantage because it saves a lot of time and energy while at the same time providing an accurate enough approximation to reality. For example, if you have a generalized, simplified schema of a tiger as a stripy predator that can eat you for dinner, you immediately know what to do when you see one. You run for your life! You do not stop and analyse how dangerous the creature is. You do not care about keeping an open mind and an unbiased perception.

Of course, a disadvantage of this is that schematic processing will be flawed when a situation changes and when past experience is no longer a good guide in new circumstances.

The pace of cultural evolution far exceeds the pace of biological evolution. Today, being open-minded about things is valued more than being quick in one's judgements. Critical thinking is valued more than basic intelligence. So whether or not schematic processing is a good inherited trait for the modern human is open to debate.

TOK

What concepts and ideas in TOK are related to the idea of cognitive schema/schematic processing in psychology?

Think about confirmation bias and other researcher biases that may affect the work of a scientist, historian, artist, or mathematician.

However, the issue may go beyond the work of an individual researcher and their individual perceptions. An entire paradigm may be affected by schematic processing. A paradigm in science is a way of interpreting data. Observational data in themselves do not make any sense without a theory that is used to interpret them, and a theory is deeply embedded in the currently accepted paradigm.

Research challenges in investigating schema

Schema is a construct. It is not something we can observe directly in behaviour. Neither is it something we can see in a brain scan. We infer the existence of schema from observable behaviour. However, an observation can usually be interpreted in different ways (this is known as **alternative explanations**).

To illustrate this, let us consider the study of office schema by Brewer and Treyens (1981).

In this study, researchers used a room that they designed to look like a typical university office of the time. Examples include: a table with a typewriter, stationery, a Skinner box, some shelves. However, some items that would normally be expected in such a place were deliberately missing—for example, books. Additionally, the room contained some items that were not consistent with the schema of an office—for example, a skull, a picnic basket, and a piece of bark. Participants had to wait in this office briefly (35 seconds) before they were taken to another room where they thought the experiment would take place. When they did go to the other room, the true aim of the study was revealed and they were given a memory task to complete. They had to remember everything they had seen in the office. Some participants simply had to list all the items in writing, others had to draw.

Results of the study showed that objects with a high schema expectancy (such as tables or stationery) were more likely to be remembered than objects with low schema expectancy (such as a skull or a picnic basket). In a number of cases, participant recall included objects that were not actually present in the room (these objects were "inferred" by participants): books, filing cabinet, coffee cups, pens, telephone, window. Brewer and Treyens say that the power of schema becomes most evident when a participant draws a window that was not there.

However, there are at least two alternative explanations for this result:

- Schema influences what we will and what we will not encode in our memory when we look around the office.
- We encode everything irrespective of schema, but schema influences what we can recall at retrieval phase.

To see which of these explanations is more likely, the researchers asked participants to carry out a recall task (writing down all the items they could remember) followed by a recognition task (looking at a long list of objects and marking the ones that they thought they had seen in the office). Brewer and Treyens observed that there were a number of items that participants recognized but did not recall. For example, they would fail to recall seeing a skull when describing the room in writing, but later when asked directly if they had seen a skull, they would answer yes with confidence. This would suggest that the office schema did not prevent participants from encoding the skull, it just prevented them from retrieving it.

This problem of alternative explanations is pertinent in psychology, as in any other subject. Findings of research studies (observational facts) can always be explained in more than one way, therefore potentially supporting not merely one theory but several. However, science is all about uncovering the true nature of things, so we want to be able to investigate alternative explanations and gradually eliminate those that do not fit.

Discussion

Can you think of any other possible alternative explanations for Brewer and Treyens' (1981) results?

ATL Research, Thinking

Exam tip

Exam questions in IB Psychology typically start with the so-called command terms that tell you what exactly you are expected to write about and to what level of depth.

It is important to know the distinction between command terms associated with Assessment objective 2 (AO2) and those associated with Assessment objective 3 (AO3).

AO2 command terms in exam questions will typically be "describe" and "explain". These set the expectation that you will provide some details and/or provide an explanation of a psychological idea, usually supporting it with an example. Paper 1 Sections A and B will only use AO2 command terms.

AO3 command terms will typically be "discuss", "evaluate", and "to what extent". They imply critical analysis and evaluation. Paper 1 Section C will use AO3 command terms.

Conceptual analysis

Causality

One might say that cognitive schemas influence behaviour, and this statement would have plenty of empirical support. For example, Brewer and Treyens (1981) showed that the existing schema of an office affects what we will remember about the objects in an office when we look around it. Schema can affect perception and observable behaviours; for example, the activation of a stereotype can affect the judgements we make about people.

However, a schema itself is influenced by experience. Piaget showed that the formation of schemas is a constant interplay between assimilation and accommodation. Therefore, causality between schema and behaviour cannot be said to be unidirectional: they influence each other (although a schema, once formed, is resistant to change, so the influence may actually become more unidirectional with the course of time).

Measurement

We need to be sure that what we are investigating in a study is indeed a cognitive schema and not something else. A schema itself is a construct, something

hypothetical that cannot be directly observed. We build our knowledge of it based on its effects on other phenomena—the effects that we can observe. However, there are always confounding variables that are hard to control—variables that can obscure the measurement.

Thankfully, cognitive schemas lend themselves to experimental studies. It is relatively easy to operationalize and manipulate a schema, making it possible to conduct true experiments and to control bias in research more easily.

Bias

The concept of schemas is very broad and different researchers have chosen multiple ways to operationalize it. Being invited to an office, being given a context picture before reading a text, being told something about a student prior to meeting them—all of these are thought to activate a "schema". If we use one operationalization to arrive at conclusions that do not replicate in other studies using different operationalizations, we have a problem. For example, imagine that Brewer and Treyens had conducted a similar study, but in a classroom rather than an office, and the results did not replicate—what would this finding suggest?

Another way in which the idea of bias may be applied to cognitive schemas is by looking at the role of schemas in our lives. One may say that it is schemas that make us biased. At the same time, schemas perform a very important function, saving us time and energy for otherwise demanding tasks.

Change

You already know that schemas are a result of an equilibrium between accommodation and assimilation and that, once formed, they may be resistant to change. This is one way of looking at them through the concept of "change".

Another possible angle is to ask ourselves if we can change schemas intentionally. It would be nice, for example, to be able to change stereotypes. Indeed, it is possible to some extent. Schema therapy is a special type of cognitive behavioural therapy directed at modifying a client's maladaptive schemas and replacing them with healthier ones. A person with depression may have a negative self-schema. In a therapeutic context the client may be led to re-examine this schema, assess how realistic it is, and exercise different ways of thinking about themselves and about the world.

Responsibility

The concept of schema is very socially relevant. The word "stereotype", for example, has negative connotations. When research is published, it may be misinterpreted, which may lead to negative social consequences. Researchers have an ethical responsibility to report their results in ways that will prevent unjustified generalizations.

One must be very careful with what conclusions can and cannot be made from a particular study. We must consider the extent to which these conclusions may be influenced by the limitations of the study itself. We must also be mindful about how the results of the study may be perceived by the general public (in case the study comes into the public domain).

2.3 Biological factors in cognitive processes

Inquiry questions

- Can cognitive processes such as memory be reduced to the biological activity of the brain?
- How can we study the role of biological factors in human cognition?
- To what extent can technology help us see what is happening in the brain as we are processing information?

What you will learn in this section

Key learning:

- It is possible that the division between biological, cognitive, and sociocultural variables only exists epistemically (in the way we study the world) but not ontologically (in the world itself).
- Biological reductionism promotes a simplified picture of reality, but it is a necessary step in scientific research.
- One way to understand the role of biological factors in cognitive processes is by using the idea of localization of function in the brain.
- Brain imaging techniques provide a breakthrough in the study of localization of cognitive processes as they allow us to register brain activity when a cognitive task is being performed.
- There is a trade-off between spatial and temporal resolution in brain scanning.
- The cognitive process of flashbulb memories could be caused by selective activation of the amygdala. This is an example of a biological factor underlying a cognitive process.

Key terms: biological, cognitive, and sociocultural perspective, epistemic division, biological reductionism, non-reductionism (holism), localization of function, anterograde amnesia, retrograde amnesia, brain imaging techniques, functional magnetic resonance imaging (fMRI), temporal resolution, spatial resolution, selective activation of the amygdala, flashbulb memory

In a wider context

Every behaviour can be looked at from three **perspectives—biological, cognitive, and sociocultural**. The biological perspective views behaviour as determined by biological factors such as brain structure, genetics, or brain chemistry. The cognitive perspective considers the way we process and interpret information about the world. The sociocultural perspective focuses on how individual behaviour is influenced by the social and cultural group that one belongs to. However, in real life these three influences are closely intertwined and human behaviour is usually determined by multiple factors at the same time.

TOK

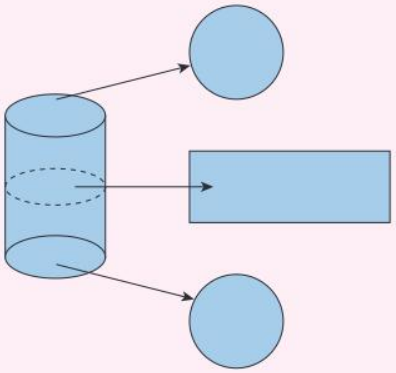
Ontology and epistemology are two large parts of philosophy.

Ontology is the study of "being". Ontological questions are questions of existence; for example: is the universe infinite? What is the nature of reality? What is morality? What is time?

Epistemology literally means "theory of knowledge". Epistemological questions are questions about knowledge; for example: how can we know the limits of the universe? How reliable are our methods of studying reality? What do we know about time? What is our evidence for the existence of morality?

There are a lot of things that exist as a unity, but we look at them from various angles in order to be able to get knowledge about them. As a metaphor, think about a cylinder (a 3D figure) that is projected on two different axes: it appears as a rectangle on one axis and as a circle on the other. Could it be that "cognition" is like this cylinder and the three "perspectives" are merely three ways of looking at the same thing?

Parts of a cylinder



▲ Figure 2.9 Projections of a 3D cylinder onto 2D planes

Biological, cognitive, and sociocultural

Arguably, the biological–cognitive–sociocultural triad is an **epistemic division**. This means that the division only exists epistemically (in the way we study the world), but not ontologically (in the world itself). In reality there is no such thing as a "cognitive", "biological", and "sociocultural" process. Many psychological processes are all these three things at the same time. It is for research purposes that we have constructed these ways of looking at them, emphasising and abstracting one aspect at a time.

When you read a paragraph and try to memorize it, you perform a cognitive activity that is explainable in terms of cognitive constructs (such as information, encoding, retrieval). At the same time, however, there is the brain and the activity of its neural networks: electrical signals go through their pathways. And at the same time you are consuming a product of human culture expressed in language as a cultural tool, using meanings created by this culture to make sense of things. It is one process that can be looked at in three different ways.

Consider a computer. A computer is a combination of hardware and software. Hardware includes all the physical things in it: the hard disk drive, the microprocessors, the circuits, and the electricity that runs through them. Software includes the applications that users interact with, written in a programming language. Programs in the first computers, such as the computational device built by Alan Turing, looked like a physical collection of on/off switches. The pattern of switches determined how the process would flow. Later on, physical switches were replaced with computer programs that consisted of long strings of ones and zeros (1 = go electricity, 0 = stop electricity).

As software became more sophisticated, these extra layers of sophistication did not remove the basic logic of ones and zeros. They just built upon it. Programmers identified the collection of ones and zeros associated with the act of recording data on the hard drive and designed a programming language where the user can type "record" without typing the whole string of digits. However, behind the scenes, the program still translates your command into the pre-defined sequence of switches that control the flow of electricity.

In this metaphor, the human brain is the hardware and human cognitive processes are the software. We describe cognitive processes in terms of "remembering", "thinking", "noticing", but we should not forget that on a very basic level these things are actually "just" complex patterns of brain activity.

Biological reductionism in research

SAQ Biological reductionism

Some scientists reject the idea of cognitive explanations entirely. They believe that we should describe cognition purely in terms of brain activity and should not invent any constructs. They would say, for example, that there is no such thing as "memory", but there are patterns of brain activity that we simply agreed to label with the term "memory". This is an extreme version of **biological reductionism**. In less extreme versions of it, biological

reductionism does not reject the existence of cognitive phenomena. However, it does try to explain all behaviour and cognitive processes in biological terms—that is, to "reduce" complex phenomena to biological factors.

Some people use the word "reductionism" in a way that implies a negative connotation: that it is somehow bad for research to be reductionist and better to be **non-reductionist (holistic)**. This is a misconception.

When you consider multiple interacting variables at the same time, it is harder to carry out a carefully controlled study. As you know, a true experiment is the only method of research that allows us to make cause–effect inferences (A causes B). In order to be able to do so, researchers typically isolate one variable and manipulate it while carefully controlling all other variables. You might say that the true experiment is a "reductionist" scientific method because it attempts to reduce the complex psychological phenomenon to one variable influencing another variable. However, this kind of reductionism is required to obtain solid experimental evidence. If you allowed all confounding variables to vary without controlling them, you would not know how exactly to explain the observed behaviours. Therefore, reductionism is a necessary step in scientific research.

Similarly, research that attempts to reduce complex cognitive phenomena to relatively simple biological factors should not be viewed as limited. Psychologists who try to "reduce" thinking to brain activity are not missing the point. It is their intention to see how much can be explained by biological variables. They know that there is a limit somewhere and their aim is to discover where exactly this limit is.

On the contrary, a drawback with holistic research is that it is rarely experimental. The only way to consider multiple interacting variables at once is to conduct a correlational or a qualitative study. These have their own inherent limitations.

In this and other chapters, there will be many examples of attempts to explain complex phenomena with biological variables.



▲ Figure 2.10 Scientist with a brain

Localization of function and cognitive processes

SAQ Localization of function

One way in which we can understand what role biological factors play in cognitive processes is through the idea of **localization of function**. The word "function" refers to any activity we perform, and that includes cognitive activity. If cognitive functions or their components are localized somewhere in the brain, then an impairment to that brain area must cause that function to change. For example, if short-term memory has a location in the human brain, then a damage to that location will result in a loss of short-term memory (a condition known as **anterograde amnesia**).

A well-known classic case study in psychology is Henry Molaison, also known as patient H.M. He was suffering from severe epileptic seizures. At age 27, in an attempt to control the seizures, he underwent surgery that removed large portions of the hippocampus and amygdala from his brain. The surgery helped with his seizures, but H.M. developed both anterograde and **retrograde amnesia**. Retrograde amnesia affects already existing memories. H.M. retained memories from his childhood, but partially lost his ability to recall what happened in the three years leading up to the surgery. Anterograde amnesia affects one's ability to form new memories. H.M. lost the ability to retain many forms of new information, such as new facts or new faces.

Anterograde amnesia is a curious condition. Imagine bumping into an old friend in a supermarket: he is clearly surprised and happy to see you, he asks how you have been, and you have a short conversation. Then you get distracted with a phone call, step outside to finish it for two minutes, and return to your friend. He sees you again and once again he is surprised and excited, and asks you the same questions, and has no recollection of this same thing happening just minutes before.

Importantly, in the case of H.M. not every kind of memory suffered. His declarative memory was impaired, but his procedural memory was not. Declarative memory is memory of facts, life events, faces, words, experiences—anything that requires conscious processing. Procedural memory is the kind of memory that makes you better at performing tasks after practice when you are learning a new skill. H.M. retained his ability to improve at tasks after practice, even though he could not remember how and when he practised. This is an important finding because it shows that different kinds of memory can be localized in different areas of the brain (Scoville and Milner, 1957).

Research that observes changes in cognitive functioning in individuals whose brain has been damaged is informative but has important limitations. One limitation is that case studies are limited in generalizability. Another limitation is that, as cynical as it sounds, the damage occurs naturally and not under controlled conditions. Usually more than one brain area gets affected, and it is difficult to say which brain area in particular has been responsible for which cognitive changes. However, case studies provide an important insight into the biological basis of cognition because they allow us to study human cognition in conditions that cannot be recreated in a laboratory.

Activity

Based on what you know about typical strengths and limitations of case studies in psychology, evaluate the H.M. study.

ATL Research, Thinking

The use of brain scanning technology

Prior to the 1970s, psychologists had to wait for a patient to die in order to extract their brain and study it. With the invention of brain scanning technology scientists had a breakthrough—they were now able to see inside the brain of a living person.

SAQ Brain imaging techniques

Brain imaging techniques can be divided into structural and functional. Structural brain imaging techniques include magnetic resonance imaging (MRI) and computerized axial tomography (CAT). They show us a static image—the structure of the brain. They are used for detecting the volume of white and grey matter in various parts of the brain, identifying the locations of lesions and tumours. Functional brain imaging includes **functional magnetic resonance imaging** (**fMRI**), positron emission tomography (PET), and electroencephalography (EEG). These techniques allow us to see brain activity in real time—they produce animations rather than static images.

In a typical fMRI study, the subject is required to carry out some tasks in which periods of activity are alternated with periods of rest—when a brain region is active during the performance of a task, the flow of oxygenated blood in that region increases. The response of blood to rapidly changing magnetic fields differs depending on the flow and the level of oxygenation. The signal that is analysed by the fMRI scanner to reconstruct brain activity is known as BOLD (blood-oxygen-level dependent signal). The flow of oxygenated blood directly correlates with the energy used by brain cells, and this directly corresponds to the level of activity in a specific brain region.

However, there are some important limitations to what brain scanning technology can do. Brain scans are limited in **spatial resolution**: they are too crude (basic) to see individual neurons. fMRI scans can currently register "voxels" within the ranges from 1 to 5 mm. Each voxel contains several million neurons and several billion synapses—this marks the limit of how clearly we can see inside the activity of the human brain. Brain scans are also limited in **temporal resolution**, their ability to capture rapid processes. Currently the temporal resolution of fMRI is approximately one second. However, a lot of important brain processes happen in less than one second, and this is just too fast for fMRI scanners to capture.

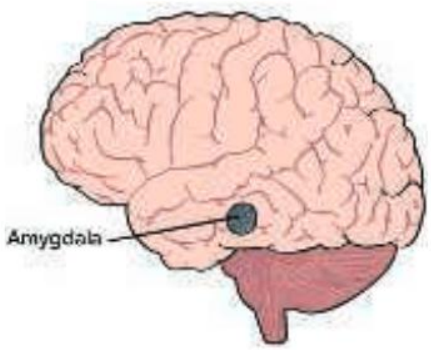
Note: a trade-off also exists between spatial and temporal resolution in brain scanning: typically as you improve one, the other declines. For example, EEG (electroencephalography) has excellent temporal resolution, but is unable to tell us where in the brain the signals are coming from.

There will be examples in this and other chapters of research studies using fMRI and other brain imaging techniques to research biological factors underlying behaviour.

Studies of biological factors in cognitive processes

Let us now look at an example of a study that used brain imaging together with ideas about localization of function in the brain. The aim was to study biological factors in cognitive processes.

Sharot et al. (2007) conducted a study of personal memories of the terrorist attacks in New York City on 11 September 2001, often referred to as the "9/11" attacks. On that day, 19 terrorists hijacked four commercial planes and crashed them into various targets. The first two planes crashed into the Twin Towers of the World Trade Center in New York. Both skyscrapers collapsed after the attack. Three years after the attacks, researchers asked 24 participants to recall events of that day as well as other control events while lying inside an fMRI scanner. Participants were selected so that at the time of the attack half of them had been in Downtown Manhattan (close to the World Trade Center) and the other half had been in Midtown, a few miles away.



▲ Figure 2.11 The amygdala

When inside the scanner, participants could see either the word "September" (the experimental condition) or the word "Summer" (the control condition) appear on a screen, 60 times in a random order. They were instructed to recall the events of 9/11 if they saw the word "September" and any other event that occurred during the preceding summer (June–August 2001) if they saw the word "Summer". This way the scanner could obtain an average picture of brain activity in response to these two types of memories, and the results could be analysed for any significant differences.

Sharot et al.'s results showed that participants in the Downtown Manhattan group exhibited **selective activation of the amygdala** as they recalled events from 9/11 but not while they recalled control events: 83% of participants in this group showed higher activation in the left amygdala during 9/11 trials than summer trials. The same difference was present in the Midtown group, but less pronounced. In fact, selective activation of the left amygdala correlated with the proximity of the participant to the World Trade Centre during the attacks ($r = 0.45$, $p < 0.05$). Taken together, these results suggest that the biological basis for **flashbulb memories** is the participation of the left amygdala. Flashbulb memories are a special type of memory that is very vivid and long-lasting, almost photographic, and believed to be triggered by shocking and unexpected events. The amygdala is known to be involved in emotional responses, including feelings of fear, anger, and anxiety. Perhaps it is a combination of memory networks with these deep emotion-related networks that creates this special kind of memory.

Discussion

What conclusions do you think can be made from this study in terms of causality? Can we say that the left amygdala causes participant memories to be vivid and detailed?

How would you evaluate this study from an ethical point of view? Are there any ethical considerations in conducting the study? What about ethical considerations in disseminating its findings?

ATL Research, Thinking, Communication

Conceptual analysis

Causality and perspective

An investigation of biological factors in cognitive processes is closely related to the concept of causality. This is because some researchers view cognitive processes as being caused solely by biological factors. Whether or not it is true is a very complicated question, which goes back to an old debate about the interaction between the body and the mind (the mind–body problem).

There is also a very important difference between saying that brain activity is involved (or plays a role) in cognitive processing and saying that brain activity is the primary cause of cognitive processes.

Bias

Research studies investigating the potential causal role of biological factors in cognition are important. They should not be criticized for being "reductionist" because this kind of reductionism is a necessary step to understand the exact place of biological factors in cognition.

Arguably, taking a reductionist stance for research purposes is not a bias. However, when it is outside of research, hypothesis-testing context, providing an explanation that reduces complex behaviour to basic biological factors may be considered a biased explanation. For example, imagine we explained student achievement at school purely with biological predisposition. The implications of this would be early testing of their IQ, placement in separate classes, labelling, and stigmatization.

Measurement

As you have seen from the research examples, problems of measurement impose some crucial limitations on the nature of conclusions that we can make. For example, brain scanning technology has been a real game changer: it has allowed us to see inside the brain of a living person (as opposed to waiting for that person to die). However, brain scanning technology has limitations—it is limited by its spatial and its temporal resolution.

Another huge problem related to measurement is isolating factors. When we conduct a study, we want to say that A (a biological factor) influences B (a cognitive factor). But for this to be possible, A has to be isolated. As we have seen, the activity of our nervous system is so complex that such isolation is difficult to achieve.

Change

One of the attractive features about this line of research is that biological factors are easy to control. This gives us a tool to change behaviour. For example, we have some idea about which brain areas are involved in which functions, so if something goes wrong, we can put together a course of therapy that will utilize our brain's ability to restore itself. For example, speech therapy for those who suffer from aphasia is aimed at training other brain regions to take over the function of the damaged areas.

Responsibility

A study of biological factors of cognition involves manipulating biological factors. That, by definition, is an invasive procedure. Animal experimentation is widely used since it would be unethical to perform such procedures on human beings, but psychologists must abide by the Ethical Code regulating the use of animals. See the section "Ethical considerations in animal research" on page 62 for details on the most important ethical regulations.

There is always a clash between ethical regulations and generalizability of research. Animal research is limited in that we never know if it is generalizable on human beings. Take HIV treatment as an example—over 85 vaccines for HIV have worked well in primates, but when subsequently trialled in humans, all of these vaccines failed (Bailey, 2008). Human research is often impossible for ethical reasons. You can cut out a piece of a rat's brain (with permission from an ethics committee, if it is justifiable for your research), but you cannot do the same to a human.

2.4 Cognitive models

Inquiry questions

- How can an unobservable cognitive process be scientifically studied?
- Which memory model is the best?
- Why is there never a perfect fit between a model and available data?

What you will learn in this section

Key learning:

- Without cognitive processes, the science of human behaviour is incomplete.
- To study cognitive processes scientifically, we create models of them. Models are hypothetical, but predictions based on models can be tested in research.
- There always exist alternative models that attempt to explain the same phenomenon. We select between models based on several criteria, such as goodness-of-fit to the available research findings and parsimony (the ability to explain a lot with only a few components).
- According to the multi-store memory model (MSMM), memory consists of three stores, each of which is characterized by a specific duration and capacity. In order for information to move to the next memory store, certain conditions (attention and rehearsal) have to be met.
- It is impossible for one research study to support a cognitive model in its entirety. Each study tests one specific prediction based on the model.
- Some research findings do not sit well with the MSMM. For example, the phonological similarity effect and the word length effect.
- The working memory model (WMM) suggests a further sub-division of short-term memory. The increased complexity allows the model to explain observations that MSMM could not explain.
- Increased explanatory power of the WMM decreases parsimony.

Key terms: scientific model, model of a cognitive process, multi-store memory model (MSMM), working memory model (WMM), goodness-of-fit, parsimony, sensory memory, short-term memory, long-term memory, rehearsal, serial position effect, primacy effect, recency effect, phonological similarity effect, word length effect, phonological loop, phonological store, visuospatial sketchpad, cognitive load theory

In a wider context

We have considered biological foundations of cognitive processes and now we are switching to the purely cognitive perspective: trying to understand cognition in itself without reducing it to biology.

Examples of cognitive processes include memory, perception, attention, thinking, decision-making, problem solving, and language. All these processes are unobservable, which led behaviourists to claim that studying them is unscientific.

However, we cannot have a complete understanding of human behaviour without them, so cognitive processes do need to be investigated. To study them scientifically, we create models of them. In this unit, we will focus on two examples of models of memory.

Historical context

Behaviourists disliked cognitive factors because they are unobservable—a "black box". To them unobservable was synonymous with speculative and unscientific.

However, with the course of time it became obvious that many forms of behaviour (both human and animal) cannot be fully explained without using at least some cognitive variables. In a clever series of experiments with rats Tolman (1948) demonstrated that rats are capable of forming "cognitive maps"—mental representations of their environment (see Unit 2.2: Schema theory). Bandura (1977) highlighted the importance of observational learning—learning that occurs without any overt behaviour and even without direct punishment or reinforcement (see Unit 2.6: Social learning theory).

This and other research played a role in the "cognitive revolution" in psychology that started in the 1950s. Cognitive psychology reinstated the role of cognitive variables and even assigned primary importance to them. After all, highly developed forms of cognition are what makes us human, and we should not avoid a scientific study of these processes only because they cannot be directly observed.

Scientific models

So now we have a problem. We want psychological knowledge to be scientific, but we also want to study something that cannot be directly observed or measured. What do we do?

The answer lies in the idea of **scientific models**. A model is a hypothetical construct that provides a description of a directly unobservable phenomenon. For example, the multi-store memory model (MSMM) suggests that memory consists of three stores (sensory, short-term, long-term). However, these three "stores" do not exist in the physical sense: if we open a person's skull, we will not find three boxes in it. Brain scans do not show the boxes either. How do we know there are three stores, not one, not five?

We don't know! A model is hypothetical.

However, it is possible to formulate various predictions based on the model and test the accuracy of those predictions. If the observations match our predictions, we grow more confident that the model is valid. If the observations do not match the predictions, we should reject the model or revise it.

TOK

Here are some questions to consider that relate to the role of models in other subject areas, as well as knowledge in general:

- To what extent are models useful as a tool for acquiring knowledge?
- Name examples of models that are used in other disciplines—for example, geography, natural sciences, mathematics.
- What are the common limitations of models? Think about simplification.
- If models simplify reality, why use them in the production of knowledge?

Models of memory

Multiple alternative models have been proposed for each cognitive process. For example, there are multiple models of memory.

This is a normal process of scientific development: models compete with each other over how well they will fit the currently available research findings. Simply speaking, the model that provides the best fit to the available data will win and will remain a winner until another model gets proposed that provides an even better fit.

The reality is not quite as simple as that, however, because **goodness-of-fit** is not the only criterion that is used to select a model. Another important quality of a scientific model is its **parsimony**. A model is said to be parsimonious if it explains a lot of phenomena with a small number of components. To understand this, consider two extremes:

1. A scientific model that attempts to explain all available observations, but in doing so it has to introduce a whole range of additional variables, parameters, exceptions, and so on. The model has a large explanatory power, but it is unwieldy, difficult to apply in practice, and difficult to test in experimental research.
2. A scientific model that is simple and compact, easy to test in experiments—but it does not explain some of the observed characteristics of human behaviour.

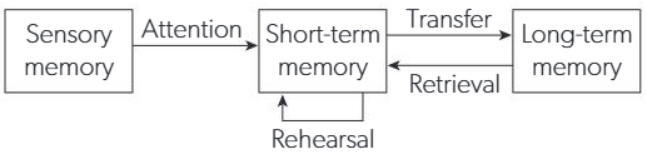
The desirable scientific model is somewhere in between: the one that provides a large explanatory power and yet does not introduce many unnecessary variables into the equation.

Multi-store memory model

The **multi-store memory model** (**MSMM**) was proposed by Atkinson and Shiffrin in 1968. In this model, human memory is said to consist of three separate components:

1. sensory memory
2. short-term memory store
3. long-term memory store.

Each of these three components is characterized by a specific duration (for how long the store is able to hold information) and capacity (how many units of information it can hold). For information to move to the next memory store, certain conditions have to be met.



▲ Figure 2.12 Multi-store memory model

The capacity of **sensory memory** is only limited by our perception. For example, for visual information, we can keep everything that enters our visual field. However, the duration of sensory memory is short. Traces in visual sensory memory decay after one second.

The condition that has to be met for information to transfer from sensory memory to short-term memory is attention. If a unit of information is attended to, it does not decay, but moves further. Therefore, potentially sensory memory has unlimited capacity, but it is only transient, and we can only attend to a small subset of information stored in sensory memory.

The capacity of the **short-term memory** (STM) store has been established to be 7±2 chunks of information. A chunk is a meaningful combination of individual units. For example, the sequence of symbols PCBMXBMWXBOXPS4 makes 15 units of information, which falls outside the capacity of short-term memory. If, however, you perform some grouping: PC–BMX–BMW–XBOX–PS4, it now becomes five chunks!

The duration of short-term memory is generally no longer than 30 seconds. If the information is left unattended, the trace fades away in this period of time.

The condition for increasing the duration of short-term memory and transferring information into the third store (long-term memory) is **rehearsal**. If we rehearse information (e.g., repeat words over and over again), it stays in the STM longer, and eventually the trace gets consolidated and the information enters the long-term memory store.

Long-term memory (LTM) is described as a place for storing large amounts of information for indefinite periods of time (Galotti, 2008).

Activity

Do some research and suggest your explanation for phenomenal memory (extraordinary memory powers). How is phenomenal memory possible? Can it be learned? You may start exploring the issue with the story of Laurence Kim Peek (1951–2009), the person who became the inspiration for the 1988 movie *Rain Man*. He read countless books and instantly memorized them. Reportedly he could accurately recall the contents of at least 12,000 books. He could also carry out complex calculations in his mind without the need for a calculator.

ATL Research, Communication, Self-management

Although the capacity of LTM is potentially unlimited, not all information that is stored in LTM is easily retrievable. It is not storing but retrieving information from memory that may be problematic.

Experimental test of the MSMM

Atkinson and Shiffrin's model has a lot of components that require testing. Here are just a few questions that scientists need to investigate as part of their efforts to test the validity of this model:

1. Are the memory stores really distinct and separate? For example, is sensory memory really separate from STM?
2. Are there really three memory stores, not more, not less?
3. Is there a physiological basis for the memory stores or are they just constructs?

TOK

The fact that we can only attend to a limited amount of information in our sensory memory at a given time links to the knowledge concept of "selectivity of perception". Recall examples from various areas of knowledge that show how selectivity of perception affects our knowledge in a discipline. How does it work in history, for example?

SAQ Cognitive models

Cognitive models are one of the content points that can appear in questions in Papers 1A and 1B. You will need to be able to explain what a cognitive model is and support your explanation with one example.

In this unit we are considering two prominent models of memory: MSMM and WMM. You could choose either of them and build your response around it.

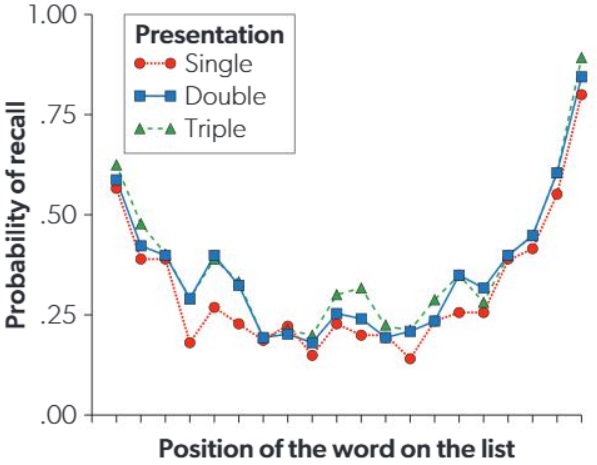
4. Is rehearsal necessary and sufficient for the transfer of information from STM to LTM? Can this transfer occur without rehearsal? Can the transfer fail to occur in the presence of rehearsal?
5. Does information really only flow in one direction (from sensory memory to LTM)? Can information flow backwards; for example, can LTM influence which pieces of data are selected from sensory memory and transferred into STM?

All these questions and predictions require a series of carefully controlled experiments. Let us consider one such example.

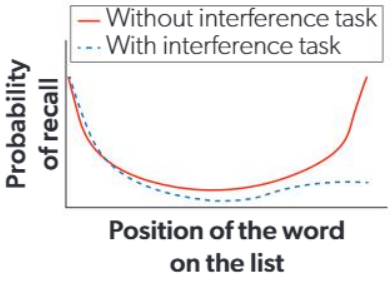
Glanzer and Cunitz (1966) are famous for their research on **serial position effect**, which serves as support for STM and LTM being separate memory stores. Serial position effect is the tendency to recall the first and the last items on a list better than items in the middle. Participants (240 army-enlisted men) were required to memorize lists of words followed by a free-recall task. A free-recall task is when you are permitted to recall the words in any order.

There were two conditions.

1. In the first condition, participants were presented with recordings of 20-word lists. Immediately after hearing the list of words they were required to do a free-recall. Results of these trials clearly demonstrated serial position effect in both its aspects: participants were better at remembering words at the start of the list (**primacy effect**) and at the end of the list (**recency effect**). This did not depend on the number of repetitions of each word.
2. In the second condition, researchers introduced a delay between the end of the list and the start of recall. During the delay, participants engaged in a filler task: counting backwards from a given number for 30 seconds. The filler task was meant to prevent rehearsal. The resulting data indicated that participants were still successful at recalling the words from the start of the list (primacy effect preserved) but were no longer able to recall the words from the end of the list (recency effect disappeared).



▲ Figure 2.13 Serial position effect



▲ Figure 2.14 Disappearance of recency effect after a filler task

Glanzer and Cunitz explained that when people are hearing a list of words with the intention to memorize them, they tend to repeat the words to themselves. The first words on the list get repeated (rehearsed) more often and enter the long-term memory, which is unaffected by the delay and the filler task. However, the last words on the list are not rehearsed enough. Without rehearsal, their trace in short-term memory decays in just 30 seconds, so the recency effect disappears after the filler task. Since one of the effects disappears and the other does not, it supports the idea that STM and LTM have separate memory mechanisms behind them.

Data that do not fit into the MSMM

Let us consider two examples of empirical findings about human memory that do not fit into the multi-store memory model.

The first example is the **phonological similarity effect**. This effect was discovered in a study by Conrad and Hull (1964). In their study, participants were required to recall lists of letters, some of which were phonologically similar (e.g., B, D, C, G, P) while others were not (e.g., F, H, P, R, X). The results indicated that rhyming lists were more difficult to remember than non-rhyming lists. The multi-store memory model does not contain anything that would account for this effect. Things get even more perplexing because in Conrad and Hull's experiment letters were presented to participants visually: participants did not *hear* the letters, they *read* them. Letters on the rhyming lists (e.g., B, D, C or P) sound similar, but they certainly do not *look* similar. Their visual traces should be sufficiently distinct, so why would it be more difficult for someone to recall them?

The second example is the **word length effect**. This was discovered by Baddeley, Thompson, and Buchanan (1975). The experiment showed that the capacity of short-term memory is greater for short words than for long ones. All other things being equal, people tend to recall a larger number of words from lists of words with fewer syllables (e.g., house, candle, table) than from lists of words with more syllables (e.g., television, candlestick, apartment). Again, there is nothing in the multi-store memory model to account for such a finding. The average capacity of short-term memory is seven chunks of information. A chunk is a chunk, no matter how many syllables are in it. Common sense suggests that words with more syllables probably somehow took up more "space", but what kind of space is that? There is no such concept in MSMM.

Activity

Get into a small group and suggest a minimal modification to the MSMM that would explain both the phonological similarity effect and the word length effect. Share your suggestions with other groups.

ATL Thinking, Communication, Social

Working memory model

The **working memory model** (**WMM**) proposed by Baddeley and Hitch (1974) focuses on the structure of STM. In this model, working memory consists of a central executive that coordinates two subsystems: the **visuospatial sketchpad** and the **phonological loop**.

The visuospatial sketchpad ("the inner eye") holds visual and spatial information. The phonological loop holds auditory information and is further subdivided into the **phonological store** ("the inner ear") and the articulatory rehearsal component ("the inner voice").

- The inner ear holds sound in a passive manner; for example, it holds someone's speech as we hear it.

Exam tip

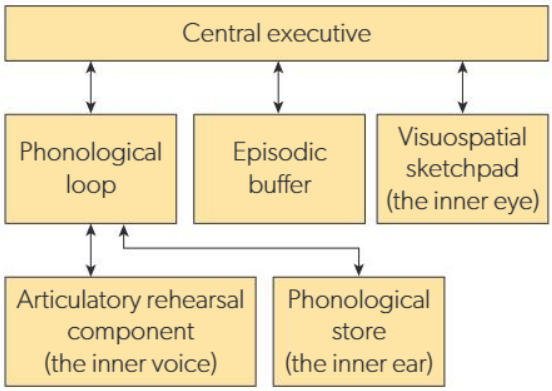
Generally speaking, IB examiners equally accept the breadth approach, where the student considers more examples in less detail, and the depth approach, where the student considers fewer examples, but with a more elaborate analysis.

In short-answer questions (such as those in Papers 1A and 1B), you will typically be asked to describe or explain one psychological idea using one example. In extended response questions (such as those in Papers 1C), the typical quantifier will be "one or more", leaving it up to you to decide whether the breadth approach or the depth approach should be used.

Having said this, make sure that your response keeps developing the argument. Avoid the situation where you use two different examples to support the same point. In models of memory using both MSMM and WMM in an extended response could be effective because these two models in combination illustrate the interaction between the model's complexity and explanatory power. Using both models together may allow you to bring up more arguments and achieve more depth than if you were only using one.

- The inner voice, on the other hand, turns visual stimuli into sounds. For example, if we are shown a list of written words, we may subvocally pronounce these words, changing the modality from visual to auditory, and the words will enter our STM through the auditory channel.

The central executive is a system that allocates resources between the visuospatial sketchpad and the phonological loop. In this sense, it is the "manager" for the other two systems. In 2000, Baddeley and Hitch also added the fourth component, the episodic buffer, as a component that integrates information from the other components and also links this information to long-term memory structures.



▲ Figure 2.15 The working memory model

Explaining the data that did not fit

How does this model explain the evidence that did not fit into the multi-store memory model (MSMM)?

The phonological similarity effect is explained by suggesting that, when a participant reads a word that is presented on paper, the inner voice subvocally pronounces it and sends it to the inner ear. Therefore the word is stored as a sound (in the phonological loop) and not as an image (in the visuospatial sketchpad). In the phonological loop the words are stored acoustically, so memory traces of similar words may be similar, which means that rhyming words and letters (e.g., B, D, C, or P) will be more easily confused even though they do not look similar.

The word length effect is explained by a similar line of reasoning. When a study participant reads a word, it is subvocally pronounced in the inner voice and enters the inner ear. Long words have longer acoustic traces, and therefore take up more capacity in the phonological loop.

Both of these explanations are supported by experiments using the experimental technique of articulatory suppression. This is when participants are asked to repeat a sequence of sounds (e.g., the–the–the or one–two–three–one–two–three) over and over again while at the same time performing the experimental task. In doing this, the capacity of the "inner voice" is filled up and so words and letters get stored as visual images rather than sounds. Since the "inner voice" is busy, they get sent to the "inner eye" rather than the "inner ear". For example, in an experiment like this Baddeley, Lewis, and Vallar (1984) demonstrated that the phonological similarity effect disappears when the articulatory suppression technique is used. This means that, when participants read rhyming letters

(e.g., B, D, C, or P), but at the same time were required to pronounce a repeated sequence of sounds (e.g., the–the–the…), the recall of rhyming letters was the same as the recall of non-rhyming letters.

Evaluation of the working memory model

Overall, the strength of the working memory model is that it is more sophisticated than the multi-store memory model and allows us to explain a wider range of observed phenomena.

However, models of this degree of complexity are harder to test empirically. They are less parsimonious. Complicated experiments must be designed to test just one specific aspect of the model (e.g., articulatory suppression). Maybe as a consequence of this, and due to the existence of multiple potential explanations of the same experimental result, the exact role of some of the components of the model (the central executive and especially the episodic buffer) remains unclear.

SAQ Cognitive load theory

Cognitive load theory is closely related to both schema theory and the multi-store memory model (MSMM).

It was developed by John Sweller (1988). It provides guidelines for those who prepare instructional materials (such as teachers and textbook writers) to decrease unnecessary cognitive load during learning and refocus the learner's attention on the important schema-related material. Cognitive load refers to the amount of information that our working memory can process at any given time. The theory makes a distinction between three types of cognitive load:

1. Intrinsic cognitive load—this is the inherent level of difficulty associated with a particular topic. For example, learning about memory in psychology may be inherently easier than learning about epigenetics, and the teacher has little control over that. The only thing that a teacher can do with an inherently difficult schema is break it down into smaller schemas, teach them separately, and combine them at a later time.
2. Extraneous cognitive load—this is the cognitive load created by the manner in which material is presented to students. Teachers have a great degree of control over this. Extraneous load is often referred to as unnecessary. This is because it consumes the limited cognitive resource and does not allow the learner to sufficiently engage with the useful information. For example, it is much easier to draw a schematic representation of the structure of the atom than to explain it verbally. If you explain it verbally, most of your students' cognitive resources will be consumed by understanding what you said and translating it into a mental image.
3. Germane cognitive load—this is the cognitive load related to the construction of schemas. In other words, it is the kind of mental processing that leads learners towards constructing an effective mental representation of the studied material, understanding this material, and connecting it to what they already know. Germane cognitive load is considered to be useful. It was Sweller's belief that to teach more effectively, we must minimize extraneous cognitive load by either eliminating it or converting it into germane load. An example of germane load is creating diagrams and flowcharts to represent complex ideas.

Chat with AI

Both MSMM and WMM are classic memory models in psychology, and you can expect generative AI to have a good understanding of them. Use this to "interrogate" your favourite generative AI to gain a deeper understanding of both models. Here are four prompt starters that you could consider using:

1. I am an introductory psychology student studying the multi-store memory model (MSMM) and the working memory model (WMM). Could you give several examples of phenomena or observations that are explained well by one of these models, but not the other?
2. Sometimes I can't remember someone's name when I need it, but then after some time it just suddenly comes to me. How can this be explained? Can MSMM and/or WMM provide an explanation for this phenomenon?
3. Could you compare WMM and MSMM in terms of their strengths and limitations? Please use bullet points and support your answer with examples.
4. I don't fully understand the difference between the phonological loop and the phonological store in WMM—could you explain with a simple example?

Long-term memory uses schemas to structure and categorize information. As you remember, the capacity of short-term (working) memory is 7±2 chunks. However, once you understand some material and create a meaningful schema of it, it becomes one chunk. For example, when you are learning about the multi-store memory model for the first time, you need to understand the difference between memory stores, how information flows, how rehearsal is related to capacity, what it means for capacity of long-term memory to be unlimited, and so on. Once you have constructed your schema of MSMM, it is now one chunk. Like in a game of Tetris®, when you fit the pieces together, they collapse and free up space for more pieces. This is why the creation of meaningful schemas is the end goal of effective learning.

Cognitive overload occurs when the combination of the three types of cognitive load exceeds one's capacity and becomes overwhelming. When this happens, even the most knowledgeable individuals can fail to process the simplest tasks.

Conceptual analysis

Perspective

Different models or theories of the same psychological phenomenon may be viewed as different perspectives of that phenomenon. It is always the case that more than one **model of a cognitive process** is possible. Models "compete" with each other. With new evidence gathered, some models provide better fit to the observed results of research, some worse.

The fit between any model (however good it is) and available research evidence is never perfect: there are always phenomena that one model can explain but another one cannot. For example, the phonological similarity effect can be easily explained by WMM, but MSMM fails at this task. Conversely, some phenomena like the sensory memory can be explained by MSMM, but not WMM.

The choice of the "best" cognitive model is guided by several factors and explanatory power is only one of them. One other important consideration is the model's parsimony.

Causality

As a rule, models of cognitive processes are causal. In MSMM, for example, information enters the short-term memory store *because* it was attended to in the 30-second window while it was available in sensory memory. Information enters the long-term memory store *because* it was rehearsed a sufficient number of times. This property allows predictions to be made based on models (because A influences B, if we create such and such conditions for A, we should observe such and such changes in B … let's set up a study and check!). This also makes the experiment the most suitable research method for investigating (testing) cognitive models.

Measurement and bias

Properly organized and carefully controlled experiments are fundamentally important in the study of cognitive models. We must be certain that our experimental evidence reflects the reality of things and not artefacts and biases associated with a poorly organized research procedure. In addition, we must try and obtain evidence that is as unambiguous as possible—that is, evidence that can only be interpreted in one way and not multiple ways.

Experiments should be replicated, and evidence triangulated. Conclusions cannot be based on limited evidence. There should always be a research programme accumulating experimental results. In cognitive studies variables are more subtle, less visible, and harder to control. For example, one of the things that must be controlled is tiny nuances in how participants in various groups perceive the experimental instruction, or how the instruction is delivered to them. Considerations of internal validity are very important here. But so is the trade-off between internal validity and external validity: the more carefully controlled our experiment is, the more artificial it is and the less confident we should be that results actually apply to how participants' cognition works in real-life situations.

Change

It is assumed that the multi-store memory model will equally apply across the world, in China as well as Venezuela, for example. At the same time, the overwhelming majority of research samples in investigating models of cognitive processes come from WEIRD (Western, Educated, Industrialized, Rich, and Democratic) contexts. There is an obvious lack of cross-cultural studies in this area.

Cross-cultural variation in the applicability of a cognitive model, if discovered, would be ground-breaking. At the very least it would mean that the three memory stores are not biologically based (e.g., different brain areas) but culturally determined. Or it may mean that the model is simply incorrect. See Unit 2.7: Cultural factors in cognitive processes, for further details on how cross-cultural research contributes to our understanding of cognitive processes.

Responsibility

In conducting experiments to investigate the validity of cognitive models, deception plays a large role. If participants know the true aim of the study, or even if they suspect what it could be, such expectations may influence their cognitive performance in all sorts of ways that may interfere with the experiment and lead to biased results.

It is important to handle deception ethically: gain approval of an ethics committee, not deceive beyond the necessary minimum, reveal the deception as soon as possible during debriefing, and give all participants a chance to withdraw their results once they know they have been deceived.

2.5 Cognitive biases and the dual processing model of thinking and decision-making

Inquiry questions

- Can human biases in thinking and decision-making be predicted?
- Do humans think rationally?
- What is meant by the duality of human cognition?

What you will learn in this section

Key learning:

- The duality of human cognition has been recognized in psychology for a long time. We rely on automatic processing most of the time, but we are also capable of rational, logical, controlled analysis.
- Kahneman's dual-process model of thinking postulates the existence of two systems of thinking (System I and System II) that act sequentially. System I thinking relies on heuristics and may result in cognitive biases. System II thinking can override the "default" decisions of System I. System II thinking is described by normative models and System I by descriptive models. A cognitive bias occurs when human thinking systematically deviates from the normative model.
- Framing effect is the tendency to make more risky decisions when outcomes are described in terms of loss and to avoid taking risks when outcomes are described in terms of gains. The normative model that is violated by framing effect is the expected-utility model of decision-making.
- Confirmation bias is the tendency to seek out information that confirms our pre-existing beliefs and avoid information that can potentially contradict them. It violates the normative model of scientific thinking about evidence where both supporting and contradictory evidence is important to test a theory (the logic of falsifiability).
- Anchoring bias occurs when an individual's decision is influenced by an initial piece of information, even when this information is not very relevant. It violates logical thinking where the order in which information is presented and/or irrelevant information should not affect the result of the calculation.
- Debiasing is a group of techniques used to reduce the influence of cognitive biases. It has been demonstrated to be relatively effective.

Key terms: thinking, decision-making, automatic processing, controlled processing, duality of human cognition, dual-process model, System I and System II thinking, heuristic, cognitive bias, descriptive model, normative model, framing effect, expected-utility theory, anchoring bias, confirmation bias, falsifiability, debiasing, motivational, cognitive, and technological debiasing strategies

In a wider context

Thinking is a cognitive process. Other cognitive processes include perception, memory, language, attention, and emotion. Thinking is the only cognitive process that can take existing information, process it, and produce new information. Thinking is very closely related to **decision-making**, which is the observable act of choosing among available alternatives. Sometimes these two are used together to underline their interdependent nature, as in the "dual-process model of thinking and decision making".

Just like there are models of memory, there are also models of thinking and decision-making. The dual-process model is one of them.

Exam tip

The dual processing model and cognitive biases are two separate content points in the IB syllabus. Exam questions in Paper 1 Section C can be asked separately about the dual processing model or cognitive biases.

However, there is so much overlap between these that it makes sense to study them together. Just make sure that you slightly adjust the focus of your response depending on the question.

History: duality of human cognition

Starting with the works of William James in the 19th century (James, 1890), psychologists brought up the idea that human mental processing may be of two types, automatic and controlled. Other names for **automatic processing** included unconscious, implicit, instinctive, intuitive. Other names for **controlled processing** included conscious, explicit, logical.

You encounter multiple examples of automatic processing every day: when people pick a more expensive item of clothing because it is believed to be better in quality, when people elect leaders based on their public appearances, when people go and see a movie because everybody else does. We frequently make decisions without much logical analysis or mental elaboration.

It is this **duality of human cognition**—relying on automatic processing most of the time, but also being able to perform a deep meaningful analysis—that has attracted the attention of psychologists. It has resulted in multiple specific theories that all share this basic premise but differ in exact details and focus. One such theory is the **dual process model** of thinking by Daniel Kahneman (2011).



▲ Figure 2.16 Daniel Kahneman

The dual-process model of thinking

SAQ Dual-process theory

Kahneman's dual-process model asserts that people use two distinctly different systems of thinking: **System I and System II**.

- System I is fast, intuitive, and automatic.
- System II is slow, controlled, and rational.

System I is for making snap judgements based on past experiences without a deep analysis of the current situation. System II is for carrying out a deep rational analysis and making balanced decisions.

An important point of the theory is that System I and System II act sequentially. This means that our first reaction always comes from System I. Later on this first intuitive judgement

System I	System II
Fast	Slow
Unconscious	Conscious
Automatic	Effortful
Everyday decisions	Complex decisions
Error prone	Reliable

▲ Figure 2.17 System I and System II

may or may not be overridden by System II analysis. There is no such thing as analysing a situation logically (System II) without first having some sort of gut feeling about it.

Psychological models of System I thinking are called **descriptive models**. Descriptive means describing the way it is. Models of System II are called **normative models**, which refers to setting the norm and saying what it should be.

One example of a normative model of thinking is formal logic. Another example is statistical probability theory. Normative models are not part of psychology. They describe how people *should* think, not how they *actually* think.

Whereas System II uses rules and calculations, System I uses **heuristics**. Heuristics are simplified rules of reasoning, also referred to as cognitive shortcuts. For example, an amateur chess player may use the following simplified rule: if you are not sure what to do, move your pawns forward. The "logical" way of playing (calculating all possible moves of the opponent) is too computationally difficult. Such heuristics are not guaranteed to lead to a positive outcome. This is especially so if the situation is different from what you encountered before, but they are a good enough general rule in most situations. They allow us to save mental energy.

While a heuristic is any simplified rule of reasoning (cognitive shortcut), some heuristics may lead to systematically inaccurate or irrational decisions. These are known as **cognitive biases**.

Cognitive bias

Cognitive biases have become the main focus of research in the dual process theory. A very large number of cognitive biases have been discovered. In this unit we will look at three such examples.

Example 1: framing effect

Framing effect is the tendency to avoid taking risks unless to avoid losses.

A normative theory that describes the correct way of thinking about risks and probabilities is the **expected-utility theory**. It explains that you need to multiply the expected pay-off (the utility) of an outcome by the probability of that outcome, and then choose the outcome that yields the highest number. For example, suppose you were choosing between the following two gambles:

1. If you choose option A, you get $10 for certain.
2. If you choose option B, you get $200 with a 6% probability.

Using the expected-utility theory, it is correct (more rational) in this situation to take a risk. This is because the expected utility of option A is $10 × 1 = $10, but the expected utility of option B is $200 × 0.06 = $12.

However, research has demonstrated that the expected-utility theory is not a good theory to describe (and predict) the choices that people actually make in this and similar situations.

In one of their most famous experiments, Tversky and Kahneman (1981) gave their subjects a problem like this:

"Imagine that the USA is preparing for an outbreak of an [unusual] disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the program are as follows:"

(The options were different for two independent groups of subjects.)

Group 1	Group 2
Program A: 200 people will be saved	*Program A: 400 people will die*
Program B: there is $\frac{1}{3}$ probability that 600 people will be saved, and $\frac{2}{3}$ probability that no people will be saved	*Program B: there is $\frac{1}{3}$ probability that nobody will die, and $\frac{2}{3}$ probability that 600 people will die*

▲ Table 2.1 Response options for the two groups in Tversky and Kahneman (1981)

If you think about it logically, both of the choice sets are identical. "200 people will be saved" is the same outcome as "400 people will die". Moreover, Program A and Program B are also the same from the perspective of expected utility. One-third probability that 600 people will be saved has the same expected utility ($\frac{1}{3} \times 600 = 200$) as saving 200 people with certainty. Logically, there should be no difference in how participants in the two groups respond to this situation and we should expect a 50:50 split between Program A and Program B, indicating no preference.

In Tversky and Kahneman's (1981) experiment, participant choices in the two groups were reversed:

	Group 1	Group 2
Program A	72%	22%
Program B	28%	78%

▲ Table 2.2 Findings from Tversky and Kahneman (1981)

This deviates from the prediction of expected-utility theory, so apparently participants used some kind of heuristic. The only difference between groups is in how the situation was described, either in terms of potential gains ("will be saved") or in terms of potential losses ("will die").

Tversky and Kahneman (1981) explain this finding in terms of a shift in the reference point. In Group 1, the reference point is the future state (600 people will be dead). The options are perceived as potential gains (how many people will I save?). In Group 2, the reference is shifted to the present (nobody has died yet). The options are perceived as potential losses (how many people will we lose?). Depending on whether outcomes are described ("framed") as gains or losses, subjects make different judgements: they are more willing to take risks to avoid losses, but they tend to avoid risks associated with gains.

Example 2: confirmation bias

SAQ Confirmation bias

Confirmation is a tendency to seek out information that confirms pre-existing beliefs and avoids information that contradicts them. People may be susceptible to **confirmation bias** because rejecting one's established beliefs is psychologically uncomfortable. It is never a pleasant feeling to realize that we have been wrong.

TOK

In the methodology of sciences it is difficult to overestimate the influence of Karl Popper's ideas on falsification and **falsifiability**. Put simply, the idea is that no amount of supporting evidence is sufficient for us to be certain that a scientific theory is true. On the other hand, even a small amount of contradictory evidence is enough for us to accept that the theory is false. For example, if your theory is that "all swans are white", then no matter how many white swans you observe, you cannot be 100% sure of the truth of the theory, but observing even one black swan is enough to reject the theory.



▲ Figure 2.19 Black swan

If that is the case, then looking for supporting evidence is meaningless. Looking for potentially contradictory evidence, on the other hand, is informative. The more we try to refute the theory and the more we fail, the more confident we are that we can accept this theory—provisionally—as true.

By extension, any statement that is not falsifiable (i.e., it is impossible to design a test that could potentially refute it) is not scientific.

Can you think of any examples of non-falsifiable theoretical claims or statements in psychology?

The normative model that confirmation bias violates is the scientific principle that all evidence should be considered equally, and we should look for objective tests of our beliefs rather than simply a confirmation. It is always possible to obtain confirming evidence, even for theories that are wrong. In science, we look for balanced conclusions obtained from objective evidence (both supporting and contradictory). However, this scientific thinking does not always translate into daily life.

There are many examples of confirmation bias. One such example (Wason's four-card problem) is considered here.

Confirmation bias can be illustrated by Wason's four-card problem (1968). Suppose you have the following four cards, and you know that each card has a letter on one side and a number on the other side. You are also given the following rule: "*If a card has a vowel on one side, then it has an even number on the other side*". Your task is to "name those cards, and only those cards, that need to be turned over in order to determine whether the rule is true or false".



▲ Figure 2.18 Wason's four-card problem

When Wason gave this and similar tasks to his participants, he observed that the most preferred answers were "A only" and "A and 4". Very few participants picked card 7. However, this is not logically correct—here are the reasons why:

- If you turn over the first card (A) and there is an even number on the other side, it will support the rule. If there's an odd number on the other side, it will refute the rule. Therefore, turning over the first card is potentially very informative.
- If you turn over the second card (D) and there is an even number on the other side, it tells you nothing about the rule. Neither does an odd number. Since the letter is a consonant, any result will neither support nor refute the rule that you are testing.
- If you turn over the third card (4) and there is a vowel on the other side, it will support the rule. However, if there is a consonant, it will not tell you anything about the rule because the rule does not say that consonants must be coupled with odd numbers.
- If you turn over the fourth card (7) and there is a consonant on the other side, it tells you nothing about the rule. However, if there is a vowel on the other side, it tells you that the rule is incorrect.

Overall, the two options that can potentially refute the rule are A and 7. These are the correct choices from the logical point of view.

Apparently, in their everyday thinking people do not follow scientific logic. They frequently pick the third card (4) which can only support the rule but not refute it—a useless card from the logical perspective. Why do they do that? Because presumably they tend to seek out information that confirms their pre-existing beliefs, i.e., confirmation bias. This would also explain why participants in Wason's experiments almost always failed to turn over the fourth card (7).

Example 3: anchoring bias

SAQ Anchoring bias

Anchoring bias occurs when an individual's decision is influenced by the initial piece of information provided, even when that information is not very relevant. An obvious example of this is asking price. Suppose you are buying a pair of shoes at a market where there is some room for bargaining. You ask the seller what the price is, and they give you a number. This asking price then becomes the "anchor". If you go on bargaining, you may eventually settle on a lower price and it will appear like a good bargain, but it depends a great deal on how high the initial anchor was set.

Anchoring bias deviates from normative reasoning because logically our calculation of probabilities should not be affected by irrelevant information. Apart from real-life situations such as price negotiation, other examples that illustrate this bias are the following studies of Strack and Mussweiler (1997) and Englich, Mussweiler, and Strack (2006).

Strack and Mussweiler (1997) was the study that first showed the empirical discovery of anchoring bias. Researchers randomly split participants into two groups and asked them the following questions:

Group 1: Did Mahatma Gandhi die before or after age 9?

Group 2: Did Mahatma Gandhi die before or after age 140?

Both of these questions are, of course, quite unreasonable even for someone who doesn't know exactly who Gandhi was, and participants in both groups responded correctly (after age 9, before age 140).

Another question on the same questionnaire asked participants at what age they thought Mahatma Gandhi died, and this is where differences were revealed. Students in Group 1 guessed age 50 on average, whereas the average guess in Group 2 was age 67. This shows how, although the first question was clearly unreasonable and not relevant to the subsequent task, the question served as an anchor and influenced students' judgements.

To show that the bias shows itself in real-life situations as well as in laboratory conditions, Englich, Mussweiler, and Strack (2006) conducted a study with a sample of practising judges. The judges had to consider a hypothetical case and arrive at a decision about the most appropriate punishment in that case. Part of the hypothetical case was the severity of the sentence demanded by the prosecutor. To make this parameter clearly random, the judges were asked to throw dice and take the outcome as the number of years of jail sentence requested by the prosecutor.

Results of this study revealed that there was a positive correlation between the final sentence awarded by the judge and the number on the dice. The larger the number they threw on the dice, the more severe the sentence they arrived at. The judges fell victim to anchoring bias. Note that the severity of the sentence demanded by the prosecutor, rationally speaking, should not have any effect on the judge's reasoning. Moreover, since the judges could so clearly see that this is a parameter determined at random (they threw the dice themselves!), they should not have taken that information into account at all. However, they clearly did, although they were probably not consciously aware of it.

Activity

Take the three cognitive biases considered in this unit (framing effect, confirmation bias, anchoring bias) and think about how you might be affected by these biases in your personal everyday life.

ATL Thinking, Self-management

Activity

The list of heuristics and cognitive biases discovered to date is impressive. Review a list of cognitive biases that you can find online (a good place to start could be the Wikipedia article called "List of cognitive biases"). Choose one bias or heuristic that captures your attention. Conduct further research into that bias/heuristic and present your findings to the class. Ideally you should find a research article that reports results of an empirical study investigating this bias.

Think of possible practical applications of your chosen bias or heuristic.

ATL Thinking, Communication

Chat with AI

Analyse signs of cognitive biases in a real-life conversation, with the help of AI.

Take a transcript of a conversation. For example, download subtitles to one of your favourite movies and retrieve a memorable conversation between the movie characters. Provide this transcript to your favourite generative AI with a prompt like this:

Here is a transcript of a conversation. Could you analyse this conversation for any signs of cognitive biases? Name the cognitive bias, say who of the people might have fallen victim to it, and give evidence from the transcript. Please only look for the following cognitive biases: [provide a list of biases].

Make sure to not take the AI output for granted. Analyse the conversation yourself and see if you agree with the AI suggestions.

Debiasing

It is understandable that people may want to reduce the influence of cognitive biases, if not to get rid of them entirely in some situations. **Debiasing** is the collective name for the group of methods and techniques designed to do so.

One could distinguish between the following three groups of **debiasing strategy**:

1. **Motivational**. This includes holding people accountable for their decisions or incentivising them to make rational choices. It relies on the assumption that people are capable of normative (rational) reasoning if they are motivated.
2. **Cognitive**. Cognitive strategies usually prescribe context-specific rules that are designed to overcome heuristics—for example, "consider the opposite". Another example would be the use of counter-stereotypical information, such as the Obama effect (see Unit 2.9: Potential for improving cognitive processes).
3. **Technological**. Technological strategies involve using external tools, such as computer-based decision support systems or even printed decision algorithms.

Ludolph and Schulz (2018) conducted a systematic review of debiasing on decision-making in a medical context. The review included 87 studies. The reviewed studies showed that 69% of debiasing interventions were partially or completely successful. The authors of 29% of the studies, on the other hand, concluded that their debiasing attempt was ineffective. In terms of different types of debiasing strategies, technological strategies appeared to be more effective (88% effective interventions), followed by cognitive strategies (50% effective interventions).

Overall, the study suggests that debiasing techniques in the medical field are indeed effective. However, studies like this one should always remember the potential role of publication bias. It could be that only "successful" studies tend to be published, and if that is the case, then estimates of effectiveness would be inflated.



▲ Figure 2.20 Confirmation bias

Conceptual analysis

Bias

The idea of bias can be looked at from various angles, two of which are: (1) humans may be biased in their behaviour, and (2) researchers may be biased when studying this behaviour. It is the inherent biases of human thinking that is the focus of Kahneman's theory. Of course, it is also always a possibility that a researcher may be biased when studying bias in human thinking.

It is not a new idea that people's decisions may be biased. It was understood well before the dual process theory. The significant contribution of the theory is the understanding that bias itself is predictable. We can predict exactly how people will be biased: in what way, to what extent, in which situation. Dan Ariely, one of the most popular and influential scientists in behavioural economics, called one of his books *Predictably irrational*, which captures this idea perfectly (Ariely, 2009).

Responsibility

If bias may be predicted, we can use it in many ways. One option is to use this knowledge to manipulate people's decisions. For example, it can be used in political campaigns to design messages and public appearances. It can be used by businesses to persuade people into buying more products they do not really need.

Knowledge of cognitive biases may also be used for noble purposes. Think about training doctors to avoid racial bias, or training judges to recognize and override the anchoring effect, or training pilots to recognize when there is a high risk of human error.

Causality

Most of the things we know in the study of decision-making come from carefully controlled experimental studies. This allows us to make cause–effect inferences: A influences B. For example, we know that framing a situation in terms of potential losses will cause participants to make riskier decisions (the framing effect).

In all true experiments, it is of paramount importance to control potential confounding variables because they can compromise internal validity and impair our ability to infer causation.

Change

Many cognitive biases are hard-wired into our brain. It takes reflexivity and critical thinking to override them with more realistic, rational decisions. To illustrate this, we could use the metaphor of an airplane and autopilot: our mind is the airplane, System I is the autopilot, and System II is the human captain. Our plane flies on autopilot most of the time. However, we have the ability to learn when to trust and when not to trust the automatic controls, and to override them at those crucial moments. A proficient captain is not the one who flies without autopilot but the one who knows when to take over.

Researchers and practitioners have designed and trialled various de-biasing programmes. Earlier in this unit (see page 140), we looked at Ludolph and Schulz's (2018) review of the effectiveness of these programmes in the context of medical decision-making.

Measurement

Not every irrational thought is a cognitive bias. One needs to be very careful about construct validity in order to be certain that what is being measured in a research study is a cognitive bias and not something else.

A useful way to think about bias is as a decision that systematically deviates from the normative model. Any measurement of cognitive bias implies that there exists a normative way of thinking about the problem, that you know what the normative solution is, and that you register a predictable, systematic deviation of participants' judgements from this normative solution. The normative model in Tversky and Kahneman's (1981) experiment is the expected-utility theory. The normative model in Wason (1968) is propositional logic. In Strack and Mussweiler (1997), the normative model is that our judgements should not be influenced by preceding irrelevant questions.

2.6 Social learning theory

Inquiry questions

- To what extent can people learn by observing the behaviour of other people?
- Is the "free will versus determinism" debate just an incorrectly asked question?
- Is social learning theory a better explanation of human learning than behaviourism?

What you will learn in this section

Key learning:

- Social learning was a reaction to the behaviourist assumption that all learning is direct and involves a trial-and-error process.
- The main proposition of social learning theory was that learning can occur indirectly (observational learning). When developed further, it introduced a number of cognitive mediating variables that play a role in this process, such as identification with the model, self-efficacy, and motivation.
- Bobo doll experiments supported the hypothesis that children imitate aggression when they see adult models behaving aggressively.
- When developed even further into a general model of human behaviour, social learning theory introduced the notions of human agency and reciprocal determinism.
- Reciprocal determinism suggests that there are bidirectional interactions among three groups of variables—environmental factors, personal factors, and behaviour.
- Cognitive dissonance is the phenomenon that explains how beliefs can change because a certain behaviour has been performed.
- Reciprocal determinism also suggests that such age-old debates as "nature versus nurture" and "free will versus determinism" are inappropriate.

Key terms: social learning theory, observational learning, direct learning, indirect learning, mediating factors, attention, retention, motor reproduction, motivation, Bobo doll, self-efficacy, human agency, vicarious punishment, vicarious reinforcement, reciprocal determinism, cognitive dissonance, bidirectional interactions, nature versus nurture, free will versus determinism

In a wider context

The idea of social learning was proposed as a reaction to the behaviourist assumption that all learning is direct and involves a trial-and-error process. Social learning suggests that we can learn by simply observing the behaviour of others.

In a wider context, social learning is a type of learning. In the beginning, it was still very close to principles of conditioning. However, it was also a step towards cognitive psychology: social learning occurs "in one's head", in the "black box".

When developed further, **social learning theory** proposed a number of cognitive variables that impact social learning, such as identification and **self-efficacy**. As an overarching model of human behaviour, social learning theory also introduced the notions of **human agency** and reciprocal determinism to explain causality.

From introducing the idea of observational learning to supplement behaviourism, it grew into one of the most influential models of human behaviour in its own right.

Exam tip

Sometimes a content point that is required for Paper 1 Section C is the same as a content point that is required for Paper 1 Sections A and B. For example, social learning theory is on the list of content that can be asked in Paper 1 Sections A and B. At the same time, social learning theory is one of the content points embedded within the contexts. This is why it appears as a separate section in this unit. Social learning theory can also appear in questions in Paper 1 Section C.

The difference is in the expected depth of response and critical analysis. When asked in Paper 1 Sections A and B, the expectation is that you can explain what social learning theory is and support your explanation with an example. No critical analysis or evaluation is required. When asked in Paper 1 Section C, it becomes an extended response that should be driven by critical analysis and evaluation, as well as connections to concepts.

An overview of social learning theory

SAQ Social learning theory

Social learning theory was formulated by Albert Bandura in 1977. Initially it was intended to explain the phenomenon of observational learning, but it expanded over years into a much more general theoretical framework explaining all human behaviour.

Here is a brief summary of the main ideas of social learning theory:

- **Observational learning** can occur by observing another person's actions and without a trial-and-error process. The person whose behaviour we are learning from is called a model.
- Consequences of the model's behaviour can affect learning indirectly through vicarious punishment and vicarious reinforcement.
- There are cognitive factors that mediate observational learning, such as: attention (identification with the model), self-efficacy ("I can do it" beliefs), motivation (anticipation of consequences).
- The first studies with a Bobo doll demonstrated that imitation of behaviour is most likely to occur if the child identifies with the model.
- Social learning theory culminated in the principle of reciprocal determinism. It suggested that there are bidirectional interactions among three groups of variables: environmental factors, personal factors, and behaviour.

In this section we will further explore each of these points. There will be a number of research studies and examples to support the theory. See also Unit 5.6: Social learning in group behaviour and Unit 4.9: Social learning and health problems.

Observational learning

Albert Bandura, the author of social learning theory, proposed that all learning can be divided into two types:

1. **Direct learning** occurs when an individual performs an action and experiences its consequences (positive or negative). This is described by the principles of conditioning.

2. **Indirect learning** occurs when an individual observes another person's actions and their consequences. This is the same as observational learning.

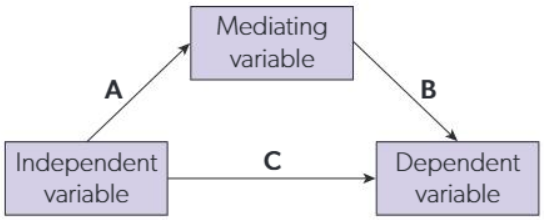
The learning described in the behaviourist approach was all trial-and-error. However, Bandura felt that this was not enough to explain the complex process of socialization, including learning such things as language or religious practices.

The idea that learning can occur indirectly through observation has the following implications:

- It is not necessary to perform the actual behaviour for learning to occur.
- In our learning we depend on available models—people whose behaviour we observe.

Every child and adult is surrounded by multiple impactful models, such as parents, teachers, characters in TV shows and books, friends in a peer group, and popular artists and musicians. These models serve as sources of behaviour to observe and imitate.

Not every behaviour is automatically imitated. There is a range of cognitive factors that may increase or decrease the likelihood that a certain observed behaviour will be learned and/or imitated. All these cognitive variables, according to social learning theory, serve as **mediating factors** between stimuli and responses.



▲ Figure 2.21 Mediating variable

In an early version of the theory, Bandura proposed four mediating factors of observational learning:

1. **Attention**. For a behaviour to be imitated, we need to pay attention to it. We are more likely to learn from the model's behaviour when we identify with the model (e.g., similar age, gender, ethnicity, interests and values, etc.).
2. **Retention**. Not only do we need to pay attention to it, but we also need to store it in memory so that later we can retrieve it in order to imitate it.
3. **Motor reproduction**. This refers to our ability to perform the behaviour—or, more accurately, our belief that we are capable of performing the behaviour. We will only attempt to reproduce a behaviour if we believe that we can do it.
4. **Motivation**. We are more likely to imitate the model's behaviour if we are vicariously reinforced for it. **Vicarious reinforcement (or punishment)** is the observed consequences of the model's behaviour. For example, if one student sees that another student gets praised for raising their hand in class, the first student is likely to also start to raise their hand.

Activity

Split into groups. Draw a diagram or any other graphical representation of social learning theory.

Compare what you have created with other groups.

ATL Communication, Social



▲ Figure 2.22 An illustration of a Bobo doll

Activity

Find the original paper of Bandura, Ross, and Ross (1961) titled "Transmission of aggression through imitation of aggressive models". It is available, for example, on the website called "Classics in the history of psychology"—a collection of pivotal historical publications that are now in the public domain. (Some other papers discussed in this chapter can also be found on this website.)

Look through the paper and list all the ways in which researchers controlled potential confounding variables.

ATL Research, Self-management

Bandura, Ross, and Ross (1961): Bobo doll experiment

A famous study that supports the notion of observational learning is the Bandura, Ross, and Ross (1961) "**Bobo doll**" experiment. A Bobo doll is a special 1.5-metre-tall inflatable doll that bounces back when punched.

It was a true laboratory experiment: researchers manipulated the independent variable (type of model) and observed the changes in the dependent variable (children's aggressive behaviour).

Participants consisted of 72 children aged 3–6 years. The children were split up randomly in the following groups:

- Aggressive role model group (24 children)—the adult model behaved aggressively towards the Bobo doll by following a script.
- Non-aggressive role model group (24 children)—the adult model followed a similar script, but aggressive actions were replaced by non-aggressive actions.
- Control group (24 children)—this group had no model.

The study was conducted in three stages.

Stage 1. A child was seated in one corner of a room and an adult model (either male or female) in another corner. The child was given prints and stickers to play with, and the model had a toy set with construction blocks, a mallet, and a 1.5-metre-tall inflatable Bobo doll. The script that the adult model followed in the aggressive condition was to spend some time playing with the toy set and then to turn to the Bobo doll and behave aggressively towards it, both verbally and physically. In the non-aggressive condition, the model just played quietly with the toys.

Stage 2. The child was taken to another room and a frustrating situation was created to "instigate" the child's aggression. The child was given attractive toys to play with, but when the interest was sparked and play began (usually within around two minutes), the experimenter explained to the child that they had changed their mind and had decided to give the toys to other children. After that, the child (presumably frustrated) was taken to the third room.

Stage 3. The third room contained a one-way mirror. The room had toys similar to those in the first room, including a slightly smaller version of the Bobo doll. Observation at this stage lasted 20 minutes. The researchers had a checklist of behaviours to observe, including such categories as imitative aggression and non-imitative aggression.

The model at Stage 1 "aggressed towards" the Bobo doll in distinctive ways: in addition to simply punching the doll, the adult lay the doll on the side, sat on it, and punched it repeatedly on the nose. Then the model raised the doll back up and hit it on the head with a mallet. After that, the adult model tossed the doll in the air and kicked it about the room. All such acts are unlikely to appear spontaneously. If they were observed in the child's behaviour at Stage 3, they were classified as "imitative". Any novel ways of behaving aggressively towards the doll were scored as non-imitative aggression.

The main result of the study was that exposure to the aggressive model increased the frequency of aggressive behaviour among the children. Children displayed

both imitative acts of aggression (hitting with a mallet, tossing in the air) and novel non-imitative ones (e.g., shooting a dart gun at the doll, behaving aggressively towards objects other than the Bobo doll, remarks such as "Shoot the Bobo", "cut him", and so on).

An additional finding was that boys were more likely to copy the behaviour of male models and girls were more likely to copy the behaviour of female models. This demonstrates identification with the model (attention) as a mediating variable in observational learning.

Further development of social learning theory

After his initial discovery of observational learning, Bandura paid increasing attention to investigating the mediating factors such as attention, efficacy beliefs (motor reproduction), and motivation. According to behaviourists, all these phenomena belong in the "black box". However, remaining true to the principles of rigorous science, Bandura did not mystify these phenomena but looked for a scientific explanation for each.

For example, regarding motivation, this is how he put it: "Anticipatory capacities enable humans to be motivated by prospective consequences" (Bandura, 1977, p. 18). This means that motivation has a clear explainable origin: we observed in the past that a particular behaviour resulted in good outcomes, so now we are anticipating the same outcomes. This allows motivation to be measured and studied scientifically.

Another important concept that was explored in the framework of social learning theory is self-efficacy (this concept grew out of "motor reproduction" in the earlier version of the theory). Self-efficacy is an individual's belief that they will be capable of acting in a way that will help them achieve the goal. In other words, it is the belief "I can do it". Bandura believed that self-efficacy explains a large variety of observed human behaviours.

Eventually when the theory acquired many additional components and broke entirely free from behaviourism, Bandura renamed it to social cognitive theory.

Reciprocal determinism

Once social learning theory developed to the point where it became a general theory of human behaviour (rather than a theory explaining only some aspects of it), the theory culminated in the concept of **reciprocal determinism**.

Reciprocal determinism suggests that there are **bidirectional interactions** among three groups of variables—environmental factors, personal factors, behaviour. "Bidirectional" means that the influence is mutual. For example, personal factors affect behaviour, but behaviour also affects personal factors. Personal factors can be thought of as "internal", environmental factors as "external", and behaviour as the bridge between the two. According to Bandura, this balanced approach avoids the reductionism of other theories that attached absolute importance to only one of the factors: social learning theory "neither casts people into the role of powerless objects controlled by environmental forces nor free agents who can become whatever they choose" (Bandura, 1977, p. vii).

Chat with AI

Try using your favourite generative AI to help you think of some possible links between the content you are learning and the six key concepts (causality, measurement, perspective, bias, change, responsibility). However, remember that you will have to provide some background information to AI because it may lack an understanding of what these concepts mean in IB Psychology.

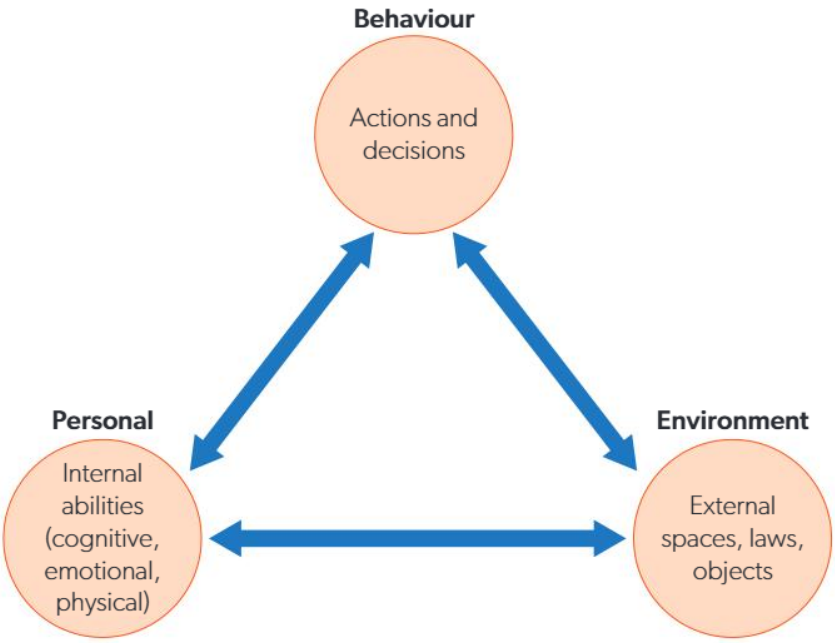
Prepare a brief description of each of the key concepts with examples. Save this text somewhere because you will need it for other topics. Separately prepare a brief summary of what you have learned about social learning theory, with some key words and most important statements. This is to give the AI an idea about how much depth is required and what exact aspects of the theory you have been learning about.

Then consider using a prompt like this:

I am an IB student studying an introductory psychology course. The course is based on the following six key concepts: [insert the description of the key concepts]. Here is a summary of what I have learned about social learning theory: [insert your summary of the topic]. Could you help me identify some possible ways in which I can link this material to the six concepts?

Of course, in order for this exercise to be useful, you should not expect the output to be perfect. Use your own reasoning and continue having a conversation with AI to make the result more relevant and insightful.

Graphically, reciprocal determinism may be represented like this:



▲ Figure 2.23 Reciprocal determinism

Let us consider each of the bidirectional influences.

A. Personal factors and behaviour

Personal factors include anything internal to the person, such as genetics, dispositions and preferences, personality traits, expectations, values, beliefs, and cognitive processes. There is no doubt that these factors affect behaviour: beliefs influence actions.

However, the opposite is also true: actions influence beliefs. This direction of causality may be less frequently studied in research, but being less studied does not make it less important. One example of how behaviour can influence personal factors (beliefs) is **cognitive dissonance**. When our behaviour contradicts our beliefs, we often prefer to change the beliefs.

SAQ Cognitive dissonance

Cognitive dissonance is the mental stress caused by the inconsistency between one's behaviour and one's beliefs. An individual who experiences such inconsistency feels stressed and uncomfortable and is driven by a desire to reduce dissonance. There are only two logical ways to reduce dissonance between beliefs and behaviour: change your behaviour or change your beliefs. Common sense suggests that beliefs drive behaviour, but research has actually shown that behaviour drives beliefs to a much greater extent.

The phenomenon of cognitive dissonance was first described by Leon Festinger in 1956 after his series of observations in a religious cult in Chicago called The Seekers (Festinger, Riecken, and Schachter, 1956). Members of this cult believed in an imminent apocalypse and took strong actions to support their belief. The leader of the cult, a self-proclaimed prophet, "received

messages" that the world would end at midnight on 21 December 1954. They also believed that they were selected to be the survivors and start a new civilization, and that for this a spacecraft was coming to pick them up and that they had to prepare for departure. As part of the preparation they had given up their jobs, spouses, and possessions, which shows how committed they were to the belief. Festinger infiltrated the group to conduct participant observations and collected valuable information on what happens if strong beliefs are not confirmed by evidence.

After the prophecy failed (the spaceship did not pick them up at midnight!), the group spent several hours in a state of tension which is very well conveyed in Festinger's detailed observation report. At 4.45am, the leader went to a separate room and received another "message" which said that the apocalypse had been called off because the members of the religious group saved the Earth by their pure thoughts and efforts: "The little group, sitting all night long, had spread so much light that God had saved the world from destruction" (Festinger, Riecken, and Schachter, 1956, p. 169). Interestingly, although the group had been notoriously closed to the outside world and had never given interviews, on the morning of the following day they began a campaign to spread the message as far as possible, calling newspapers and setting up interviews. This shows how people change their beliefs when it is impossible to change (take back) their behaviour.

Cognitive dissonance is often used in compliance techniques: it links to the principle of "commitment and consistency" proposed by R. Cialdini (see Unit 5.5: Compliance techniques). For example, foot-in-the-door is the compliance technique where you first get the person to do you a small favour and then formulate a larger request. When someone complies with your smaller request, they perform a behaviour that possibly contradicts their existing beliefs. Later when you formulate the larger request, they will experience cognitive dissonance: they do not believe they should help you, but then why did they help you before? Since they cannot take back their behaviour, the only way to reduce dissonance is to change their belief: perhaps they want to help you after all! Cognitive dissonance is known as one of the most powerful mechanisms of influence (Cialdini, 2007).

B. Personal factors and the environment

Again, it is intuitively obvious that the environment shapes us. However, the opposite is also true: personal variables affect the environment.

For example, all your classmates attend the same school and sit in the same classes, therefore they are immersed in the same environment, but they are not all equally affected by it. This is because they choose to attend to different aspects of this environment, and to attach varying degrees of importance to it. Their personal variables influence which aspects of this common environment they will ultimately be surrounded with.

C. Behaviour and the environment

Similar bidirectional relationships exist between environmental variables and behaviour. On the one hand, it is obvious that the environment influences behaviour. This is what behaviourism is all about.

Discussion

Experiments are great for investigating a unidirectional influence of one variable on another, such as the influence of the environment on behaviour.

But what do you do if you want to investigate all three groups of variables at the same time, allowing them to reciprocally influence each other and dynamically develop over the course of time? Experiments are great for capturing the idea of "influence", but they fall short when we want to capture a reciprocal "interaction", let alone a dynamically developing one.

Any ideas?

ATL Research, Thinking

However, our behaviour in real-life has a degree of influence on the environment itself. For example, the way we behave towards people determines how they will behave towards us, and therefore what kind of social environment we will find ourselves in. "Lecturers do not influence students unless they attend their classes, books do not affect people unless they select and read them, fires do not burn people unless they touch them" (Bandura, 1977, p. 195)

Chicken and egg in reciprocal determinism

You may be familiar with such debates as "**nature versus nurture**" and "**free will versus determinism**". The former asks if it is biological factors (nature) or environmental factors (nurture) that is the primary cause of human behaviour and development. The latter asks if our actions are strictly determined by preceding causes, or whether we are free to shape our behaviour ourselves.

The notion of reciprocal determinism suggests a unique answer to both debates: it suggests that the debates themselves are artificial and inappropriate. In other words, the question is asked incorrectly.

This is because the debates assume the factors to be independent of each other, but in reality, they are not. Nature can influence nurture and the other way around. Reciprocal determinism in social learning theory also suggests that free will and determinism are two sides of the same process, depending on how you choose to look at it.

Attempts to assign a causal priority to either personal or environmental factors reduce the problem to the "chicken-or-egg" argument (what came first, the chicken or the egg?). However, as Bandura put it, "In the regress of prior causes, for every chicken discovered by a unidirectional environmentalist, a social learning theorist can identify a prior egg" (Bandura, 1977, p. 203).

Conceptual analysis

Perspective

Social learning theory grew out of behaviourism, initially introducing only one additional idea that people are capable of learning by imitation, but eventually becoming progressively more and more cognitive until ultimately it was renamed social cognitive theory. It makes prominent the role of cognitive factors in regulating human behaviour. It also emphasizes that humans are social creatures, therefore their learning is social. It is more holistic than some other theories, in the sense that it considers the interplay among multiple groups of factors regulating behaviour.

Causality

The idea of observational learning suggests that people may learn by imitating the behaviour of others, but only if there are favourable cognitive conditions (such as attention, motivation, and self-efficacy). This inclusion of cognitive factors as "mediating variables" was an important step towards allowing cognitive processes into the focus of scientific investigation.

The idea of reciprocal determinism considers bidirectional causality in three groups of factors that constantly influence each other: personal factors, environmental factors, and behaviour.

Bias

On a specific level, there exist dangers of introducing bias in research studies because the topics under investigation are frequently sensitive (e.g., aggression) and the research methods include registering participant behaviours through techniques such as a structured observation.

On a more general level, social learning theory, through its idea of reciprocal determinism, tries to eliminate bias associated with a one-sided view of human behaviour as a product of environmental influences (environmental determinism).

Measurement

As you have learned, social learning theory postulates the existence of bidirectional influences (for example, environment influences behaviour, but behaviour also influences environment). However, the experiment can only model unidirectional influences of an independent variable on a dependent variable. It seems impossible to measure everything at the same time, and the idea of reciprocal interaction is difficult to capture in measurement.

Change

The concept of change is related to social learning theory because it provides a practical way to change people's behaviour through exposing them to role models. Social learning theory has therefore found a variety of applications in education, social work, health promotion, mass media, and other areas (see Units 5.6: Social learning in group behaviour and 4.9: Social learning and health problems).

Responsibility

Research studies under the umbrella of social learning theory have often focused on sensitive topics such as learning aggression. In Bandura, Ross, and Ross (1961), children were exposed to aggressive behaviour, then frustrated, then given a chance to behave aggressively themselves. One may claim that it may have done the children some harm because the experiment itself taught them to behave aggressively. In other studies, children had to watch violence on TV and in cartoons and then were given a chance to reproduce that behaviour in real-life situations.

Social learning theory has also been used as a guiding principle for designing numerous interventions and prevention programmes for mental disorders and health problems. Researchers bear the responsibility for implementing such studies with the highest standards of credibility because our findings in this area directly translate into the extent to which we can help people who experience life problems.

2.7 Cultural factors in cognitive processes

Inquiry questions

- Do people from different cultures remember things in the same way?
- How can we study culturally specific ways of processing information?
- Are we only able to fully understand another culture if we become part of it?

What you will learn in this section

Key learning:

- Cultural factors in cognition can be studied in a variety of ways, most prominently by comparing cognitive functioning of individuals who grew up in different cultures.
- Some cognitive abilities are universal. All people and even newborn babies have them no matter what culture they come from.
- Other cognitive skills are less universal. For example, depth perception may depend on cultural experiences and is not present at birth.
- Two approaches can be used for cultural research: the etic perspective and the emic perspective, which can be thought of as a perspective of an "outsider" and an "insider".
- A danger in using the etic approach to study a new culture is the possibility that the culture will be understood through the lens of concepts that are alien to it—the imposed etic bias.
- An illustration of this is the application of the concept "intelligence" together with IQ tests to compare cognitive functioning across cultures. The Western idea of intelligence may not be appropriate in some cultural contexts.
- Berry's proposed steps in research were imposed etic, emic, derived etic, and cultural universal.

Key terms: culture, depth perception, emic approach, etic approach, imposed etic, derived etic, cultural universal, free recall, narrative recall

In a wider context

Cultural factors in cognitive processes become visible when we compare individuals who grew up in different cultures. **Culture** plays a key role in how cognitive processes are acquired, shaped, and performed. This topic is closely connected to a number of others, especially:

- Sociocultural factors in development (Human development)
- Stage theories and continuous models (Human development)

- Cultural dimensions (Human relationships)
- Culture (HL extension)

Cultural differences in cognition

There certainly are cognitive universals—characteristics of cognitive processes that are shared among all humans. For example, binaural hearing—the ability to hear with both ears. It also comes with the ability to detect which direction the sound is coming from. This ability is important in the animal world because it allows an animal to quickly detect the source of danger and run away from it. No matter what culture people belong to, they are cognitively capable of processing this information. Even newborn babies can detect the source of a sound and turn their heads towards it.

On the other hand, there are components of cognitive processes that are less universal. Take for example **depth perception**. This is the ability to perceive the world in three dimensions and understand that more distant objects will appear smaller.

Anthropologist Colin Turnbull described the time he spent among the Bambuti Pygmies in the Ituri forest in Congo in the 1950s. His local guide Kenge grew up in a thick forest and never experienced vast open spaces. Kenge travelled with Turnbull and, for the first time in his life, he saw a herd of buffalo grazing on a plain several miles away. Kenge thought these were some kind of insects. His brain could not process the idea that these were large things far away, so they only appeared small. When Turnbull drove towards the buffalo and Kenge saw them rapidly increase in size, he got scared and thought that some witchcraft was involved (Turnbull, 1961).



▲ Figure 2.24 Grazing buffalo

Exam tip

There are three HL extensions: culture, motivation, technology. Each of them provides an additional lens through which to consider the four contexts (Learning and cognition, Health and well-being, Human relationships, Human development).

If you are an HL student, consider studying this unit in conjunction with Unit 6.10: Culture in learning and cognition, as there is a lot of overlap between these two units, and they enrich each other.

The emic and etic approaches

The exploration of cultural influences in cognitive processes brings forward the fundamental difference between the emic and the etic perspectives in psychological research.

SAQ Emic approach and etic approach

The **emic approach** in cultural studies is the perspective of an insider who is trying to understand a culture from the point of view of its people. It seeks to understand the unique beliefs, values, and practices of a given society and pursues no goal of comparing it to other cultures. Researchers working from the emic perspective frequently conduct qualitative studies by being immersed in the culture of interest and studying it from the inside. They often live in that culture for a prolonged period of time.

In contrast, the **etic approach** in cultural studies is the perspective of an outsider. They are trying to understand a culture from the point of view of an external framework. They apply general theories and concepts to this new culture and analyse its similarities and differences with other cultures. Researchers working from this perspective often employ standardized measures that allow them to quantify cultural similarities and differences (e.g., Geert Hofstede's cultural dimensions—see Unit 5.2: Cultural dimensions).

The emic approach makes it necessary to adopt an insider's point of view, whereas the etic approach requires the adoption of an outsider's point of view.

It is argued that "emic and etic data do not constitute a rigid dichotomy of data, but often present the same data from different points of view" (Pike, 1967, p.41). A dichotomy is a division between two things that are presented as entirely opposite to each other. On the point of emic and etic data not being a rigid dichotomy, note (for example) that psychologists rarely think of emic findings as the end of the investigation. The study of unusual behaviours in a remote society is interesting by itself, but its real value lies in its ability to change what we know about other cultures and culture in general.

A prominent problem in cross-cultural research is the **imposed etic**. It occurs when the study attempts to use measures and categories that are not appropriate in the context of a given culture.

In this section, we will consider Cole and Scribner's (1974) research as an example of (mostly) emic research of memory. The etic approach is closely linked to Hofstede's idea of cultural dimensions. There are other examples throughout this book also.

Imposed etic

Imposed etic is the problem of studying a new culture by using concepts and tools that are natural and relevant in the researcher's culture but could be alien and inappropriate in the culture under investigation.

In order to make cross-cultural comparison possible, one needs to define the comparison criteria beforehand, then apply them consistently to each culture. An example of such "criteria" are the cultural dimensions defined by Geert Hofstede, such as individualism–collectivism, masculinity–femininity or power distance. According to Hofstede, these are universal characteristics that

capture cross-cultural differences. For example, all cultures may be positioned somewhere between the two extremes on the individualism–collectivism scale. However, there is no guarantee that the scale itself is of equal relevance to all the cultures that it is trying to describe. See more about cultural dimensions in Chapter 5: Human relationships.

Another example is the idea of intelligence, which originated in European research with Western children attending a Western educational system. IQ tests are designed to assess children's performance on cognitive tasks that Western children are typically required to perform. These are to categorize objects, find differences and similarities, and apply logical analysis.

When cross-cultural research begins, researchers may be interested in studying "intelligence" in another cultural context. However, the researcher is equipped with the original construct of intelligence and the original measures of it (IQ tests). Applying this construct and these tests to the new culture is imposed etic. It could be that the idea of a "cognitively competent individual" in the new culture is different, and that an intelligence test will be a completely invalid measurement.

Derived etic and cultural universals

You may think that imposed etic is a negative thing, but John Berry (1989) pointed out that "all research must begin somewhere" (Berry, 1989, p. 726). He said that the initial step in any cultural study is when the researcher takes a concept or tool that is emic to their own culture and, with that concept or tool, attempts to study another culture. Since the concept or tool is brought to the new culture from the outside, Berry called this approach an "imposed etic".

As the researcher gets more and more familiar with the new culture, they become capable of setting aside the imposed etic and exercise a more emic approach. Once the emic approach is firmly established and the researcher has acquired concepts and tools that are meaningful in that particular culture, the culture can now acquire its truly emic characteristic.

However, the process does not end here. If the researcher now follows the same process with some other culture, its emic characteristics may appear to be similar to the emic characteristic of the first culture. Berry called these commonalities that emerge from emic research the "**derived etic**". They allow us to make cross-cultural comparisons, but unlike imposed etic, they were derived inductively from the emic study of each separate culture.

Hypothetically, we can imagine a situation when all existing cultures have gone through the process and some commonalities have been identified. These commonalities would then be referred to as **cultural universals**.

In summary, the research process that Berry proposes is:

1. imposed etic
2. emic
3. derived etic
4. cultural universal.

Chat with AI

Do some research to find more examples of imposed etic. If the generative AI you are using has the function of carrying out an internet search and providing links to its sources, then you might find the following prompt fragments useful:

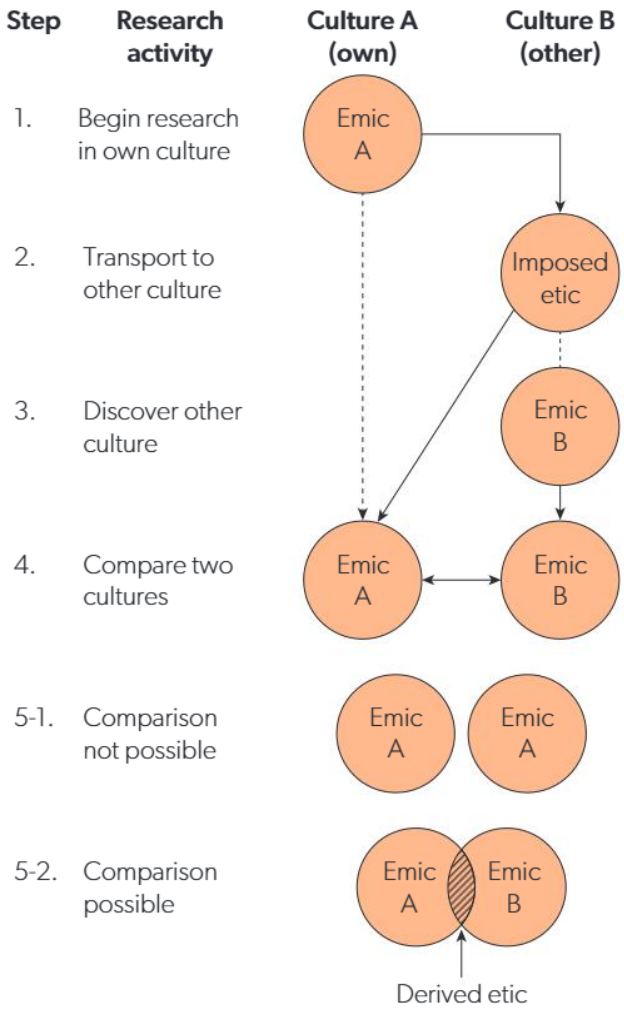
- Could you provide several examples of psychological research studies where the emic approach was used to understand participants' behaviour and cognition? Please provide a link to each study along with a brief summary and an explanation of why you think the emic approach was used in the study.
- Could you provide examples of imposed etic in psychological research—situations when a culture was studied through the lens of concepts and measures that were alien to this culture?

TOK

The emic and the etic approach can be broadly thought of as the "insider perspective" and the "outsider perspective". Slightly outside of cross-cultural context, the terms that bear a similar meaning are nomothetic research and idiographic research.

Which of the perspectives is more valuable in the acquisition of knowledge? What does it depend on?

Is the study of history emic or etic?



▲ **Figure 2.25** The process of obtaining derived etic

In emphasizing how both the etic and emic approaches are necessary in cross-cultural research, Berry said: "We cannot be 'cultural' without some notion like *emic*; and we cannot be 'cross' without some notion like *etic*" (Berry, 1989, p. 729).

The study of memory in the Kpelle society

Cole and Scribner (1974) described a series of studies that they carried out among the Kpelle people in Liberia.

In a typical study, their participants would have to memorize one of the following:

1. The clusterable list included objects that came from different categories: clothing (e.g., trousers, singlet, headtie), tools (e.g., cutlass, knife, hammer), foods (e.g., potato, onion, banana), utensils (e.g., plate, pan, cup).
2. The unclusterable list included a list of objects that were difficult to group into categories (e.g., bottle, nickel, chicken feather, box).

The experimental approach that they started with was **free recall**. Free recall is when participants are free to remember the objects in any manner they choose.



▲ Figure 2.26 A Kpelle village in Liberia

Results of Cole and Scribner's free recall experiments demonstrated the following:

- American children, especially those 10 years old or more, use clustering in their recall. This means that items from the same category (e.g., utensils) are typically said together.
- This was not true for Kpelle children who showed little to no semantic clustering.
- However, Kpelle children who attended school used more clustering, especially older children.

These studies seemed to suggest that Kpelle children's semantic memory is inferior to that of their American counterparts, that they are not so capable of organizing material into logical categories and thus not so efficient at memorizing, and that this improves with schooling.

This could have been the end of the story, but Cole and Scribner noticed that Kpelle participants never recalled items in the same order as they were presented. This is something you would typically expect if a person were using rote memorization. Therefore, it cannot be said that the Kpelle used simple rote memorization. Apparently they did use some organizing principle, but that principle was not the one the experimenters expected them to use.

To probe this further, Cole and Scribner tried another experimental approach where they incorporated to-be-remembered objects into a narrative (a story). It was hypothesized that this kind of **narrative recall** is more relevant to how the Kpelle people organize material in their daily lives. There were two stories used by Cole and Scribner. Story 1 used the same semantic categories as were used in the previous experiments, only this time the categories were wrapped into a relevant real-life situation. Story 2 followed a different kind of organization: a sequential structure where the order of objects was important.

Chat with AI

It is inevitable that when reading an academic text there are certain terms that you are not entirely familiar with, or that you may have doubts about. For example, we have just referred to Kpelle children's semantic memory, although we never gave a precise definition of this kind of memory. Generative AI can be very helpful here. As you are reading, use simple prompts like this:

What is semantic memory in psychology? How is it different from some other types of memory?

Whenever you have doubts, just ask. It is a useful habit to acquire in the interaction with AI.

Activity

Below are two stories adapted from Cole and Scribner's original study of narrative recall (some elements have been shortened and some items have been renamed).

STORY 1. A chief had a beautiful daughter, and many young men wanted to marry her. Following tradition, each of them brought many presents for the young woman and left them with the chief. One brought a knife, a hammer, a pen, and a spade. Another brought an onion, a banana, an orange, and a potato. Another brought a plate, a pot, a pan, and a cup. And another brought a hat, a dress, a shirt, and a shoe.

What things did the young woman receive? Whose presents were the best? Why?

STORY 2. An evil spirit came to town one day, disguised as a very handsome man. He met a young woman who agreed to marry him. On the night they married, she discovered he was an evil spirit. He told her she must come with him to his farm, but she said to wait a bit while she got her things together. She knew where the evil spirit's farm was, and so she put many things on the ground in her house to show her people the way to reach his farm.

She put her *plate* first, since she always ate at home. Then she put the evil spirit's *hat* to show that he took her away. Then she put a *pot* to show that he took her first in the direction of her family's kitchen behind the house. Then she put a *knife* to show that they went past the woodcarver's house. Next was a *dress* showing that they passed the store where she bought it. Next was an *onion* to show they passed the market, and a *cup* to show that they passed the shop where they sell tea. Next was a *hammer* to show they passed the house being built on that trail. She then put down a *pen* to show that the house belongs to the teacher. Next came a *banana* to show they took the road with the banana trees, and a *shirt* to show they passed the place where they wash clothes. Then she put an *orange* to show that they took the trail with the orange tree, and *a shoe* to show they passed the shoemaker's place. Then came a *spade* to show that she was on a farm, and a *potato* to show that it was a potato farm, and finally a *pan* to show that she was at the kitchen at the farm. The woman's people saw all these things and understood where she had gone and came and rescued her from the evil spirit.

Tell all the things she put on the ground and their meaning so that if you were the woman's family, you could find her.

(Adapted from Cole and Scribner, 1974, pp. 133–134)

Use these stories to conduct your own study with your classmates. Which one were they better at?

ATL Research, Social, Self-management

Results of the narrative recall study showed that participants who heard the first story had a strong tendency to recall items by category, whereas participants who heard the second story recalled the items sequentially with little to no clustering.

Therefore, the first conclusion that could have been made based on the study of free recall (namely, that the Kpelle people are not capable of semantic clustering) is incorrect. They are capable of semantic clustering, and they do use it effectively in their recall. However, they only use it when this structuring of information makes sense to them in the context of a real-life situation.

The first explanation based on the studies of free recall was imposed etic: the differences between American and Kpelle children were explained in terms of "good" and "bad" memory, or the "ability to categorize". The second series of studies (narrative recall) posed the question differently: what kind of categorization do traditional Kpelle people apply to recall tasks? How do they decide whether or not to use categorization in a specific situation? This is an example of an emic approach. Apparently, the Kpelle can use semantic categories when required, but it is not their preferred way to remember things.

This conclusion has quite a different connotation.

Activity

If you were sitting on an ethics committee and Cole and Scribner approached you to get an approval for their study, what questions would you ask them?

ATL Communication, Self-management

Conceptual analysis

Perspective

Although we usually refer to the triad "biological–cognitive–sociocultural" in this course as "perspectives", the emic and the etic approach to the study of culture equally deserve this status. We can think of them as the emic and the etic "perspectives".

We have seen in this unit that the emic and etic approaches are fundamentally different, yet, as demonstrated by Berry (1989), both are needed at some stage in the process of research. We also looked at such related phenomena as imposed etic, derived etic, and cultural universals.

Bias

Any study of cultural influences involves cross-cultural comparisons, and any cross-cultural study is susceptible to the effects of imposed etic bias. As Berry (1989) suggested, this bias may be unavoidable in the process of a cultural study. However, what we can do is continue the study to ensure that the bias has been eventually accounted for or removed. Cole and Scribner's (1974) research, that we considered here, is a good illustration of that.

Causality

Applied to cognitive processes, the etic and the emic perspectives lead to a debate: are cognitive processes universal or a product of culture? This is directly related to causality.

- If cognitive processes are universal, this would mean that they are fundamentally influenced by factors shared by all members of the human species, such as the biology of our brain. Any variations observed from culture to culture would be considered non-essential fluctuations. Researchers would not feel a need to make their samples more culturally diverse because, if cognitive processes are universal, it does not matter who you study.
- If cognitive processes are a product of culture, this would mean that cultural factors are superior to biological factors in determining cognition. Cultural variations would be considered important and viewed as essential. Researchers would try to make sure that their samples are as culturally diverse as possible, and that research covers a large number of various cultures.

Measurement

The most obvious challenge is that a true experiment is not possible in cross-cultural psychology. When we compare a sample from one cultural group to a sample from another cultural group, culture is never the only thing that differs between these samples. They may also speak different languages, have different traditions of education, and be exposed to different experiences.

A related challenge is that participants from different cultural backgrounds may not be the same in how they experience the research procedure. What if the instruction to the participants, when translated from one language to another, acquires a slightly different meaning? Translation of research materials and instructions is a challenging task because you need to ensure equivalence. A separate question is whether or not the researcher running the study and

interacting with the participants should be the same in all groups or rather a representative of the respective culture. People may change their behaviour around a foreigner, especially if we are talking about Indigenous communities.

Change

The study of cultural variations in cognitive processes allows us to understand exactly how much in human cognition is innate and how much is acquired (learned). This sheds some light on the evolutionary development: if certain things do not change from culture to culture, then we must have inherited them from animal ancestors. The study of variations within the same culture (for example, how cognitive processes depend on the amount of schooling) sheds some light on individual development.

Responsibility

There is the issue of reporting on the results of the study and applying them in real-life practices. Unsupported, hasty generalizations may be dangerous. It is far too easy to jump to the conclusion that societies such as the Kpelle are somehow inferior in terms of semantic memory abilities—a conclusion that would be a misrepresentation. Hasty generalizations go hand in hand with labelling. Results from cross-cultural research that lack accuracy and diligence can lead to labels being unfairly attached to certain cultural groups.

2.8 Environmental influences on cognitive processes

Inquiry questions

- Can working conditions such as room temperature affect our cognitive functioning?
- How certain are we that having breakfast is good for learning?
- Are effects of poverty on cognition reversible?

What you will learn in this section

Key learning:

- Environmental influences on cognitive processes include poverty, nutrition, and ecological quality.
- Broadly speaking, all social and cultural factors are a part of a person's environment. However, when used in the narrow sense, environmental influences refer to physical rather than social characteristics.
- There is little bidirectional ambiguity in this area: environmental variables affect cognition, but not the other way around.
- Research studies have discovered various effects of working conditions (e.g., noise, lighting, and temperature) on cognitive processing. However, research is complicated by the fact that all such factors act in combination, so results of studies that isolate one variable may not always apply to real-life conditions.
- Breakfast consumption has been shown to be generally beneficial for learning, but it depends on a lot of factors, such as the kind of food consumed. There is a lack of studies investigating long-term effects.
- Poverty is a complex phenomenon that includes a range of interacting factors. There are two competing models: the family stress model and the parental investment model.

Key terms: environmental variables, working conditions, breakfast consumption, glycaemic index, poverty, family stress model, parental investment model, statistical separation of effects

In a wider context

Although broadly speaking, social and cultural influences are all part of a person's "environment", factors such as **poverty**, nutrition, and pollution deserve special consideration. It is of great practical value to understand the impact that these factors may have on cognitive processes.

One may argue that the direction of causality in this area is more clear and less ambiguous than anywhere else. There is bidirectional ambiguity with biological

Exam tip

The range of examples of environmental influences is quite broad. Some of them are more obviously physical, such as air quality or climate (they do affect cognitive processes!). Others may be more social in nature: for example, poverty and malnutrition are linked to the socioeconomic status of the family. The line between physical and social aspects of the environment is blurred. We will consider a range of examples.

Activity

What working conditions do you find most comfortable? Do you like studying with background music or in silence? Do you prefer a brightly lit room or dim conditions? Do you like it when it is warm in the room or slightly cold and fresh?

Exchange your observations with a partner. Do you think your cognitive processes are objectively affected by these conditions or is it simply a matter of personal habit and preference?

ATL Self-management, Social



▲ Figure 2.27 Working conditions

factors and cognitive processes. Biological factors can influence cognition, but cognition can also influence biological factors through neuroplasticity (the ability of the brain to reshape itself in response to experience). The same can be said about social and cultural factors. They influence the development of cognitive processes, but cognitive processes (such as identifying with one's culture) also determine the extent of this influence. Environmental factors are different: insufficient nutrition may affect the development of memory, but it is hard to imagine how memory may affect nutrition.

Effects of working conditions on cognitive functions

One example of "environmental influences" on cognition is the effect exerted by **working conditions**, such as lighting, temperature, level of noise, and ventilation. Do any of these factors in your classroom affect your ability to learn?

Available evidence in this area suggests that passive exposure to heat only leads to a change in cognitive function when body temperature is raised to above 37 °C. Even then, it only affects complex cognitive functions but not simple ones. The reason these changes occur may be related to altering blood flow in the regions of the brain responsible for attention and voluntary control (Taylor et al., 2016). It was also discovered that nutritional supplements containing tyrosine (an amino acid) help maintain cognitive function in extreme environmental conditions. Findings like these might be useful in such occupations as mining, farming, factory work, firefighting, and the military.

Wang et al. (2021) conducted a systematic review of how indoor environmental quality (IEQ) affects cognitive functions. They conducted a review of 66 focused research studies and concluded that indoor environmental quality conditions are "not always associated with reduced cognition". For example, one of the findings was that the influence of noise on attention is complicated. Students work faster in noisy environments, but they also make more mistakes. In other words, in the trade-off between speed and accuracy, noise tends to make us faster. Older people are more affected by noise than young people.

To illustrate how complex the relationship between two variables may be, consider the following two findings:

1. When the room air temperature was 22°C, increasing illuminance from 200 lux to 1500 lux improved attention.
2. However, when the room temperature was 37°C, the opposite effect was found (Mohebian, Dehghan, and Dehghan, 2018).

Research in this area shows that isolation of single variables may be better for internal validity, but it may also compromise applicability to real life. Real-life working conditions are always a combination of factors: noise level, lighting, temperature, air quality, and other factors all exist in combination.

Effects of breakfast consumption on cognitive performance

Breakfast is sometimes referred to as the most important meal of the day, but it is also well known that school students do not like getting up early in the morning. Therefore, they may miss breakfast entirely. If having a healthy breakfast is indeed associated with better cognitive growth, then a huge impact on the cognitive growth of a nation may be achieved by relatively simple campaigns promoting **breakfast consumption**.

Martins et al. (2020) conducted a review of 30 relevant studies. They found general consensus that having breakfast is beneficial for cognitive function. But the exact mechanism of this impact is not fully understood. Evidence points to the role of blood glucose. Some studies have shown that foods that have a low **glycaemic index** (GI) yield better effects on cognitive performance than high GI foods.

Although most studies suggest that low GI breakfasts are beneficial for cognitive performance (versus high GI breakfast or skipping breakfast altogether), results are never unanimous. One study with Australian children included in the Martins review, for example, found no significant differences in cognitive functions in children eating low-GI versus high-GI breakfasts. The only way to resolve this controversy is to conduct reviews and meta-analyses involving multiple studies, and carefully analyse the discrepancy in their results.

Research in this area faced the following typical challenges:

- The sample participating in the study. Effects of breakfast on cognition may be different depending on participants' age, education level, poverty status, and sociocultural background. Most studies will be limited in their sample composition, and their generalizability. This is why reviews and meta-analyses are so crucial.
- The complexity of cognitive processing itself. Breakfast consumption can have different effects on different cognitive functions, or perhaps even different components of the same cognitive function. Does it even make sense to talk about effects on "cognitive processes" in general?
- The complexity of breakfast itself. What matters is not just breakfast or no breakfast, but what kind of breakfast it is. Even though there seems to be consensus about low-GI foods that do not increase blood sugar levels too quickly, any breakfast is always a combination of multiple types of food, and they all affect each other and moderate each other's effects. What would be the formula for an ideal breakfast?

Overall, the review of Martins et al. (2020) suggests that there are positive effects of having a breakfast regularly on cognitive processes (such as attention, memory, and thinking). These effects also translate into better school performance (better subject grades).

They also point out that there is currently a lack of longitudinal studies investigating the long-term effects of breakfast consumption on cognitive development. The overwhelming majority of studies focus on immediate, short-term effects, such as cognitive performance the following day.

Activity

The GI shows how quickly food affects your blood sugar (glucose) once it is consumed. High GI foods cause a rapid increase in blood sugar. They include: sugar and sugary foods, sugary soft drinks, white bread, white rice, potatoes. Low GI foods include fruit and vegetables, and wholegrain foods (e.g., porridge oats).

Make a list of the food items that you had for breakfast today. Look these items up online to find out their GI and estimate the total. Compare your results with a partner.

Discuss how, in your experience, your schoolwork is affected by what you eat.

ATL Thinking, Social, Communication

Activity

Design a creativity, activity, service (CAS) programme for your school to promote healthy breakfast consumption among IB DP students. How would you justify the importance of this project to school management? What research would you draw on?

What factors will be important for you to consider as you plan the project? What risks do you need to avoid?

ATL Thinking, Communication, Self-management



▲ Figure 2.28 A healthy breakfast

Effects of poverty on cognitive functions

Another **environmental variable** that has been of interest to researchers for a long time is poverty. The question is "How does living in poverty affect cognitive processes?" A large number of studies that simply compared children from poor families to children from financially well-off families found that poverty is indeed associated with worse outcomes, such as lower academic achievement. However, such a comparison is flawed.

There are at least two challenges here:

1. We must understand that poverty is not just the amount of money in the family. It is a combination of multiple interacting factors, for example, quality of parenting, rates of crime, working hours, rates of substance abuse, quality of education accessible to the children, and so on.
2. It is not enough to demonstrate that children from poor environments have lower scores on standard tests of cognitive development. Would they get back on track and catch up with the other children once placed in a more financially stable family environment? Or are the effects of poverty on cognitive development irreversible? Does it matter if poverty was experienced early or late in life? Does it matter if it was one separate episode or continuous poverty? To find out, we need to look at children who shifted from one environment to another, and we need to look at their cognitive development longitudinally.

There have been two major groups of hypotheses about which aspect of poverty is responsible for delays in cognitive development. They are:

1. The **family stress** model. This approach places the main importance on such factors associated with low income as the home environment, the quality of parent–child interactions, and the amount of time spent with the child. In other words, not a lack of money itself, but the lower quality of interpersonal relationships that may be associated with it.

2. The **parental investment model**. This approach places emphasis on more material things, such as the ability to provide the child with material goods, services, and experiences. It also includes malnutrition: children in a persistent state of poverty may suffer the effects of poor nutrition that might have physiological consequences for the brain (Barajas, Philipsen, and Brooks-Gunn, 2007).

Dickerson and Popli (2016) used data from the UK Millennium Cohort Study, obtaining a sample of 19,000 children born in the UK from 2000–2001 and tracing their cognitive development up until the age of seven years. All participants in this study were tested when they were aged nine months, three, five, and seven years. At each testing point, information was gathered in face-to-face interviews and included a wide range of characteristics about the child, the family, parenting activities, and cognitive assessments.

One major result of the study is that the timing of poverty episodes was shown to have a large impact on subsequent cognitive development. More specifically, experiencing poverty earlier in life had a more detrimental effect than experiencing it later. Being born into poverty affected cognitive development the most.

Dickerson and Popli (2016) also applied advanced statistical techniques to separate parental investment from family stress. This technique led them to the conclusion that it is parental investment rather than family stress that is the most statistically significant aspect of poverty.

Conceptual analysis

Perspective

Speaking of environmental influences on cognitive processes is somewhat unusual in the context of the triad of perspectives that we normally use: biological, cognitive, sociocultural. A person's environment would normally be thought of as a sociocultural influence. It may be argued that factors like poverty are sociocultural.

However, environmental factors in cognition deserve a separate place in the IB Psychology syllabus because of their practical applications. Nutrition, air quality, and noise are all parameters that are easily measurable and that potentially can be controlled. If something as simple as providing children with a healthy breakfast can achieve the same cognitive outcomes as sending them to a better school, then that is a great, cost-effective investment.

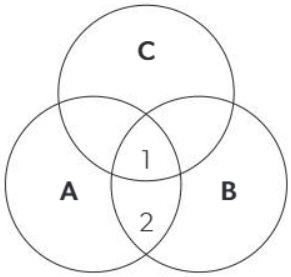
Causality

Unlike with some other factors, such as biological or sociocultural ones, there is more confidence that environmental factors affect cognition and not the other way around. For example, with biological factors you could argue that brain structure affects cognition, but you could also argue that exercising cognitive processes has an effect on brain structure through neuroplasticity.

In contrast, with environmental factors, it makes sense to hypothesize that air pollution may affect one's cognitive processes, but it is hard to imagine how one's cognitive processes may directly influence air pollution.

Activity (HL only)

Statistical separation of effects involves using complex statistical methods to analyse correlations between multiple variables and calculating correlations between A and B after the correlations with C (confounding variables) are statistically "removed". This idea is illustrated in the Venn diagram below: if three variables (A, B, and C) all correlate with each other, it is possible to mathematically "remove" area 1 and estimate the "pure" correlation between A and B—area 2 in the image. The correlation between A and B is now "unaffected" by the confounding variable, C.



▲ Figure 2.29 Venn diagram illustrating the statistical separation of effects

In research papers, such confounding variables are frequently referred to as "covariates". For example, if an author describes an analysis where they investigated the effect of one variable on another variable "with age and gender as covariates", this means that the effects of age and gender were considered and could be statistically removed in the way described above.

Find one study you have already encountered where covariates were used in results analysis. Why do you think researchers included those variables as covariates in their analysis?

ATL Research, Thinking

Measurement

Measuring the effects of environmental variables on cognition is not easy. The problem lies in the isolation of variables.

Environmental factors always come as a package of multiple variables that act together and mediate each other's effects. Even in something seemingly simple, such as having a breakfast, there are multiple variables to consider: some components of your breakfast may enhance or negate the effect of others.

The trade-off between internal validity and ecological validity of an experiment is very relevant here. When we isolate one variable and keep all other variables constant or controlled, it increases the credibility (internal validity) of our experiment, but it also reduces its ecological validity, making it more artificial and less generalizable to real-life situations.

Bias

Any complicated research design runs the risk of biased conclusions based on research that oversimplifies things. Such conclusions as "heat makes your memory worse" or "eating breakfast improves your school performance" or "poverty impairs your cognitive development" are all simplified statements that ignore a whole range of important aspects. While simplified does not necessarily mean biased, it may lead to biases if such statements are systematically misunderstood.

Will an increase in temperature always lead to a decrease in memory performance? Not necessarily—it depends on how other variables are changing. If someone took these research findings and used them to jump to some ill-informed conclusions, such as that firefighters have worse memory, that would be biased, and it would raise ethical concerns.

Responsibility

One of the problems with applying results of research in this area is misrepresentation of effect size. For example, suppose you had two groups of participants performing a simple memory test in either a normal or a high-temperature environment. Suppose the first group remembered 75% of material, but the second group only remembered 73%. However, the groups participating in your study were very large and the result is statistically significant.

Indeed, this finding has theoretical significance. But looking at it from a practical point of view, fluctuations far greater than that may occur for other reasons—for example, how well-rested the individual is. Imagine an article in mass media with a catchy title "High temperatures damage your memory". The journey from research findings to their being reported is long, and it is the ethical responsibility of anyone who writes a scientific report to present information in a way that is balanced and does not misrepresent the actual findings.

Change

There is a difference between short-term effects and long-term effects. When studying the effects of environmental factors on cognitive processes, it is important to investigate both. The study of poverty makes it obvious. It would be a gross oversimplification to only base research studies on a comparison

between poor and non-poor populations. Here are some questions relevant to the idea of change:

- Are the effects of poverty cumulative? Is it true that the longer you stay below the poverty line, the more negatively your cognitive functions are impacted?
- Are the effects of poverty irreversible? Is it possible for cognitive functions to go back to "normal" after you have climbed above the poverty line?
- Is the timing of poverty episodes important? Is it true that the earlier you experience poverty in life, the worse it will affect your cognitive development in the future?

2.9 Potential for improving cognitive processes

Inquiry questions

- Can a cognitive function be restored to a normal state after injury?
- Could we do something to eradicate implicit cognitive biases and stereotypes?
- Is it possible to read someone's thoughts with brain scanning?

What you will learn in this section

Key learning:

- Understanding how human processes work allows us to try to improve them.
- One example of "improving" a cognitive process is restoring a cognitive function to its normal state after an injury. Using one sensory modality to process information that is normally associated with another modality is known as sensory substitution.
- Another example is when cognitive function has not been lost or impaired, but we know that cognitive performance could be better. Debiasing strategies such as using counter-stereotypical information have demonstrated effectiveness in reducing implicit bias and making human judgement more rational.
- Implicit association tests are useful to capture the difference between implicit biases and explicit attitudes.
- A third example is using brain–computer interfaces to acquire new cognitive skills, and the related field of neuroprosthetics. Early research has been conducted on animals and has achieved some promising results. New research with humans (such as turning thoughts into words) relies not only on measuring the activity of the brain, but also on using AI to decode the signals.

Key terms: neuroplasticity, sensory substitution, sensory augmentation, debiasing, counter-stereotypical information, Obama effect, implicit association tests (IAT), implicit prejudice, explicit attitude, brain–computer interface (BCI), neuroprosthetics, large language model, locked-in syndrome

In a wider context

If we understand something, we can explain and predict it. If we can explain and predict it, we can control it. It would be reasonable to believe that understanding human cognitive processes gives us a degree of control over them as well.

When we think of "improving" cognitive processes, many ideas come to mind. However, in this unit, we will consider the following three examples:

1. Restoring a cognitive function to its normal state after it has been impaired (e.g., restoring memory after amnesia or returning limb sensitivity after paralysis).
2. Increasing the efficiency of cognitive functions (e.g., increasing retention of learned material or making one's judgements under uncertainty less biased and more rational).
3. Using brain–computer interfaces to acquire new cognitive skills.

Restoring a cognitive function to its normal state

Cognitive functions may get impaired or disrupted due to an injury. There are plenty of examples of how brain damage affects cognition. In fact, such examples are used as evidence for localization of cognitive functions in the brain. However, thankfully, there are also plenty of examples of how these cognitive functions may be restored or compensated for due to neuroplasticity. **Neuroplasticity** is the ability of the brain to form new connections in response to experience. It is the making and breaking of synaptic connections.

Paul Bach-y-Rita (1934–2006) was an American neuroscientist, known as "the father of **sensory substitution**" (Twiley, 2017). Sensory substitution is when you use one sensory modality to perceive information that would normally be perceived by another sensory modality.

One of Bach-y-Rita's inventions is a chair that allowed blind people to "see" (Bach-y-Rita et al., 1969). It was a dental chair with 400 small, vibrating plates built into the back. The plates were arranged in a 20×20 grid. They could be switched on in a pattern, resulting in an "image". There was also a camera on a tripod that the patient could manipulate, pointing it at various objects around the room, zooming in and out. The image from the camera got "translated" into the pattern of the plates vibrating against the patient's back. Six blind subjects were trained with the device. Most of them had been blind from birth. Most of them underwent 20–40 hours of training.

Participants were able to achieve remarkable results. They were able to recognize common objects such as a telephone, a toy horse, or a cup. They could discriminate overlapping objects when more than one was presented at a time. They could recognize persons standing before the camera and describe their posture and movements, and recognize letters. Remarkably, training was shown to be very effective. For example, during the first trials it normally took the subjects between five and eight minutes to recognize a common object, but after ten hours of training, recognition time fell to 5–20 seconds! Interestingly, participants reported external localization of stimuli—that is, subjectively they felt like the sensory information was coming from in front of the camera rather than from the sensors contacting their back. It's like they learned to use the apparatus as their "eyes". As Bach-y-Rita put it, "You don't see with the eyes. You see with the brain" (Bach-y-Rita et al., 1969).

Exam tip

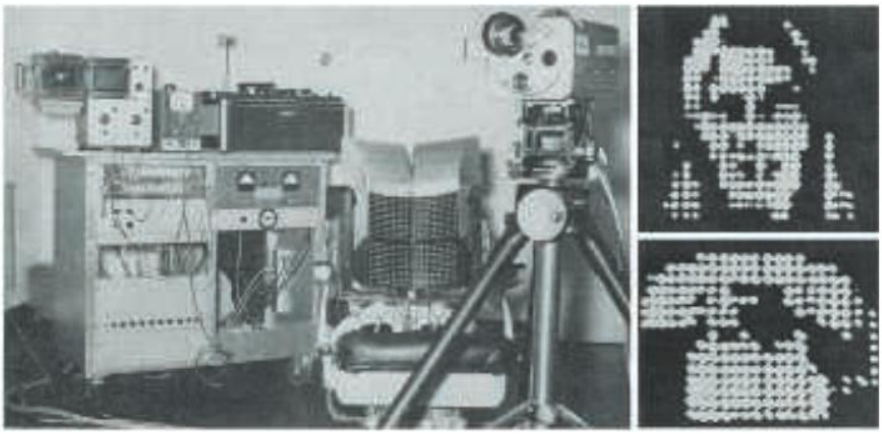
Neuroplasticity is a content point that can appear in questions in Papers 1A and 1B. You will need to be able to explain it and support your explanation with one example.

The necessary content to answer exam questions on neuroplasticity is summarized in Unit 3.1: Brain development.

TOK

It is commonly believed that "seeing something with your own eyes" is the most reliable evidence.

Now that you know the story of Bach-y-Rita's invention, what would you say to someone holding that belief?



▲ Figure 2.30 The "tactile television" of Paul Bach-y-Rita

Discussion

If there was an advertisement recruiting volunteers to sign up for a trial of a sensory augmentation device, would you sign up for it? Why or why not?

What "superpower" would you like to get with the use of sensory augmentation technology?

ATL Communication, Thinking

Technology has developed since 1969. Bach-y-Rita's initial invention has been transformed into a device that connects a small camera on a patient's brow to a structure—similar to a plastic plate the size of a postage stamp—that the patient holds in their mouth against their tongue. This way the patient can "see" with the tongue. Erik Weihenmayer, who was blind since high school, is famous for having climbed the tallest peaks on all seven continents, and for being the first blind person to climb Mount Everest. He used the device—but not for Everest—to help him climb other challenging summits (Twiley, 2017).

David Eagleman, another neuroscientist who works in the area of sensory substitution, believes that the adult brain can even be trained in the future to develop super-senses, such as the ability to "feel" electromagnetic fields, stock market data, or the weather in space (Twiley, 2017). This is referred to as **sensory augmentation**. Imagine being able to feed data from a GPS tracker to a device connected to your skin so that you could "feel" the geographical location of your family members in real time.

Increasing the efficiency of cognitive functions

In this section, we consider the situation where the cognitive function has not been lost or impaired, but could be better.

One example of this is **debiasing**. Debiasing is an attempt to reduce the impact of implicit cognitive biases and restore rationality of judgement, especially in crucial situations. See also the discussion of debiasing in Unit 2.5: Cognitive biases and the dual processing model of thinking and decision-making.

What makes this issue difficult is the implicit nature of the bias. "Implicit" means outside of conscious awareness and voluntary control. In fact, implicit bias is often dissimilar and even opposite to **explicit attitude**, so a lot of people who claim (and genuinely believe) that their attitudes are unbiased are actually wrong. Effects are only observable through behaviour, and in many instances those who have an implicit bias will find the results surprising and shocking.

The good news is that some effective techniques of reducing bias in judgement have indeed been discovered. For example, it is known that **counter-stereotypical information** can be effective at reducing negative stereotypes. Counter-stereotypical information is any information we encounter that contradicts our existing stereotypes.

Columb and Plant (2011) discovered what they called the **Obama effect**. According to this effect, positive examples from a social group can decrease biased judgement and **implicit prejudice** towards that group. For their experiment, they recruited 51 non-Black undergraduate psychology students. The study included two stages.



▲ Figure 2.31 Barack Obama

Stage 1: participants performed a task where they had to decide if a string of letters was a word or non-word (e.g., compare "table" and "belat"). Before each trial, participants were primed with a name—this means that the name was flashed on the screen for 55 ms. The name could be "negative", for example a person who was perceived negatively by the majority of participants at the time of the study (e.g., well-known criminals), or "positive", i.e. Barack Obama (who was the first Black person to be elected president of the USA). There were three randomly formed groups depending on the exact nature of this priming.

- Group 1: primed with a "negative" name on the first set of trials and the neutral "XXXXXX" on the second set.
- Group 2: primed with the "negative" name on the first set of trials and "Obama" on the second set.
- Group 3 (control): primed with XXXXXX on both sets of trials.

Stage 2: participants were given a Black/White **implicit association test** (IAT). An implicit association test is a procedure based on reaction times where participants have to quickly assign stimuli to groups of "positive" or "negative". This measured their implicit anti-Black prejudice. They were also given a measure of explicit attitudes (a questionnaire).

Results of the study showed that participants in the Obama condition displayed significantly less implicit Black/White bias on the implicit association test than the participants in the negative condition. There was no difference between the control condition and the Obama condition. Conscious, explicit attitudes (as measured by the questionnaire) were not affected by the experimental manipulation.

It was concluded that exposure to Obama "undid" the harmful effect of negative exemplars on implicit racial bias. In other words, positive counter-stereotypic exemplars (such as Barack Obama) can reduce implicit anti-Black bias in judgement.

Discussion

How could you put the Obama effect into practice in your school to address problems with stereotypes, prejudice, and discrimination?

ATL Thinking, Self-management

Activity: implicit association tests (IAT)

In a Black/White IAT, you are given a stimulus, which you need to categorize with a click of a button into one of two groups, as quickly as possible. The stimulus is either a face (Black or White) or a word (negative or positive).

In some trials, the two groups are presented like this:

a. Black or negative
b. White or positive.

In other trials, the groups are presented like this:

a. Black or positive
b. White or negative.

The idea here is that, if you have an implicit association between "Black" and "negative" in your mind, then the first trial will be cognitively easier for you than the second trial, therefore you will cope with it faster and make fewer mistakes. On the contrary, if there is no implicit prejudice, then completion times on both trials should be similar, as should be the number of errors.

Similar tests exist for other typical prejudices, for example against women in science, older people, and people with disabilities. Many of these tests are available for you to try for free on Harvard's "Project Implicit" website.

Take a few of their IATs. Share your reflections (but not your results) with a partner: Did you find the results surprising? Do you think they are accurate?

ATL Research, Social, Self-management

Brain–computer interfaces (BCI)

Today's technology makes it possible to establish a direct connection between the brain's electrical activity and an external device through a **brain–computer interface** (BCI). This means literally controlling the external device with your brain.

A field of research related to BCI is **neuroprosthetics**. A neuroprosthetic device aims to supplement or replace an impaired sensory organ, limb, or cognitive function. Think about such examples as a robotic arm, an artificial eye, or a "memory chip" implanted in your brain to enhance the capacity of your memory.

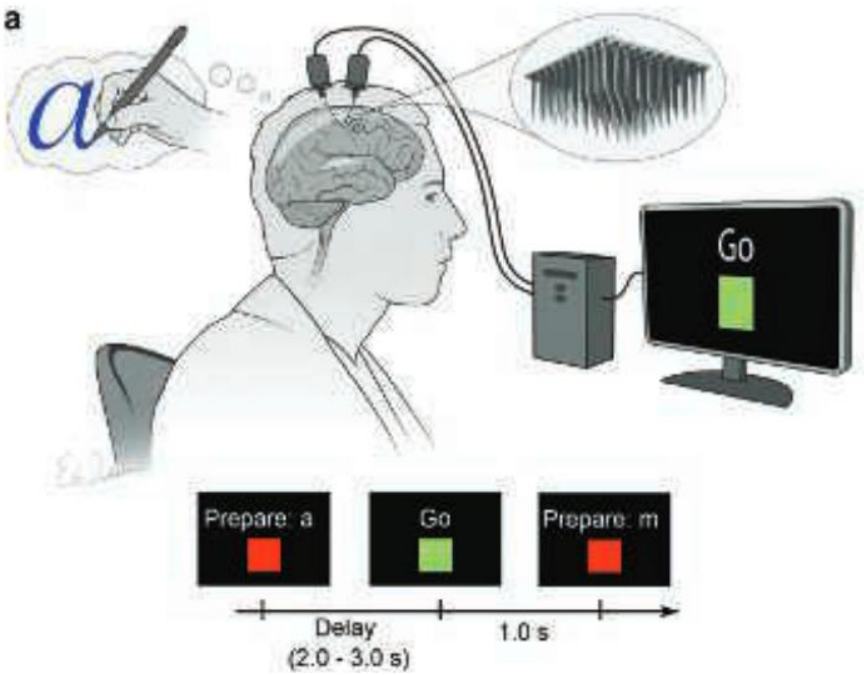
Early BCI research was conducted on animals such as monkeys and sheep, and even with the recent inventions the technology is first extensively trialled on animals before any human trial can be allowed. For example, Elon Musk's Neuralink was implanted in a monkey in 2021.

Stanley, Li, and Dan (1999) installed electrodes in the thalamus of cats and managed to decode visual signals from the cat's retina. The cats were shown eight short movies and the researchers were capable then of reconstructing recognizable images on a screen. Velliste et al. (2008) succeeded at training a monkey to pick up pieces of fruit and marshmallows and to feed itself using a brain-controlled robotic arm.

Recent promising results demonstrate the potential of BCI to decode speech signals and translate thoughts into spoken words, which may eventually be used to help people who have lost the ability to produce articulate speech.

Willett et al. (2021) had one participant who could not speak and had been completely paralysed and unable to move from the neck down. The participant was instructed to imagine that his hand was not paralysed, that he had a pen in his hand, and that he was writing on a piece of paper.

Electrodes were installed in his prefrontal cortex (precentral gyrus). Signals from the electrodes were decoded using a range of data analysis techniques, including machine learning, which is a subset of artificial intelligence (AI).



▲ Figure 2.32 Willett et al. (2021) setup

Researchers achieved a high accuracy of "translating" these signals into handwritten letters. The patient's "typing" speed reached 90 characters per minute with a 99% accuracy rate (if using autocorrect), which is comparable to the smartphone typing speeds of healthy subjects of the same age (115 characters per minute).

Researchers were able to decode complete handwritten sentences in real time. The participant was allowed to use some special symbols (such as the apostrophe) and punctuation, but for easier detection had to write spaces as ">" and periods as "~". The results are shown in Figure 2.33 (the top sentence is what the patient imagined writing, the sentence at the bottom is what the machine decoded. Mistakes are highlighted in red).

Prompt: you must be the change you wish to see in the world
Raw Output: y ou> m u s t >be >t he >c hang e >y ou> wi s h>t o>s e e> i n>t he >w orl d



Prompt: daisy leaned forward, at once horrified and fascinated.
Raw Output: d a i s y >l ea r e d >f or wa rd , >a t >once >ho o r r i fi e d > a r d >f a s c i na t e d .



▲ Figure 2.33 Willett et al. (2021)—real-time examples

The study demonstrated that the motor intention for handwriting remains intact in the cortex even after paralysis. It means that, although your hand is no longer able to actually carry out the writing, the neural signals with the writing intention still get generated in the motor cortex, even years after the accident.

The reason Willett et al. (2021) achieved such impressive results in their study is because they used **large language models** (LLMs) to analyse the preceding sequence of signals and to estimate the probability of which letter or word will appear next. Consider the sentence "you must be the change you wish to see in the world". When it is time to decode the neural activity underlying the production of the word "see", it is not just the messy signals that we have at our disposal. We can use the preceding sentence ("you must be the change you wish to…") and estimate the probability that each word of the English language will appear next in this sequence (it is similar to autocomplete in a text editor). We can then use these probabilities to make better sense of the messy signal.

The use of such techniques is difficult and requires a lot of computation. That is why the research is currently small (only one participant in Willett et al.'s study) and serves as a proof-of-concept. However, even with just one patient, such achievements inspire hope that, in the near future, we will be able to help patients with severe impairments, such as the "locked-in syndrome".

Activity

Find out more about the "**locked-in syndrome**": what it is and what techniques have been invented to help such patients. Can you think of more ways that we might find to work with the "locked-in syndrome"? Do you think the development of technology could make a difference in this area?

ATL Thinking, Research

Discussion

The first case of direct "telepathic" communication between nervous systems of two individuals was recorded in 2002, when Kevin Warwick (a researcher in cybernetics and artificial intelligence) had a 100-electrode array implanted in his left wrist and connected to the internet. Warwick was able to control a robotic hand remotely across the ocean and receive signals (such as the sensations of force and pressure) back from it. Later on, his wife Irena also joined the experiment and had similar electrodes installed in her arm. They "connected their nervous systems together" over the internet. Every time she would open or close her hand, he would receive the pulses, which they described as "hand-holding" (Giaimo, 2015). Of course, this is nothing like what we imagine when we say "telepathy"—i.e., directly transmitting thoughts to each other over long distances. Nevertheless, it was the first of its kind.

If you were in a long-distance relationship, would you be willing to buy such a device and "install" it in your body?

ATL Self-management

Conceptual analysis

Perspective

Some improvement strategies considered in this unit are based on the idea of neuroplasticity (training the brain to perform a function by re-wiring it) and localization of function (learning to decode brain activity by registering where signals are coming from). These approaches try to leverage biological factors in cognition. We have considered some purely cognitive approaches as well—for example, using cognitive interventions to overcome cognitive biases. Of course, there are also approaches to improving cognitive functions based on sociocultural factors: in fact, education is a sociocultural way of improving cognition.

Causality

In improving cognitive processes, the idea of perspective and the idea of causality closely overlap: perspectives are guided by what we believe is the major factor that influences cognitive processes.

There is also another dimension of causality that is highly relevant to this area of research: in order to claim that X improves a cognitive process, we must make sure in our research that X is carefully isolated and that it is indeed X, and not something else, that necessitates the change in the dependent variable. This closely links to internal validity of experimental research. Think about how dangerous it might be to jump to incorrect conclusions based on incomplete or biased research.

Bias

Suppose it has been discovered in research that early training of attention and cognitive control in pre-school years improves memory and literacy scores in primary school. This finding alone is not reason enough for us to start modifying all our pre-school curricula to fill them up with attention training.

There are all sorts of questions around it that need to be answered first. How large is the improvement? Are we talking "making a real difference" or "statistically significant, but practically negligible"? What is the cost of the intervention—can teachers be easily trained to deliver this new curriculum? Can a similar result be achieved with less effort using some other strategy? It is inevitable that we will fill some of the gaps with our assumptions and jump to some conclusions based on prior theory or common sense. This is where bias may easily creep in, especially when there is commercial or other interest involved in the intervention that is being developed.

Measurement

To improve a cognitive function, we need to have a solid theory that explains it. Theories are based on research which is limited by multiple considerations related to measurement. All the problems of measurement in research related to neuroplasticity, localization of function, implicit biases, brain imaging, and other areas are also relevant here. For example, attempts to create BCIs are fundamentally limited by spatial and temporal resolution of brain imaging technology.

Responsibility

Researchers should also make every attempt possible to investigate various aspects of the proposed intervention: comparing its short-term effects to long-term effects, and looking at possible side effects, costs of implementation, and unintended consequences. It is never possible to achieve this level of investigation in one research study, so no practical application should ever be based on a single study.

Practitioners constantly look through scientific publications to find inspiration for their proposed programmes (e.g., school curricula, rehabilitation techniques, therapeutic approaches). They should always conduct a thorough search and make balanced decisions while recognizing the inherent uncertainty of available

evidence. When evaluating evidence, practitioners have the responsibility of conducting a comprehensive analysis of the credibility of research and generalizability of its results to their specific practical context.

Change

Improving cognitive processes is directly related to the concept of "change". Armed with our knowledge of a psychological phenomenon, we purposefully change it.

One important distinction to remember in this regard is that between change and development. This distinction is particularly important when we are looking at a growing individual. As we grow up, our cognitive functions undergo a natural process of maturation. To the extent that they do, they will develop even without any external influences such as education. When we evaluate research, we should always rule out maturation as an alternative explanation of the observed changes.

Exam-style practice questions

Paper 1 Section A (4 marks)

Using one example, explain classical conditioning.

Paper 1 Section B (6 marks)

Anomic aphasia is the neurological speech impairment where a person finds it difficult to retrieve the correct words and express what they want to say (especially nouns and verbs). They are unable to come up with the right words for what they want to talk about. When they speak, it may seem like they are talking "around" the subject matter without naming it directly. The same happens when they write. All other abilities remain intact.

Research has indicated that anomic aphasia is associated with brain damage in the white matter under the surface of the left hemisphere of the brain; more specifically, the nerve tract called the arcuate fasciculus. This nerve tract provides a connection between the back of the brain and the front, but its exact functions have not been fully understood yet.

Explain this finding with reference to the idea of localization of function in the brain.

Paper 1 Section C (15 marks)

Hypothetically, knowing how cognitive processes work should enable us to change them in desirable directions. To what extent can one cognitive process be improved?

(Concept: Change. Content: Potential for improving a cognitive process. Context: Learning and cognition)



3 Human development

Introduction

Developmental psychology is the study of how and why humans change across the span of their life. The first major focus of developmental psychology was on early years (infancy and childhood), but gradually the field was expanded to other periods such as adulthood and senior age.

The changes are driven by two interacting processes: maturation and learning. Maturation refers to the gradual unfolding of the genetic programme. We develop certain abilities at a certain age because that is how our genes work. Learning refers to changes in response to environmental stimuli. Maturation and learning interact. On the one hand, maturation creates a "predisposition" to learning, which manifests itself in the form of "critical periods". For example, there is a critical period for the development of language abilities, and if language is not developed at that time, there is a high chance that the brain will never be able to learn language at all (e.g., this is evidenced by the case studies of feral children). On the other hand, learning is often a necessary condition of development. Learning is essential to bring the genetic plan to life. The relative contributions of maturation and learning to human development have been the focus of the "nature versus nurture" debate.

As you know, there are three important dimensions of human behaviour: biological, cognitive, and sociocultural. Human development is studied through the lens of these three broad aspects. They are not isolated. For example, the genetically programmed development of certain areas of the brain at a certain age makes it possible for a person to develop certain cognitive abilities. Cognitive abilities will influence the way we interpret social interactions and so lead to a formation of identity—a representation of one's self in a larger social group. Identity will determine a lot about our interactions with people and in turn these will influence both our cognitive and brain development. This complex process of interaction between the various dimensions of human development is the focus of developmental psychology.

It is essential to understand the major factors that influence cognitive and social development. This research has a lot of practical implications. For example, what is the role of trauma in a child's cognitive and social development? This will influence how we approach treatment and education. What is the role of poverty and socio-economic status? Knowing the answer to this question is important for various enrichment programmes.

From the standpoint of theory of knowledge, it is logical that our understanding of something might be enhanced by our knowledge of how it has developed. For example, the nature–nurture debate can benefit a lot from the knowledge of how maturation and learning interact in the growing child. This has not always been understood. The first philosopher who emphasized the idea of change and its importance was Heraclitus in Ancient Greece, who said, "Everything flows" and "No man ever steps in the same river twice". For a long time after that, however, the idea of change was not given too much attention and much of the discussion focused on things and phenomena as they are (in their current state). The ideas of change and development were later reintroduced by a German philosopher Georg Hegel (1770–1831). Think about this: the genius of Charles Darwin largely depended on his idea to trace the existing species back to their ancestors, and to look at the development of species rather than their current state.

Before Darwin, there were numerous attempts to classify and explain existing species (for example, by the French naturalist, Lamarck). These were largely unsuccessful because they ignored evolution. Darwin did his research shortly after the publication of Hegel's influential philosophical ideas.

Similarly, it seems natural today to look at childhood and adolescence as special periods in human development, but this was not always so. The concept of childhood only emerged in the 17th century when the new advances of society required prolonged education and protection. Previously, children had mostly been viewed as little versions of adults. Once they became physically capable, they were used as apprentices in their parents' workshops or engaged in other manual labour activities. It was expected they would have similar behaviour, similar experiences, and similar responsibilities to adults. The idea that people go through certain "stages" in their psychological development is a relatively recent one.

Why and how does a person become what he or she becomes? What are the contributions of genetics and environmental influences? What is the role of culture? How do people develop their values, attitudes, relationships, and abilities?

Children are not "small adults". They are qualitatively different. They do not perceive the world in the same way as adults do. Also, a child cannot understand adult reasoning, so it is the responsibility of adults to understand how children's minds work.

Child-rearing practices have also changed dramatically in history. In Medieval Europe, for example, children were often restricted in their cribs so that they could stay motionless and "humble". It is probably too early for you to have children, but imagine you do. If you have experience in taking care of younger siblings, it will be easier for you to imagine what it is like. If you were responsible for these small people's lives, would you:

- have talked to them while they were still in the womb
- raise them in a bilingual environment assuming that this will boost their cognitive development
- impose a strict set of rules that they have to follow
- punish them when they break the rules
- shout at them if they have done something potentially dangerous for their health?

Here is an overview of the topics in this chapter as presented in the IB DP Psychology Guide:

Models of development	Development of self
• Brain development • Sociocultural factors in development • Stage theories and continuous models • Theory of mind	• Attachment • Enculturation of social norms • Peer influence • Role of childhood experiences

The IB Guide states that for models of development, students should understand one or more of the following: cognitive, social, moral, or language development.

Note: The order of topics in this chapter will be changed to ensure a logical sequence of learning.

3.1 Brain development

Inquiry questions

- How does the brain change as it grows?
- To what extent does biological development of the brain depend on life experiences?
- How good is the brain's ability to reshape itself?

What you will learn in this section

Key learning:

- Developmental neuroscience has accumulated rich knowledge about the biological process of brain maturation. This process can be viewed in terms of four stages: neurogenesis, migration of neurons, differentiation, pruning.
- Neuroplasticity is the making and breaking of synaptic connections. It is the neural mechanism of learning.
- Experience-expectant neuroplasticity works by overproducing neurons and connections and then eliminating the ones that are not used (pruning). There are critical periods of time when the brain is particularly responsive to relevant experiences.
- It is not the number of connections that determines the level of cognitive functioning, but how efficiently they are organized.
- Experience-dependent neuroplasticity has evolved to deal with experiences that are unique to the given individual. New synapses between neurons are created depending on how frequently the two neurons are activated together.
- The study of structure–function relationships in the developing brain is important, but the interaction between structure and function is complex. Developmental neuroscience should rely heavily on triangulation of evidence.
- There are important limitations associated with measurement. For example, young participants are not suitable for long brain scans, and the contrast between grey matter and white matter also changes, affecting brain imaging.

Key terms: biological maturation, neurogenesis, migration of neurons, differentiation, pruning, neuroplasticity, synaptic plasticity, cortical remapping, experience-expectant neuroplasticity, experience-dependent neuroplasticity, sensitive (critical) periods in development, discrimination paradigm, structure–function relationship, localization of function, triangulation of evidence, developmental neuroscience

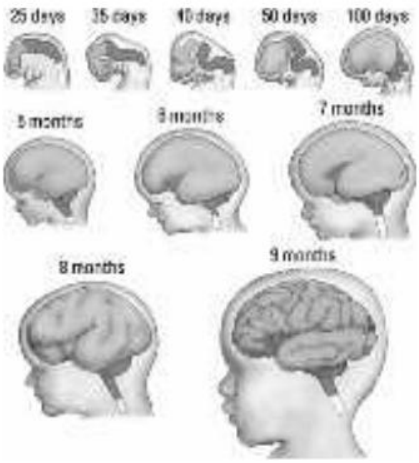
In a wider context

Brain development is the biological process of the growth of neurons, the formation of new synaptic connections, and so on. This is different from cognitive development, which is the development of cognitive functions such as memory, thinking, and decision-making. Cognitive development is going to be the focus of the following sections in this unit.

Biological process of brain development

The brain is a complex system of highly interrelated elements. The successful development of this system requires that each component be formed fully in a timely way and integrated correctly with other components. Broadly speaking, this process can be viewed in terms of four stages:

1. neurogenesis (the birth of neurons)
2. the migration of neurons to their correct location



▲ Figure 3.1 Prenatal development of the human brain

Exam tip

You will not be asked any questions specifically about the structure of the brain, so there is no expectation that you will have a thorough knowledge of brain anatomy. Having said that, various examples you use to support your responses will entail knowledge of some specific parts of the brain and their functions.

For an example you know you will want to use in exams, it will help to look up any brain parts that you do not understand. There are apps available showing 3D models of the human brain. They let you see where the various parts are located, along with their functions. Carry out some research on the internet to find an app that suits you.

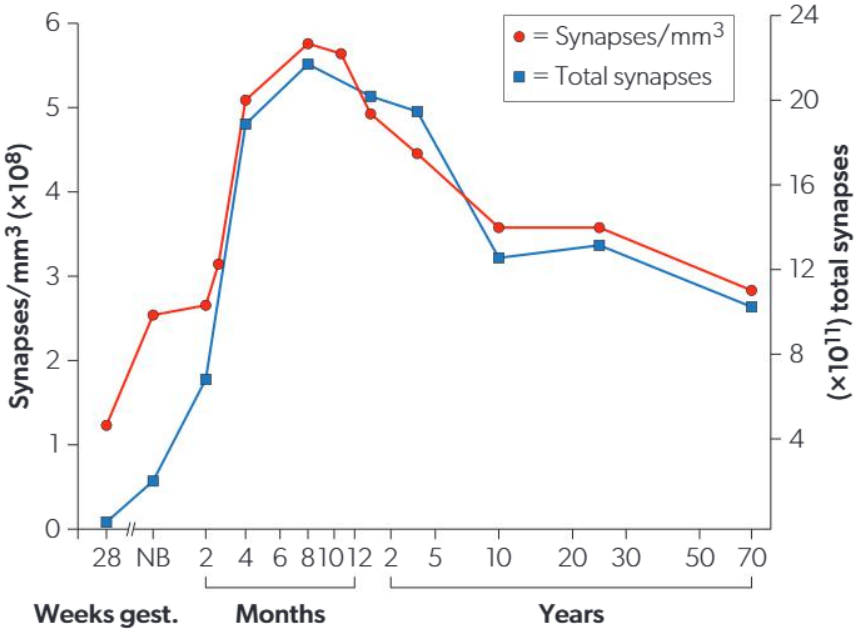
3. differentiation (the development of connections between neurons)
4. pruning (the elimination of these connections as well as neurons themselves).

Neurogenesis is the production of new nerve cells. It is mostly finished during the gestation period—that is, the neurons in your brain are all produced by the time you leave your mother's womb. Curiously, the cortex actually overproduces neurons to account for normal cell death in the future. This means you have fewer neurons now than when you were just born.

Migration occurs approximately nine weeks after conception when neurons start to migrate to their "correct" positions. They travel along special glial fibres that form very early in the foetal brain and extend from the brain's inward structures to the cortical layers. The neuron "climbs" along the fibre like a snake on a tree trunk.

It should be noted that neuronal growth in the brain cortex happens layer after layer, "inside out". That is, new neurons are formed below the previously formed layers and then travel outwards, passing the existing layers of neurons on their way. New cortical structures are added to the old ones in a complementary manner, literally "on top" of them. It would be curious to see if the development of cognitive abilities follows the same pattern (Jean Piaget would say it does. See Unit 3.2: Stage theories and continuous models of human development).

Differentiation is the growth of neural connections (synapses), and **pruning** is the elimination of these connections. Synapses start growing shortly after conception. After some critical point this growth becomes rapid (40,000 synapses per second) and continues until humans are nearly two years of age. This is followed by a plateau and then rapid reduction in the number of synapses. Elimination of synapses continues at a rate of up to 100,000 synapses per second and lasts until the end of puberty. Around 50% of synapses that are initially



▲ Figure 3.2 Synaptic density (number of synapses per mm³) and total number of synapses in the primary visual cortex as a function of age

formed are eliminated during this period. This is followed by another plateau, and then there is another drop in the number of synapses in old age.

According to Kolb and Fantie (2009), "the mechanism that controls synapse formation is one of the major mysteries of developmental neurobiology, largely because synapses are perceptible only by electron microscopy, which does not allow direct observation of their sequence of development in living tissue" (Kolb and Fantie, 2009, p. 25).

One of the puzzling observations about this process is that synapses between neurons form very rapidly and abruptly, although the two neurons may be positioned next to each other for days.

Pruning and death of neurons are perhaps the most intriguing findings of **developmental neuroscience**. Both are linked to environmental stimulation. A lack of such stimulation at a certain critical period may lead to a loss of neurons and/or connections between them (the "use it or lose it" principle). Synaptic density in the brain of an infant is larger than that in the adult brain. So, if you are using the logic "more connections = higher intelligence", then infants should be a lot smarter than adults.

SAQ Neuroplasticity

Neuroplasticity is the making and breaking of synaptic connections between neurons. During this process, neural networks in the brain literally change their shape. The reasons for such changes are both genetic (normal pre-programmed development of the brain) and environmental (e.g., injury, brain damage, or simply learning a new skill).

Neuroplasticity can be observed on different scales. On the smallest scale, at the level of a single neuron, it takes the form of **synaptic plasticity**. This is the ability of the neuron to form new synaptic connections and break old ones. On the largest scale, neuroplasticity takes the form of **cortical remapping**. This is the phenomenon when one brain area assumes the functions of another brain area—for example, due to injury.

Synaptic plasticity depends on the activity of neurons. If two nearby neurons are frequently activated at the same time, a synaptic connection between them may form. Similarly, if two neurons are rarely activated together, the existing connection may fall apart. This has been summarized like this: "neurons that fire together, wire together" (which was originally said by Carla Shatz and is quoted in Doidge, 2007) and "neurons that fire out of sync, fail to link" (Doidge, 2007, pp. 63–64).

As you will learn later in this section, in the context of brain development it is important to make a distinction between **experience-expectant neuroplasticity** and **experience-dependent neuroplasticity**. Experience-expectant neuroplasticity refers to the mechanism when the brain "anticipates" certain typical experiences and becomes particularly responsive to them at a predetermined period of time. This links closely to the idea of "critical periods" in human development. Experience-dependent neuroplasticity refers to experiences that are unique to each individual, where patterns of neuroplasticity depend entirely on the activity itself.

Pruning is also an example of neuroplasticity.

Chat with AI

If you are curious to know more about the process of **biological maturation** of the brain, use your favourite generative AI. This is a highly specialized, but well documented, field of knowledge where AI can be particularly helpful. Research papers may be too technical for you, but generative AI has a great ability to rephrase and simplify things.

You could consider the following prompt fragments and conversation starters:

- I am an introductory psychology student and have been learning about brain development (maturation). Could you tell me a bit about the process of neurogenesis—what key things do I need to know at an introductory level?
- [Ask all sorts of clarification questions, for example:] Does this mean that no new neurons are produced after birth? Is it true that neurons have a limited lifecycle and die after a certain number of years?
- Could you give me some examples of how pruning can explain cognitive growth?
- How can elimination of synaptic connections (pruning) result in better cognitive function? I don't understand this. Could you explain?

Discussion

Pruning is counterintuitive—why do cognitive processes improve when neurons are eliminated?

Discuss with a partner and share your hypotheses.

Would your commute from home to school be more efficient if there were more available routes?

ATL Communication, Social, Thinking



▲ Figure 3.3 Hindi alphabet (Devanagari script)

Experience-expectant neuroplasticity, pruning, and critical periods

Greenough, Black, and Wallace (1987) suggested that there are two types of neuroplasticity in the developing brain: experience-expectant and experience-dependent.

Experience-expectant neuroplasticity is related to the well-known idea of **sensitive (critical) periods in development**. Critical periods are brief windows of opportunity when experience can influence development. For example, there is only a relatively brief window of opportunity for language abilities to be developed in early childhood. If this opportunity is missed and the child is deprived of language and communication, it will be impossible to acquire a language ability at a later time.

Greenough, Black, and Wallace (1987) suggest that certain types of experience have been common in the evolutionary history of a species, and therefore the brain has evolved to "anticipate" them. The way this is achieved is by over-producing synaptic connections and then pruning the unnecessary ones. Specific experiences with the environment determine which synaptic connections will survive. Since these typical experiences are "expected" by the brain, they are incorporated into the unfolding genetic programme. This opens the window of opportunity exactly when the brain is ready to benefit the most from such experiences.

An example that illustrates over-production and subsequent pruning is the study of linguistic abilities of infants by Werker et al. (1981).

Werker et al. (1981) aimed to investigate the ability to discriminate between Hindi phonemes in infants (six to seven months of age), English-speaking adults, and Hindi-speaking adults. A phoneme is a sound of speech. The researchers used pairs of stimuli, such as /ta/ versus /Ta/ in Hindi. This distinction between pronounciations is not used in English, but it is meaningful in Hindi because replacing one sound with the other can sometimes change the meaning of the word.

Participants were tested in a **discrimination paradigm**:

- For infants, this procedure involves first conditioning an infant to turn their head towards the loudspeaker when there is a change in the auditory stimulus. This is reinforced for the infant with the presentation of an interesting toy animal each time they turn correctly.
- Adults press a button when they think they detect a change in stimulus.

Results of the study showed that infants (independent of nationality) were just as able to discriminate between Hindi phonemes as Hindi-speaking adults. At the same time, English-speaking adults were not able to discriminate between Hindi phonemes.

Although there is no direct measurement of synaptic connections in this study, results may be explained with the concept of pruning. In fact, the infant brain has an excess amount of neuronal connections that allows it to be responsive to speech sounds. Over time, the connections that are not involved in the infant's actual linguistic experience gradually disappear.

Activity

The idea of critical periods in brain development is also illustrated by case studies of children who were deprived of certain experiences when they were growing up. A well-known one is the case study of a girl named Genie who was severely abused, neglected, and isolated. Her father tied her to a seat and kept her locked alone in her room until the age of 13 years and 7 months. She was not exposed to any interpersonal interaction and so she did not acquire language in the critical period (Curtiss et al., 1974).

Find out more about this well-known case study. What happened after Genie was discovered by social workers? Did she ever learn how to use language? What happened after she turned 18 and returned to live with her mother?

ATL Research, Self-management

Experience-expectant learning highlights the dynamic nature of interaction between biological and environmental factors in the development of the brain. Even the genetic programme itself, due to how it has evolved, is not independent of environmental influences.

Experience-dependent neuroplasticity

By contrast, experience-dependent neuroplasticity has evolved to deal with information that is unique to a given individual. This is also quite important because in addition to typical circumstances and experiences, every organism during its lifetime deals with circumstances that are unique and unforeseen. According to Greenough, Black, and Wallace (1987), the mechanism of experience-dependent neuroplasticity is somewhat different to experience-expectant learning. There is no overproduction followed by pruning. Instead, new synapses between neurons are created depending on how frequently two neurons are activated together.

Blakemore and Cooper (1970) conducted an experiment with newborn kittens. Immediately after birth they were placed into a completely dark room. At the age of two weeks old, they were randomly allocated into one of two conditions: either a vertical or a horizontal environment. The environment was a tall cylinder, the inner walls of which were painted with black-and-white stripes (vertical or horizontal depending on the condition). The kittens also had to wear a wide black collar that did not allow them to see their own body. All they could ever see were the stripes. In this world there were no edges or corners. When they were not in the cylinder, the kittens were brought back to the dark room. This experiment lasted five months. Researchers made an interesting remark about the ethics of the study: "The kittens did not seem upset by the monotony of their surroundings and they sat for long periods inspecting the walls of the tube" (Blakemore and Cooper, 1970, p. 477).

Once the experimental manipulation ended, researchers observed the behaviour of the kittens under normal well-lit conditions, in a room furnished with tables and chairs.

All cats showed some behavioural impairments. They followed moving objects with "very clumsy, jerky head movements" (Blakemore and Cooper, 1970, p. 477).



▲ Figure 3.4 Cat in a cylinder with vertical stripes

TOK

Development is not always a smooth, linear process. Think about various patterns of development that may be observed in different areas of knowledge. How are these patterns different? Here are some examples.

- Evolution of species: combinations of genes occur randomly. However, as certain genotypes die out because they are less adapted to environmental demands, more successful genotypes are gradually consolidated.
- Revolution in the development of political systems: governments are overthrown and replaced, which sometimes causes dramatic changes in society.
- Paradigm shifts in sciences: new evidence triggers the process of rethinking the very basis of a field of science.

Which of these examples do you think is most similar to the development of a human child?

They tried to touch things that were far away out of their reach, on the other side of the room. As they ran around the room, they often bumped into table legs. Moreover, the cats were practically blind to shapes and contours that they had not experienced. For example, in one of the trials researchers used two cats at the same time—one from each condition. If the researcher took a long rod and shook it vertically, one of the cats (the one raised in the vertical condition) ran towards it, followed it, and played with it, while the other cat remained indifferent. If the researcher turned the rod horizontally and shook it, the cats swapped roles—now the second cat was chasing the rod and playing with it while the first one ignored it entirely.

One kitten from each group was anaesthetized and paralysed, then stripes of various spatial orientation were presented to their visual field and the activity of neurons in their visual cortex was recorded. In the cat that was raised in the horizontal environment, visual neurons fired when they were exposed to horizontal stripes and remained "silent" for the vertical stripes. The opposite was true for the other cat. Researchers concluded that the nature of visual experiences in these cats' early childhood modified the physiology of their visual cortex.

Structure–function relationship

Brain development and cognitive development have both been extensively studied. A big issue in developmental psychology is that these studies occurred in parallel but relatively independently from one another. There is a lack of research that considers brain development (structure) and cognitive development (function) in interaction.

The study of **structure–function relationships** is complicated by the fact that **localization of function** is not absolute. Some functions are distributed over large areas of the brain. For example, there is no single brain area responsible for memory or logical thinking. Moreover, localization of function in the brain of a child is not necessarily the same as that in an adult brain.

All this means that developmental neuroscience should heavily rely on the **triangulation of evidence** coming from multiple sources. An example of a research study that used this approach is Kolb and Fantie (1989). In their study, they focused on one specific cognitive ability called "categorization based on linguistic features". One popular measure of this ability is the Chicago Word Fluency Test. In this test, participants are required to write as many words as they can think of beginning with the letter "S" in five minutes. After this, the task changes and requires them to write as many four-letter words beginning with the letter "C" in another four minutes. This study tests a very specific cognitive ability that requires processing language information.

Kolb and Fantie (1989) summarized the following observations of participants' performance on the Word Fluency Test:

1. Frontal lobe regions are active when healthy adult subjects are performing the test.
2. Adult patients with confined frontal lobe lesions do very poorly on the test.
3. At the same time, the same patients perform normally on the modification of the test that requires categorization based on non-linguistic features (e.g., when asked to write the names of as many objects or animals as they can think of).

4. Children perform poorly on the Chicago Word Fluency Test when very young and gradually improve performance as they age.
5. In contrast, even very young children perform well on an adult level on the modification of the test (categorization based on non-linguistic features).

Observations 1, 2, and 3 seem to suggest that categorization based on linguistic features is localized in a specific region of the frontal lobes. Observation 4 suggests that this area gradually develops with age (note: there was no direct measurement of the activity of frontal lobes in children, so this part is an inference). Observation 5 suggests that different areas of the brain (and different cognitive abilities) mature at a different rate. The explanation for this seems to be that frontal lobes in the brain are the last to mature. Other brain areas develop faster.

As you can see, we are using data triangulation here to arrive at inferences about structure–function relationships in a developing brain.

Problems and challenges in developmental neuroscience

The invention of MRI and fMRI advanced the study of the developing brain because it became possible to study the living brain ("in vivo"). This bears the promise of bridging the gap between the study of brain development and the study of cognitive development. Obviously, the two processes should be studied together (because, in fact, it is one and the same process).

Brain imaging techniques such as MRI and fMRI are sensitive to various kinds of movement. Participants have to remain perfectly still in the scanner apparatus while their brain activity is being measured (as long as 40 minutes in some cases). Otherwise the data may get corrupted. This presents a problem with young children. Even more problematically, the comparison between older and younger children is affected. Researchers often use simple designs to minimize the time that children have to remain in the scanner, but this also means fewer variables are measured.



▲ Figure 3.5 Child in an MRI scanner

TOK

An inference is a theoretical conclusion derived from evidence. There is always a gap between evidence and inference, in that the inference is not directly given in the evidence. This gap is closed by logical reasoning. We analyse the available evidence and apply logical reasoning to it, and that is how we arrive at an inference.

Some inferences are stronger, some weaker, depending on the quality of evidence that they are based on. They also depend on the quality of logical reasoning that is applied. It is an important critical thinking ability to be able to analyse the strength of an inference.

Apart from this, the growing brain is not always well-suited for brain scans. For example, the quality of MRI and the interpretation of data from it depends on the contrast between grey matter and white matter. This contrast changes dramatically over the first years of life, and so the visual appearance of the young brain on an MRI also changes.

Conceptual analysis

Perspective and causality

From a purely biological perspective, one may think about brain development as a genetically predetermined sequence of events—maturation. With today's state of research, we know the main components and phases of this process: it is an interplay of neurogenesis, migration of neurons, differentiation, and pruning. We know which parts of the brain are expected to go through which process at approximately what time.

However, even a purely biological perspective cannot do without incorporating environmental influences. In fact, the idea of experience-expectant neuroplasticity suggests that the division of biological and environmental variables is somewhat artificial. The brain has evolved to mature in a way that considers, and makes use of, typical environmental events. Of course, environmental events in the life of a human are, to a large extent, sociocultural.

What about cognitive variables? Evidently, biological development of the brain is connected with the development of cognitive functions. However, the exact nature of this structure–function relationship remains to be fully understood. It is possible that the process of maturation pre-determines the development of cognition. However, the opposite process also takes place—engaging in acts of cognition influences the development of the brain through experience-dependent neuroplasticity.

Bias and measurement

We have considered the problem of structure–function relationships. Traditionally, research of brain development and research of cognitive development have been conducted in parallel, but there is a lack of studies that explicitly connect the two processes. We have seen that conclusions in this area are often speculative, based on incomplete evidence. Simply establishing a co-occurrence of two events in time (for example, maturation of frontal lobes and an improvement in memory performance) is not enough to make any definitive conclusions. One research study is never enough to understand the nature of structure–function relationships. We always need the triangulation of multiple sources of evidence.

We have also considered the limitations of measuring the process of brain maturation itself. For example, brain imaging technology relies on the relative distribution of white matter and grey matter in the brain, which may not be the same in a growing brain, making comparisons of brain scans across age groups somewhat challenging. Additionally, brain imaging technology cannot be fully utilized with young participants. We cannot expect them to lie still in a scanning machine for 40 minutes.

Change

There are two types of development: maturation and change. Maturation is the type of development that is determined by an unfolding genetic programme. Change is the type of development that is caused by purposeful environmental influences. Traditionally, the study of a growing brain has focused on processes of maturation. Interesting facts have been established. For example, we know that the brain grows "inside out". There is not much we can change in this process. At the same time, it is evident that a full understanding of biological maturation of the brain is impossible without considering environmental influences. Neuroplasticity caused by gaining experience is an example of how the brain gets changed (as opposed to maturation).

Responsibility

Invasive research where we can manipulate biological variables provides better opportunities for understanding how the brain develops. For obvious reasons, such research is limited to animals. On the one hand, generalizability of such research to humans may be limited. On the other hand, it becomes especially important to follow the existing principles and regulations of ethical experimentation with animals.

3.2 Stage theories and continuous models of human development

Inquiry questions

- Does cognitive development of children occur in stages?
- Is cognitive development more like climbing a staircase or more like climbing a slope?
- Do individual differences in rates of cognitive development outweigh universal trends?

What you will learn in this section

Key learning:

- Proponents of discontinuity in cognitive development believe that there is a set sequence of stages, that people move through these stages in the same sequence at approximately the same rate, and that this progression is likely caused by biological maturation of the brain.
- Proponents of continuity in cognitive development do not necessarily reject the idea of stages. However, they do claim that stages have fuzzy boundaries, that individual differences in the rates of cognitive development outweigh common patterns, and that environmental influences on development are more important than biological maturation.
- Piaget's theory of cognitive development distinguished clear-cut stages: sensorimotor, pre-operational, concrete operations, formal operations. Studies used a range of typical tasks, such as conservation tasks and the three-mountains task. They established key cognitive phenomena at each stage, such as object permanence, reversibility, conservation, and cognitive egocentrism.
- If clear-cut stages of cognitive development really exist, then variations in task performance between age groups should be larger than variations between individuals from the same age group.
- Challenging this, subsequent research has uncovered significant variation in task performance within the same age group depending on environmental variables such as task difficulty, task content, and the amount of scaffolding provided by the researcher.
- For example, Borke's variation of Piaget and Inhelder's study allowed children to cope with a cognitive task one stage ahead of their age by simply changing the instructions without changing the nature of the task.
- Overall there is evidence that supports both sides. The truth is somewhere in between, but there is a lack of studies that would consider both biological maturation and individual differences in combination.

Key terms: continuity, discontinuity, stages of development, genetic epistemology, sensorimotor stage, preoperational stage, concrete operational stage, formal operational stage, object permanence, cognitive egocentrism, irreversibility, lack of conservation, conservation task, three mountains task, developmental unevenness, scaffolding, biological maturation

In a wider context

In the previous section, we discussed how the brain goes through a genetically predetermined process of **biological maturation**, following predictable patterns at predictable times.

The question is: does cognitive development follow the same pattern? If the answer is yes, then we can identify some predictable stages in cognitive development and describe at what age children are expected to be capable of what cognitive tasks.

The arguments for and against stages of development: an overview

One of the big ideas debated in developmental psychology is **continuity** versus **discontinuity**. Those who believe that development is continuous maintain that human development is relatively gradual, stable, and incremental. Those who believe that human development is discontinuous say that there exist distinct **stages of development**. There are periods of radical change between stages and periods of relative stability within them, with each stage uniquely different from all others.

Metaphorically you could think about continuous development as climbing a slope while discontinuous (staged) development is more like climbing a set of stairs.



▲ Figure 3.6 Child climbing up a slope



▲ Figure 3.7 Child walking up the stairs

If you are a proponent of a stage theory of human development, you are likely to also hold the beliefs that:

1. there is a set sequence of stages: people always progress from one stage to another and can never skip a stage
2. all people move through the same stages in the same sequence at approximately the same time
3. stages are influenced by the process of biological maturation of the brain: you cannot move to the next stage of development if your brain is not sufficiently mature—that is, if certain biological structures have not been formed.

To these three crucial ideas those who view human development as a continuous process suggest the following rebuttals:

1. stages are not clear-cut; boundaries between them are blurred
2. there are huge individual differences in when and how people mature, and the importance of such differences outweighs the significance of any common patterns in this process
3. development is also influenced by the environment—this influence is more important than genetic maturation.

Activity

Before you read further, imagine that your task is to design a ground-breaking study that will establish which of the two approaches (continuity or discontinuity) is true. You have unlimited money and resources. How would you design your study? Write a research proposal outlining your sample, method, procedure, and how you will analyse the results.

ATL Communication, Thinking, Research

As it frequently happens, the truth probably lies somewhere in between continuity and discontinuity. This is because, in real life, both processes occur simultaneously and interact with each other. Unfortunately, such artificial debates often lead to a situation where research studies are designed to test one of the theories but not both at the same time. For example, a great number of research studies have focused on uncovering common, universal stages of development without paying attention to individual differences. On the other hand, a large number of studies have focused on environmental influences and individual differences without any measurement of genetic predisposition or signs of universal stages. The process is vulnerable to confirmation bias. A researcher's theoretical orientation determines how a research study is designed, which in turn determines what kind of conclusions can be derived from it.

As Fischer and Silvern (1985) put it, "Because evidence is strong for both arguments, the fault must lie in the dichotomous nature of the arguments" (p. 614).

Jean Piaget's theory of cognitive development



▲ Figure 3.8 Jean Piaget

Jean Piaget (1896–1980), a Swiss clinical psychologist, is famous for the first theory of cognitive development. What he created was a stage theory. His ideas are sometimes referred to as **genetic epistemology** (epistemology means theory of knowledge). Much of Piaget's extensive research was done on his own three children.

According to Piaget's theory, children develop progressively in a series of stages that are related to age. He distinguished four key stages of cognitive development (there were other sub-stages within each of these, but we are only considering the "big" ones).

At the **sensorimotor stage** (from birth to two years of age), children's reasoning is driven by movement and sensation. Think of a toddler who accidentally shakes a rattle, it produces an interesting sound, so they shake it again. Two important experimentally observed phenomena characterize intellectual development at this stage: **object permanence** (understanding that objects exist even when they are outside your perceptual field; this is formed by the age of one year) and **cognitive egocentrism** (inability to understand other people's viewpoints).

Activity

You might have seen how amused young children are when adults play "Peekaboo" with them. The adult hides their face with their hands, then pops back into the view of the child and says "Peekaboo, I see you!". Why is this so much fun? Well, because until object permanence is formed, the child actually thinks that the adult's face disappears from reality and then pops back into existence! When you cover your face with your hands, you do not exist anymore. When you remove your hands, you exist again. Spooky, isn't it? That is how an infant's mind works.

Similarly, you might have witnessed how young children, when they play hide-and-seek, hide by simply covering their eyes. This is explained by cognitive egocentrism—if they do not see you, they believe that you do not see them either.

Search for some videos online demonstrating lack of object permanence in infants. What are your observations?

ATL Communication, Self-management

The **preoperational stage** (age two to seven years) is when children engage a lot in playing and pretending. For example, they can pretend that a stick is a spoon, a box is a house, and so on. Unlike during the previous stage, children at this stage are not manipulating objects as such, instead they play with meanings associated (sometimes loosely) with these objects. However, the child is still not capable of operations—that is, performing a task mentally in the absence of a physical object or at least a symbol such as a drawing. Characteristics of this stage are such phenomena as **irreversibility** and **lack of conservation**. Irreversibility is when the child is unable to mentally reverse a sequence of events. The concept of conservation means that, physically, an object remains the same even when its appearance changes: children at the preoperational stage do not understand this. Both of these phenomena can be seen in Piaget's classic **conservation tasks**.



▲ Figure 3.9 Piaget's conservation task

An example is when the child is shown two equal glasses (or beakers) of water and asked which glass has more water in it (the child correctly says that the amount of water is the same). Then the experimenter pours water from one of the glasses into another glass, narrower but taller. The child sees this. When asked which of the glasses now has more water in it, the child points to the tall glass. Piaget's interpretation is that the child is not yet able to mentally reverse the action (irreversibility), so there is a lack of conservation in the child's reasoning about the properties of water. Since the water appears taller, then there must be more of it.

Egocentrism is also characteristic of the preoperational stage. In fact, children overcome egocentrism by the end of this stage, when they turn seven or eight years old. A classic experiment demonstrating cognitive egocentrism in children is known as the **three mountains task** (Piaget and Inhelder, 1956). In this task, the child is shown a three-dimensional model of three mountains. There are some features (e.g., a cross, a house, or snow) that are visible from certain angles but not visible from others. After the child has spent some time exploring the model, the researcher introduces a doll. The doll is positioned facing the mountains from a viewpoint different from the child's, and the child is asked to describe what the doll can see. The child gives an answer by choosing one from a range of pictures. Piaget and Inhelder (1956) found that the ability to choose the picture corresponding to the doll's viewpoint dramatically varied with age: four-year-old children almost always chose the picture that corresponded to their own viewpoint (without realizing that the doll's perception might be different), and only by seven to eight years of age did they begin to pick the correct picture consistently.



▲ Figure 3.10 The three mountains task

At the **concrete operational stage** (age seven to eleven years) children are no longer egocentric. They start solving problems logically. However, they can only solve problems that apply to concrete events or objects. Inductive reasoning (drawing inferences from observations) is developed thoroughly, but children

Chat with AI

Have a conversation with your favourite generative AI to find out more about Piaget's theory of cognitive development. Here are some ideas of things to ask:

- Could you show me the full list of stages of cognitive development according to Jean Piaget, including the four main stages and all the different substages? Please provide a brief explanation for each stage and substage.
- What typical experiments did Piaget conduct to investigate cognitive development at the concrete operational stage? Could you give me some examples of tasks that he used?
- I don't fully understand why Piaget's experiments with the three mountains task support the ideas that: (1) children are cognitively egocentric and (2) there are clear-cut stages in cognitive development. Could you explain both of these?

still find it difficult to reason deductively. However, children at this stage acquire mental reversibility, which means that they can understand relationships between mental categories (e.g., if a dog is an animal, then some animals are dogs). Conservation tasks are passed successfully.

The **formal operational stage** (age 11–16 years) is when children develop abstract thought. Another new formation of this stage is metacognition ("thinking about thinking"). Adolescents become capable of deductive reasoning.

Piaget's theory was groundbreaking in developmental psychology. It introduced the world to the idea that children's intellectual development is a series of milestones. It also established a tradition of measuring development in a range of carefully designed tasks, and it asserted the existence of developmental stages.

Challenging the idea of stages: individual differences

Piaget's research suggested the existence of clear-cut stages in cognitive development. When we use a standard assessment task such as the three mountains, children of one age group almost never solve it correctly and children of the next age group almost always do. Surely this suggests that a certain brain structure matures at that age, making it possible for children to perform this sort of cognitive task.

However, the idea of stages of cognitive development rests on one important assumption: that variations between age groups are larger than variations between individuals from the same age group. Let us explore this assumption (see the "Activity" box).

Activity

Take two large objects (e.g., cups) and a large collection of small ones (e.g., marbles or pieces of candy). The small objects should be of two types (e.g., red and blue).

Draw a scale from 1 to 100. You could simply assume that your table surface is the scale.

Take the red marbles and cluster them somewhere around the beginning of the scale (e.g., the left side of your table). Take the blue marbles and cluster them somewhere around the high end of the scale (e.g., the right side of your table). The marbles will represent children of two different age groups, hypothetically belonging to two different stages of cognitive development, and their performance on a cognitive task. All blue marbles cope with the task successfully, but none of the red marbles do. Now put the two cups in the middle of the respective clusters to represent the average values of the groups.

What you have now is research data that obviously supports the existence of stages. However, real-life data is always fuzzier. The differences between groups are not always that dramatic (move the two cups a bit further away from the table edges and move the marbles with them). There is also individual variation within groups. Some red marbles actually perform very well on the task, and some blue marbles cannot cope with it. Spread the marbles further away from their respective cups while ensuring that the cup still represents the average of all marbles in that group. You will see that marbles will start mixing.

As you move these objects around, discuss the following with your partners:

At what point do we start saying that the boundaries between stages are so fuzzy that stages actually do not exist?

ATL Communication, Thinking, Research

A number of research studies have uncovered significant "**developmental unevenness**"—individual variations in the rate of development. Developmental unevenness means that there are a variety of factors that may affect the individual's progression through stages.

For example, it has been established that modifying task difficulty affects performance.

Borke (1975) replicated Piaget and Inhelder's (1956) classic study, but the task was simplified. Borke kept the basic design of the study but changed the content of the task, avoiding the use of pictures and giving children some initial practice. Grover, a character from the popular children's TV show *Sesame Street*, was used in this study instead of Piaget's doll. There were two identical objects for the display: one for the child and Grover to look at, and another one on a turntable next to the child. The display itself showed a farm area with details such as people, trees, animals, a lake, and a house. Borke reasoned that children would find Grover more relatable than the doll and therefore easier to identify with, and that turning the table to match the view is a task that is more hands-on and so easier to understand for young children. Results showed that children as young as three years old were able to use the turntable to correctly match Grover's viewpoint, and this ability was quite obvious in four-year-olds.

The researcher concluded that results of the original study (Piaget and Inhelder, 1956) might have been biased due to the nature of the task, which was cognitively complex and unfamiliar to the children.

The task still required the same skill though: the assumption is that, if you are cognitively egocentric, you will still be unable to cope even with this simplified task. However, a lot more children coped. At the very least, this means that boundaries of each stage are fuzzy and less clear-cut than we might have originally imagined.

Apart from task difficulty, research studies uncovered other factors affecting developmental unevenness, such as:

1. Task content. The same skill of perspective-taking is required in both a task where children look at an object from different angles and in a task where children describe a social situation from the perspective of various people involved in it. However, research has shown that children capable of one of the tasks at a certain age are likely to be incapable of the other task. How do we resolve this problem? Why have we decided to take the first type of task as the true indicator of that cognitive skill?
2. Minor variations in procedure. This has been established in studies where the researcher varied the degree of support or **scaffolding** provided to the child, such as Nedospasova (1985). See Unit 3.3: Sociocultural factors in development.

Overall, the pace of development can greatly vary in different children, different cultures, and different learning environments. In some cases (some specific cognitive abilities) reversal of stages was observed. If that is the case, the question is: how justified are we in talking about the existence of stages—if they do exist, how clear-cut are they?

Activity

Spend some time on the internet watching videos about Piagetian tasks with children. The key words for your search could include:

- Piaget tasks with children
- cognitive egocentrism task with children
- three mountains task with children
- Piaget conservation task with children.

Even more fascinating is to watch how children and chimpanzees compare on their performance of the same cognitive tasks. Try searching for:

- cognitive tasks, human children, and chimpanzees
- problem-solving chimpanzees
- children versus chimpanzees on cognitive tasks.

ATL Research, Self-management

TOK

Consider the following statement: "Children do x because they are in stage S". For example: "Children believe that the shape of a glass changes the amount of water in it because they are in the preoperational stage".

To make such a statement, we must establish three facts:

(1) That children are in the preoperational stage (this should be measured).

(2) That children believe that the shape of a glass changes the amount of water in it (this should also be measured).

(3) That (1) always precedes (2) in time.

Fact (1) must be established independently of (2). Otherwise, the argument becomes circular. We observe that children's judgements about the amount of water are affected by the shape of the glass. We describe this behaviour as the "preoperational stage". If we do not have any other measure independent of the conservation task to establish the fact that they are in this stage, then we are simply inferring (1) from (2). However, if this is the case, we cannot say that (2) happens because of (1) because it becomes a faulty, circular argument. According to Brainerd (1978), saying "Children do x because they are in stage S" in a situation like this is equivalent to simply saying "Children do x".

Therefore, the question is did Piaget have any procedure to establish that a child was in a certain stage of development, a procedure that would be independent of the cognitive tasks? It appears not.

Do you think it would be fair to say that Piaget's stages of cognitive development provide a description of the process, but not an explanation of it?

Conceptual analysis

Perspective

One way to use the concept of perspective as a lens to look at stage theories of cognitive development is through the relative importance of biological and environmental (sociocultural) factors in development. Stage theories traditionally rest on the assumption that cognitive development is driven by biological maturation of the brain. Continuous models of cognitive development do not necessarily reject the idea of stages per se, but they emphasize the importance of individual differences caused by environmental factors (such as task difficulty, task content, and amount of scaffolding provided to the child).

Another angle is the perspective we use when we interpret research results. Every study will always reveal a trend (such as a difference in average performance between two age groups) and some noise (such as the individual spread around the average in each of the groups). However, a stage theorist will claim that it is more important to focus on the trend, dismissing the noise as a random error of measurement or random fluctuations. In contrast, those who believe that development is continuous will claim that the noise is more important than the trend, and their studies will be designed to identify the most prominent sources of noise.

Causality

The investigation of stages of cognitive development has been more descriptive than explanatory. Researchers tried to establish which would be a better

description of the process: a "staircase" or a "slope". The question of what cognitive development is driven by (caused by) remained in this research as an implicit assumption.

Note that research supporting stage theories is mostly based on comparisons across ages: they take a cognitive task and see what percentage of children will cope with it in which age group. This is a quasi-experimental design. In contrast, researchers who investigated developmental plasticity were in a better position to use true experiments: for example, they could modify the difficulty of a standard task and see if this manipulation would affect the rate of performance.

Measurement

As always, our theories are only as good as the measurements that they are based on. You have seen in this section that there has been some controversy surrounding Piaget's typical tasks of cognitive development. Other researchers (e.g., Borke, 1975) have slightly changed the tasks and obtained results that were quite different. It is important for researchers to be able to compare their results with the results of previous research, so the classic Piagetian tasks have migrated from study to study for decades. Even when a researcher introduces a modification to the standard task (such as a simplified instruction or a more relatable character on the three mountains task), they keep the standard task as a control group or at least make modifications minimal in order to be able to compare to the original.

Bias

Any Piagetian task is a scripted scenario of a child interacting with the researcher around solving a particular task. Since interaction with a researcher is involved, it is more difficult to standardize such tasks. Bias may be introduced by small variations in how the procedure is carried out.

Then there is the big question: if we observe variations in how children perform on a task depending on slight variations in how the procedure is organized (e.g., how much the task is scaffolded by the researcher), what is it? Is this biased measurement or a genuine reflection of the fact that children are more influenced by social interactions than biological factors?

Change

The study of a child's cognitive development is a study of change. This makes it particularly challenging. How do we determine the major factors that affect this process? How do we separate universal patterns from individual variations? How do we account for the fact that younger children may be less likely to solve a task correctly not because they lack the cognitive ability to do so, but merely because they find it harder to understand the instruction? All of these challenges are typical when we study a phenomenon that undergoes change over time.

Responsibility

Experiments conducted to study children's cognitive development are typically non-invasive, friendly, and game-like. They are similar to what children would normally do in their interaction with an adult—for example, a teacher. This adds to the ecological validity of these studies, but also helps to ensure that the study follows all ethical rules. As participants are not of legal age, parental informed consent is required.

Exam tip

Many links between the specific content that you study in every unit and the six key concepts are transferable. This means that the same kind of link can be established in other parts of the course with some other content. For example, here we considered the idea that stages of cognitive development have been more descriptive than explanatory. This "descriptive versus explanatory" argument may be applied to other topics as well. In fact, it may be applied to any theory, and this would be relevant to the concept of causality. For example, how would you evaluate schema theory, the multi-store memory model (MSMM), the dual processing theory, and the social learning theory from the "descriptive versus explanatory" point of view?

Such transferable links are very valuable because they prepare you for anything that you may encounter on exam day. When you read the suggestions for conceptual analysis presented at the end of each unit, take special note of such transferable links and keep a list of them in a separate notebook or document. They will be useful when you revise for exams.

3.3 Sociocultural factors in development

Inquiry questions

- What is the role of education in cognitive development?
- Is culture more important than biological maturation in child development?
- How does a child learn cultural knowledge?

What you will learn in this section

Key learning:

- The most prominent theory in this field is Lev Vygotsky's sociocultural theory of human development.
- The theory made a distinction between low-order and higher-order cognitive functions (e.g., voluntary attention, semantic memory, conceptual thinking). It claims that low-order cognitive functions may be influenced by biological maturation, but higher-order ones depend on sociocultural factors.
- The theory stresses the importance of social interactions in psychological development. The source of cognitive development is social interaction with other people, and every cognitive function "appears twice": first interpersonally, then intrapersonally.
- The theory suggested that learning plays a leading role in development and that it should always be in the "zone of proximal development"—that is, one step ahead. This is different from Piaget's theory where learning remains "behind" biological maturation and merely supports this process.
- To achieve development in the proximal zone, the child needs help from others and such help may be provided by scaffolding.
- Vygotsky introduced the idea of externally mediated activity (activity with the use of tools). Just like we use physical tools to transform the physical reality around us, we use cultural (mental) tools to transform our own internal world. All cultural tools are signs and symbols, and the most important system of signs is language.
- Vygotsky's approach to research was different from the research methodology widely accepted in the West. He believed that the experiment was not an appropriate method to investigate the development of a child in an ever-changing interaction with others. Instead, Vygotsky and his followers used a large number of "formative experiments".

Key terms: sociocultural theory of human development, low-order and higher-order cognitive functions, voluntary attention, semantic memory, conceptual thinking, interpersonal process, intrapersonal process, zone of proximal development, out-of-reach zone, scaffolding, externally mediated activity, cultural tools, signs and symbols, language

In a wider context

So far, we have mainly considered how the brain and cognitive processes are affected by biological maturation, driven by genetic factors. The brain undergoes changes and at certain critical periods it becomes particularly responsive to certain experiences. This could be reflected in cognitive development which also appears to go through a predefined sequence of stages.

Some researchers disagreed with the idea of clear-cut stages in cognitive development and have discovered significant individual differences in this process. The differences are probably created by environmental factors. The biggest environmental factor in the life of a human is, of course, the society and culture they are growing up in.

The focus of this section is on these sociocultural influences. The most prominent researcher in this area was Lev Vygotsky and his **sociocultural theory of human development**.



▲ Figure 3.11 Lev Vygotsky

Sociocultural theory: history

The sociocultural theory of human development was formulated by Lev Vygotsky and his collaborators in Russia in the 1920s and the 1930s. The term used to refer to this theory is sometimes "cultural-historical", which is a closer translation of its original name. Vygotsky's work was widely known in Russia but not abroad. Even within Russia, his work was suppressed for political reasons for several decades. Vygotsky died aged 37 years from tuberculosis. However, in the early 1960s, his work started gaining prominence and recognition both in Russia and abroad, and a large number of scientists further developed his theoretical views.

Vygotsky is frequently mentioned in contrast to Piaget. Contrary to popular belief, Vygotsky's and Piaget's theories are not actually in direct contradiction to each other. Like Piaget's, Vygotsky's theory is built upon the idea that the child is an active explorer of the surrounding world and develops through interactions with the world. Piaget only read Vygotsky's work after his death. He openly praised Vygotsky for the demonstration that a child's egocentric speech was a stage in the process of internalizing external conversations (e.g., with the adult) into internal thoughts.

We will now outline the main components of the theory one at a time.

Activity

Split into groups and find out more about the biographies of either Vygotsky or Piaget.

Exchange some interesting facts that you find.

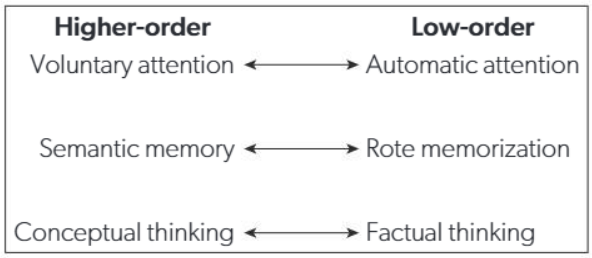
ATL Communication, Social

Low-order and higher-order cognitive functions

Vygotsky was particularly interested in cognitive processes that he called "**higher-order functions**". Examples of higher-order functions include:

1. **Voluntary attention**. This is when we deliberately focus attention on something, even though it may not be intrinsically attractive. Take a student preparing for an exam as an example. The **low-order** counterpart of voluntary attention is the kind of automatic attention that is triggered by something colourful and shiny. For example, think of a toddler whose attention is immediately captured when they see a cartoon.

2. **Semantic memory**. Semantic memory is memory of meaningful material, such as your memory of the role of biological maturation in Piaget's theory. Semantic memory requires you to understand the material and connect it in your mind to other relevant information. The alternative is rote memorization, which is a lower form of memory. You can rote-learn something even if you do not fully understand it.
3. **Conceptual thinking**. This is when you understand how factual material is related to underlying concepts and how concepts themselves are interconnected. Conceptual thinking may be opposed to factual thinking, which is when people operate with specific ideas and facts without seeing the bigger picture behind them.



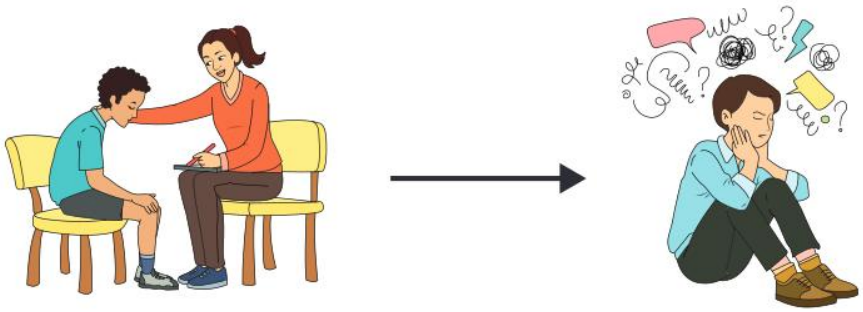
▲ Figure 3.12 Low-order and higher-order cognitive processes

Social sources of development

The crucial point of Vygotsky's sociocultural theory is that the basis of learning is interacting with other people (parents, teachers, peers, and other representatives of the child's culture). Vygotsky believed that every cognitive function in a child's development "appears twice". First it appears on the social level as a characteristic of interaction between people, then it gets internalized and appears on the individual level as a child's cognitive function. An **interpersonal process** (inter = between) becomes an **intrapersonal process** (intra = within) (Vygotsky, 1978).

For example, external control (e.g., when a parent tells the child to turn off the screen to limit screen time) gradually becomes self-control (when the child is capable of controlling their own screen use). Conversations with our parents gradually become our own thoughts.

You might have heard how young children use the "running commentary" to accompany everything they do. They might play with some plastic blocks, for example, commenting on their own actions: "Now put this block here, and this goes there…". According to Piaget, this "egocentric speech" is a stage in cognitive maturation. The child is not trying to speak to anyone, they are just trying out the new skill that is gradually growing from the inside. In contrast, according to Vygotsky, egocentric speech is compressed social speech. This means that these are parts of phrases and instructions that the child previously heard from adults and other children. The child is using these instructions to guide behaviour, but with age the instructions become internalized. They become thoughts.



▲ Figure 3.13 Internalization

Zone of proximal development

What is the role of learning in the development of higher-order cognitive functions? An important part of Vygotsky's theory is the concept of the "**zone of proximal development**" (ZPD). The ZPD is broadly defined through the comparison of three areas:

1. Zone of actual development: what the child can do by themselves.
2. Zone of proximal development: what the child can do with the help of an adult.
3. **Out-of-reach zone**: what the child cannot do even with a helping adult.

According to Vygotsky, the ZPD is where learning and education should happen.

Learning is (and should be) the driving force of development: it is always one step ahead of the natural maturation processes. In Vygotsky's interpretation, Piaget's theory viewed learning as something that follows "behind" maturation: material should be taught to children at a level that is suited to their current stage of intellectual development. Vygotsky did not agree. He thought that education should be built in the ZPD. Material should be too challenging for children to accomplish by themselves, but perfectly manageable in communication with a helpful knowledgeable adult. An example is the role of **scaffolding** in education.

Activity

Think about your own learning from the perspective of the zone of proximal development. Are the things you are learning in the IB DP programme in your zone of proximal development? Do you always ensure that your studies provide an optimal level of challenge: not too easy and not too hard? Do you effectively use the teacher's help? What kind of scaffolds are there for you to acquire the skills that are currently "one step ahead" of you?

ATL Thinking, Self-management

Cultural tools

Vygotsky introduced the idea of **externally mediated activity**—the actions that require the use of external means, or a tool, to reach a goal. For example, chopping wood with an axe is an externally mediated activity, but so is making a shopping list, or using your fingers to count something. This led to the idea that it would be impossible to understand cognitive processes without understanding the tools that mediate them.

Cultural tools are the elements that children and adults can use to help them process information. For example, consider memory. Suppose you need to keep track of the number of days it rained this year. If all you have is your own memory, then this may be challenging. However, culture has provided us with many tools that can help. We can take a sharp rock and put a scratch on the wall every time it rains. Or we can take a rope and tie a knot on it. The knot on the rope, the scratch on the wall, and, in more advanced cases, a calendar or an online weather database are all examples of cultural tools.



▲ Figure 3.14 Cultural tools: cave painting

TOK

What is the difference between a sign and a symbol? Can animals use signs or symbols? What is the difference between language and communication? Can animals communicate? Do animals have language?

Chat with AI

Try using your favourite generative AI as a (cultural) tool to further explore the perspectives of Vygotsky and Piaget on cognitive development. This will require some preparation.

- Write an outline of Vygotsky's theory in a few key statements. This is to provide AI with the context of exactly what you have learned and to what level of detail.
- Write an outline of Piaget's theory for the same purpose.

Use a prompt like the one below. Consider using the following endings supplied in the list.

- I am an introductory psychology student. I have been learning about the developmental theories of Vygotsky and Piaget. Here are a few key statements to give you an outline of what I have learned: [paste your outlines]. Could you help me understand how each of these theories would…?
- … explain egocentric/private speech
- … explain the fact that some children learn to read and write faster than others
- … design education in a way that would promote cognitive development
- … explain how children develop their understanding of moral principles.

According to Vygotsky, all cultural tools of cognitive development are **signs and symbols**. Our mental world reflects the physical world around us. Just as we use a physical tool to transform the physical reality around us, we use signs to transform the world within us: our cognition. If we neglected culture and only relied on what has been provided to us by biology, our reasoning would be very primitive.

The most important and culturally universal system of signs is **language**. This is why Vygotsky stressed the importance of language in internalizing the cultural content.

Vygotsky's approach to research

Vygotsky believed that the standard research methods (e.g., an experiment) were not suitable to test his theories. Instead, he relied on so-called "formative experiments". These involved situations similar to regular classroom practice in which the experimenter also acted as the teacher. The experimenter recorded the child's initial attempts to solve a problem without any help and then provided some scaffolding or tools through which the problem could be solved, supporting the child in trying to solve it. This type of "mediated assistance" implied that the interaction between the teacher/experimenter and the child would not always follow the same script, so it was not highly standardized.

Western psychology is based on a research methodology that emphasizes objectivity and elimination of potential biases. As a result, theories are tested in a relatively small number of studies that are specially organized, carefully controlled, and documented. Instead, Vygotsky encouraged every teacher to be a researcher. He relied in his theory on a large number of less controlled "formative experiments" happening in real classrooms. These studies were referenced in published papers, but they were rarely described in sufficient detail for another researcher to be able to replicate them exactly. Instead, papers focused on theoretical discussions.

Supporting study: Leontyev (1931)

One of the numerous formative experiments that supported the theory is by Leontyev (1931). The aim was to investigate how higher-order cognitive functions are developed through the use of cultural tools (signs). In this study, participants of different ages were required to memorize a list of 15 simple words read out by the experimenter. There were two conditions:

1. Participants were given 30 cards with pictures that they could use as a tool to help them memorize words. When participants heard a word, they could look at the cards and put aside one that would help them remember the word later.
2. There were no cards. In this condition, participants simply heard the words and had to memorize them.

Comparison across various age groups in these experiments showed the following:

- In pre-school children, memory was not aided by the cards. Performance on the task was equally poor in both conditions.

- In adults, performance was equally good in both conditions. Adults did not need the cards because the words were not difficult, and they had already internalized some memory strategies.
- In schoolchildren, however, performance on the task was poor (on the pre-school level) when they were not using the cards and good (on the adult level) when they were using them.

	Performance on the task		
	Pre-school children	Schoolchildren	Adults
With cards	Poor	Good	Good
Without cards	Poor	Poor	Good

▲ Table 3.1 Results of Leontyev (1931)

This formative experiment shows that school-age children can use cards as a cultural tool for extending memory, and that the adult level of performance on this task is in their ZPD.

Supporting study: Nedospasova (1985)

Nedospasova (1985) aimed to investigate Vygotsky's idea that social interaction with a knowledgeable adult in the ZPD and scaffolding from the adult can boost cognitive development. In particular, she was looking at the phenomenon of cognitive egocentrism in preschool children.

Nedospasova measured cognitive egocentrism (the dependent variable) with a variety of tests, including the three mountains task.

In the experiment, a child (a boy for this example) was presented with three dolls and told they were all brothers. The dolls were given names, and the child was asked to identify with one of the dolls with a prompt such as "This is going to be you, your name is…". He was then asked to say how many brothers one of the other dolls had. Cognitively egocentric children would answer "one" because they lack the understanding that they are their brother's brother. If the child said one, mistakenly, he was prompted with the dolls, and then he gave the correct answer.

After this happened, the same procedure was repeated using graphical representations, with an introduction such as "These three circles are three brothers, and this circle is you…".

When the child was successful in this trial, the researcher went on to the third stage and repeated the procedure verbally ("Imagine there are three brothers…").

This three-step procedure was meant to be a scaffold for children to learn to overcome cognitive egocentrism. Indeed, results showed that children who had participated in this formative experiment were much more likely to solve the three mountains task correctly with little or no help from the experimenter. This is despite the fact that, according to Piaget's stages, they should not be capable of doing so at this age.

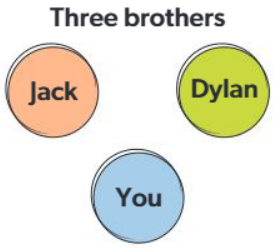


▲ Figure 3.15 Girl playing with cards

Exam tip

There is a lot of overlap between the topics "Sociocultural factors in development" and "Stage theories and continuous models of human development". This is because continuous models emphasized the idea that sociocultural factors are of primary importance in driving human development.

In this and other similar cases, remember that you are encouraged to make links between topics. It is also acceptable in many cases to use material from one section to support your response to a question that is based on another section. Unlike a book, human behaviour cannot be easily divided into chapters. Ideas covered in different chapters may be related in many different ways.



▲ Figure 3.16 Scaffolding task in Nedospasova's study

Activity

If you are an educator who fully embraces the principles of Vygotsky's sociocultural theory, how will this translate into your teaching practice?

- You will try to ensure that your teaching remains in the student's zone of proximal development.
- You will focus on providing cognitive tools rather than simple content—teaching students how to learn by themselves and how to think critically will be more important to you than teaching them some specific facts.
- You will emphasize the social nature of learning.
- You will recognize that students develop differently and that their learning is also embedded in a cultural context. For this reason, you will provide opportunities for students to choose their own pathways of learning.

How does this apply to the IB Diploma Programme in general and to IB Psychology in particular? The IB Programme has been heavily influenced by the sociocultural theory of development. Can you identify examples of this in your curriculum and the way lessons are conducted in your school?

Hint: think about the following:

- IB courses that set a challenging, but achievable target (e.g., TOK).
- ATL skills as opposed to memorization of content (consider the way exams are designed).
- Discussion-based activities and group projects.
- Individual choices embedded into the IB curriculum (e.g., choosing subjects, choosing topics for IA and EE).

ATL Communication, Thinking, Self-management

Conceptual analysis

Causality and perspective

Vygotsky's theory takes a sociocultural perspective. However, this does not mean that it denies any biological influences or neglects the biological maturation process. On the contrary, biological maturation is acknowledged, but the theory claims that it is low-order cognitive functions that are mostly affected by biological factors, whereas higher-order cognitive functions are more social (and cultural) in nature.

According to sociocultural theory, any cognitive function first exists as a function of interaction between two or more people. The interpersonal becomes the intrapersonal, and through this process the entire human culture is fed into the child's mental world.

Both Piaget and Vygotsky viewed children as "active explorers" who learn by interacting with the environment. For Piaget it was the physical environment that was his key interest, whereas for Vygotsky it was the cultural world of signs and symbols. In this sense, a part of causality also belongs to children themselves.

Measurement

Vygotsky's theory is very broad and makes general statements not only about child development, but about the very nature of the human mind. Some parts of the theory are philosophical and not directly testable in research. For example, how do you test the idea that every cognitive function has its beginnings in social interaction? In a theory that is so broad in nature, there is considerable distance between a broad theory and specific research.

An example of this gap between the construct and its operationalization is the study of Leontyev (1931), where cards with pictures (something very specific) were used to measure the role of "cultural tools" (something very generic). If children perform better on a memory test when allowed to use cards with pictures, does this necessarily lead to the conclusion that cultural factors are more important than biological factors in human development?

Bias

Vygotsky's theory was based on research studies that did not always follow the widely accepted Western research methodology. Vygotsky did not consider the experiment an appropriate method to test his theory. This is because the interaction between a child and a more knowledgeable other (see Unit 3.5: Peer influence; this could be a teacher, a researcher) cannot be standardized and cannot follow the same script for every child. However, the obvious objection from the Western research methodology is, of course, that such research is vulnerable to all sorts of biases.

Change

If we define change as purposeful development, then we might say that Vygotsky's theory is a lot more relevant to the concept of change than Piaget's. This is because of the role these two theories assign to learning and education. According to Piaget, child development follows a set of stages driven by biological maturation. There is not much we can do to accelerate this process. According to Vygotsky, education and scaffolding makes all the difference. Educators have to ensure that their instruction remains in the child's zone of proximal development at all times, and they must supply the necessary scaffolds for the child to develop efficiently.

Responsibility

Researchers are responsible for organizing their studies in a way that will not harm the participants. Research in this area is not invasive and not much different from what normally happens in the child's classroom. However, one must remember that participants are not of legal age and therefore informed consent should be obtained from parents or guardians.

Additionally, educators are responsible for ensuring that they provide the best possible education to support the development of a child in an evidence-based way. Vygotsky's theory suggests that the way we teach has an enormous influence on the way a child develops.

3.4 Enculturation of social norms

Inquiry questions

- How do children learn social norms?
- Can we learn social norms by simply observing cultural activities without participating in them?
- Can animals understand the concept of social norms?

What you will learn in this section

Key learning:

- Enculturation is the process by which people internalize the surrounding culture. Enculturation may occur in one of three ways: direct education, participatory learning, and observational learning.
- Most social activities are guided by social norms that share the following characteristics: normativity, generality, context sensitivity, and conventionality. Many social norms are implicit, unwritten, and situation-dependent.
- It has been observed that children as young as two years old understand the concept of social norms and try to enforce them upon others when they violate them (but only if the context is appropriate).
- Human beings seem to be the only species that develop social norms.
- A study has shown that five-year-old children are not merely recipients of pre-existing social norms created by adults. They can create social norms on their own and transmit them to other people. They understand the concept of a social norm—a rule that exists above individuals and is different from a whim or a wish of a single person.
- Another study has shown that in some cultures, observational learning is more important than active (participatory) learning as a mechanism of enculturation.

Key terms: enculturation, acculturation, socialization, participatory learning, observational learning, normativity, generality, context sensitivity, conventionality

In a wider context

In the previous section, we considered sociocultural influences on human development. The focus was on cognitive development—that is, the origin of an individual's cognitive functions (intrapersonal) in patterns of interaction with other people (interpersonal).

However, it is not only cognitive processes that undergo development. Another important dimension of a person's functioning in society is social norms. We are not born with social norms, they are acquired. The acquistion of these occurs in interaction with other people. How exactly it happens is the focus of this unit.

SAQ Enculturation

Enculturation is the process by which people internalize the surrounding culture. They learn social and cultural norms and acquire the appropriate values. The term "enculturation" usually refers to children growing up in a culture. When a person who has already acquired the norms of the culture they grew up in moves to another culture, that process is referred to as **acculturation**. We will discuss acculturation strategies in Unit 5.3: Acculturation.

The mechanisms of enculturation include direct instruction (education), social learning (when children observe the behaviour of adults and imitate them), peer interaction, and exposure to various other transmitters of culture such as mass media.

The concept "enculturation" is closely related to **socialization**, so much so that some authors use the two terms almost interchangeably. According to one interpretation, socialization is a process of deliberate enculturation (e.g., education, specially organized rituals, and activities). Enculturation may take the form of socialization, but it may also be non-deliberate (informal), such as subtle learning in the process of observing others and participating in joint activities.

Broadly speaking, it could be suggested that enculturation occurs in one of three ways: (1) direct education, (2) **participatory learning**, (3) **observational learning**.

We will consider some examples of enculturation in this section.

Social norms

Social norms are mutually accepted behavioural standards existing in a certain group or society. They are very powerful: people who do not follow social norms risk being ostracized or punished. Children begin conforming to social norms very early. This takes the form of following the rules imposed on them by adults. However, how exactly do children understand the concept of social norms? Do they see social norms as a wish of the single person imposing them, or do they understand that the person merely expresses a larger norm existing in the society? When a parent tells a child to clean up before guests arrive, is it simply because the parent wants it this way or is there a bigger rule behind it?

Most social activities are guided by conventional social norms. Such social norms share the following characteristics (Rakoczy and Schmidt, 2013).

1. **Normativity**. This means that social norms have the power to influence one's own behaviour and also be used to judge and evaluate the behaviour of others.
2. **Generality**. This means that norms apply equally to everybody, any individual, in any given circumstances.
3. **Context sensitivity**. This means that social norms always apply in a specific context. For example, it is fine to do stretching exercises in a park, but not in an elevator with other people.
4. **Conventionality**. This means that social norms are a product of agreement—people have agreed at some point that things should be done this way. It also implies that social norms are to some extent arbitrary.

TOK

You might have heard the expression "necessary and sufficient conditions". Example: We say that A is a necessary condition for B if B cannot exist without A. For example, lightning is a necessary condition for thunder. We say that A, B, and C are sufficient conditions for D if nothing else is required for D to occur. For example, a square has four equal sides and four equal right angles. Both conditions (having four sides and having four equal angles) are necessary for a figure to be a square. These conditions are also sufficient in that nothing else is required.

Here is a question: are normativity, generality, context sensitivity, and conventionality necessary and / or sufficient conditions of social norms? How would you approach analysing this question?

Chat with AI

Ask your favourite generative AI to provide more real-life examples of social norms. You could consider using the following prompt fragments (use them sequentially in a single conversation):

- Social norms in psychology are mutually accepted behavioural standards existing in a certain group or society. Could you give me five specific real-life examples of social norms, explaining (one for each example) why it counts as a "norm" and how children get to understand it as they are growing up?
- Social norms have four key characteristics: [insert a paragraph with definitions of normativity, generality, context sensitivity, and conventionality — this is not common knowledge, and we need to provide AI with some context]. Could you take the five examples you outlined before and apply these four criteria to each of them one-by-one, explaining how exactly each of the examples satisfies the conditions of normativity, generality, context sensitivity, and conventionality?

Having social norms is important for any group because norms promote social order and cooperation. Many social norms are implicit, unwritten, and situation-dependent. That is why a person arriving in a different culture may find it difficult to adapt, and that is why it is particularly interesting to investigate how exactly social norms are learned in the first place in the process of enculturation.

Early development of social norms

How can we check if a child understands the concept of a social norm, with its key features of normativity, generality, context sensitivity, and conventionality? The operationalization that has been used in research studies is the child's reaction to mistakes by others. When someone else (e.g., another child) breaks the norm, does the observed child try to enforce the norm by criticizing or protesting or sanctioning the mistake?

Children as young as two years old play social pretence games. For example, they can set up a pretend party and agree that leaves are plates and sticks are spoons and a chocolate wrap is the menu. Observations show that children understand these conventions in a normative way and when a new child joins the game and breaks some of the rules (e.g., they use an object for a purpose other than the intended one), the children will enforce these rules, criticize the new child, and teach him or her to play "properly". However, norms are not enforced if the new child announces that they are playing a different game. This means that context specificity is also understood (Rakoczy and Schmidt, 2013).

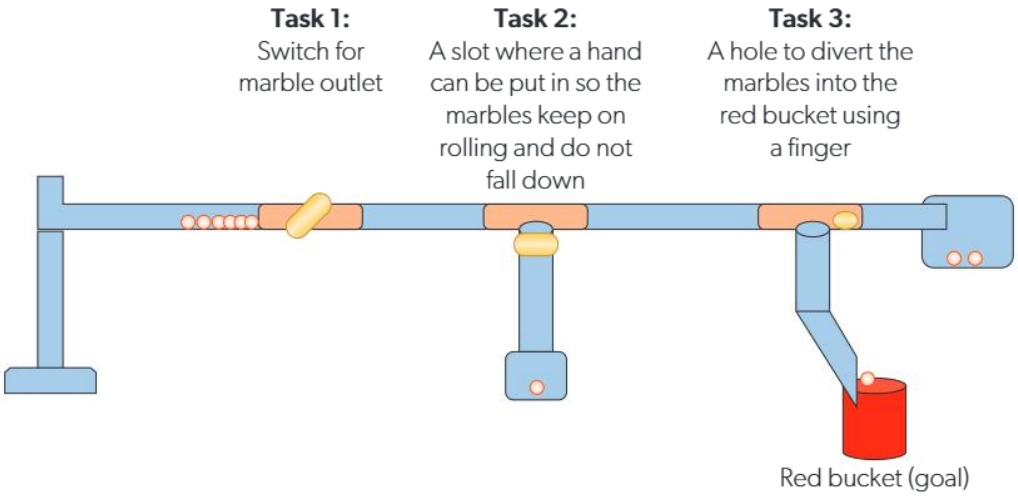
Interestingly, humans seem to be the only species that develop social norms. A properly developed social norm implies that we follow it ourselves and also enforce it upon others in a neutral way (that is, independent of our own interests in the given situation). Some primates demonstrate some primitive forms of social expectations, but their behaviour does not fit the criteria of a social norm (normativity, generality, context sensitivity, conventionality).

Some researchers believe that the understanding of social norms is what enables humans to build traditions. If children understand how things "ought to be done", they can remain faithful to tradition and contribute to it. In turn, faithful tradition building enables cumulative cultural evolution (when generation after generation builds upon the prior achievements). Cumulative cultural evolution is what really sets us apart from animals.

Example: how young children create and transmit social norms (Göckeritz, Schmidt, and Tomasello, 2014)

Göckeritz, Schmidt, and Tomasello (2014) investigated how young children create and transmit social norms.

Participants in the study were 54 German children, all five-year-olds. Children participated in triads (groups of three). They were presented with a marble run: a series of interconnected, steadily declining tubes, approximately four-metres long (see Figure 3.17).



▲ Figure 3.17 Marble run used in Göckeritz, Schmidt, and Tomasello (2014)

A total of 20 marbles were loaded through the opening at the start of the apparatus. The three critical locations in this system were:

1. a rotary switch that could let a marble through
2. an opening where the rolling marble would fall vertically through another tube into a dead-end, unless someone closed the opening with their hand
3. a hole through which someone could put their finger to divert the moving marble into a red bucket (which was the goal): if the marble was not diverted it would continue rolling into a dead-end.

The goal was to collect as many marbles as possible in the red bucket. Since the whole construction was four metres in length, the only way to win this game was for the three children to cooperate, with each child located at one of the three crucial points. Children won rewards depending on how many marbles they managed to collect in the red bucket.

As triads of participants were randomly formed, the experimenter pretended to not know how the marble run worked and told the children to work it out on their own. The trial ended when the children had rolled all 20 marbles. Up to seven trials were allowed.

On Day 2, the same procedure was repeated (again up to seven trials were allowed). No further instructions were given.

On Day 3, the triads were split and randomly recombined in such a way that each triad contained a child who had played the game on the two previous days (the expert) and two children who had never played the game before (novices). Again, up to seven trials were allowed.

For data analysis, video recordings of all sessions were transcribed and coded for the use of normative language. A total of 9,614 utterances were analysed. Examples of normative utterances were "You must put the marble in here", "The marble goes in here", "That's not how it goes", "You are supposed to use it like this". Such utterances used words and phrases like "must", "ought to", "may", "right" or "wrong", "good" or "bad". They were called "normative" because they clearly referred to existing rules or norms of behaviour.

Discussion

Can you remember what you know about inductive content analysis? What considerations do you need to keep in mind in order to ensure that the results of inductive content analysis are unbiased?

ATL Research

Exam tip

One of the most important skills for IB Psychology exams is the ability to select material that is most relevant to the question, and to clearly explain its relevance. Whether or not something is relevant to the question is always a judgement, and there are no strict rules. However, your judgement will become more mature and well-informed as you study the course, and it is assessed.

Even different contexts overlap sometimes. For example, a study of the creation and transmission of social norms could be related both to human development and human relationships. Depending on the question, you might be able to use content from one context to answer questions about another context. Always ask yourself if this is clearly relevant to the question, or only marginally relevant. Avoid marginally relevant content in your responses.

This was different from imperative utterances: requests made from one person to another without any reference to more general rules. Examples of imperative utterances were "Give me a marble" or "Put the marble in there".

Results of the study were as follows:

- The number of normative utterances peaked on the first trial on Day 1 (the first encounter of the marble run), then steadily declined from trial to trial, remained low on Day 2, peaked again on the first trial of Day 3 (the introduction of novices), then gradually declined again from trial to trial. Experts used significantly more normative utterances than novices.
- Observations during the study also showed that children created their own "arbitrary" social norms within the game. For example, one group of boys created the rule that in case a marble does not land in the red bucket directly but instead lands on the floor next to it, then the marble has to be put to the start of the marble run again. This arbitrary rule, once established, was passed on to all novices in the new groups using normative language. It was accepted by novices with no objections.

What this study demonstrates is that five-year-old children can create social norms on their own and transmit them to other people. Children created social norms to align their behaviour effectively even when there was no authority to guide them or give them feedback. At the same time, once established, these norms were passed on without questioning or negotiation, readily accepted as objective "standards" even when they were not completely adaptive (like the requirement to place the marbles that fell out of the bucket back at the start of the marble run). When experts explained the rules to novices, they did not use descriptive language such as "This is the way we played yesterday, we could play it like that again today". Instead, they used generic normative language such as "This is the way it is supposed to be done". This indicates that children understand the power of social norms and use them accordingly.

Example: the role of observational learning in enculturation (Odden and Rochat, 2004)

Most studies conducted with children from Western countries provide evidence suggesting that active (participatory) learning is a more powerful mechanism of enculturation than simple observational learning.

However, this belief may not be universally applicable. A study by Odden and Rochat (2004) suggests that, at least in some cultures, passive observational learning is the key mechanism of enculturation.

Participants in this study included 28 children aged 4–12 and their parents, from a village in Samoa. The methods used in the study were: (1) longitudinal naturalistic observation, (2) semi-structured interviews, and (3) questionnaires. One of the researchers (Harold Odden) lived in the village for 20 months.

Here are the key results of the study:

- Observations and interviews showed high specificity of Samoan attitudes towards education and enculturation. As there was a cultural value in power distance (the unequal distribution of power), questioning in class was discouraged as it was viewed as a sign of disrespect to the adult. Children

were largely left to learn things on their own, and adults did not try to engage them or motivate them. Caretakers believed that children learn by observing and listening to their elders.

- Samoan houses are often built without walls (due to the hot and humid climate), and activities of the nearby households may be easy to observe. Children and adults can freely observe the activities of others since the concept of privacy is not established in the society and can even be viewed as a form of secrecy or immorality. Although adults do not purposefully model their behaviour to children, children have plenty of opportunities to observe adults.
- By age 15, adolescents spend a significant amount of time doing chores (e.g., washing clothes, feeding domesticated animals, childcare, cooking). However, parents were never seen explaining the chore to the child or explicitly teaching the child in any other way. The same was true for fishing. By the age of 12, boys were quite skilled fishermen, but never in their lives had they been fishing under the supervision of an adult who would correct their mistakes. For example, if spearfishing was used, a boy would accompany his father, but there would only be one spear and one set of goggles available, and the child would never use it. Boys would just wait on the shore while their father was fishing in the lagoon. Then after a certain age, they would simply borrow the fishing equipment and start fishing by themselves without adult supervision.
- Knowledge of the intricate social hierarchy, ritual practice, and cultural concepts, such as power and authority, was also acquired in the same way.



▲ Figure 3.18 A village family in Samoa

The authors of the study concluded that each culture may use a different blend of mechanisms of enculturation. In the Samoan context, observational learning plays the central role in children's enculturation. It could be different in other cultural contexts.

These results are important for educators, especially in a multicultural classroom situation. Children who grow up in an environment that emphasizes observational learning as a mechanism of enculturation may be at a disadvantage

Discussion

Is Odden and Rochat's (2004) study etic or emic? Are there elements of imposed etic?

What would you say are the strengths and limitations of this study from the point of view of credibility?

ATL Communication, Thinking, Research

Activity

Find the original article of Odden and Rochat (2004). This is called "Observational learning and enculturation" and is available on the website of the Department of Psychology of Emory University. Being a qualitative study, it collected rich data. Look through the article and identify three additional findings that you think are significant for conclusions about cross-cultural differences in learning.

ATL Research, Self-management

in a Western classroom, which is typically focused on direct instruction, scaffolding, and active participation.

Conceptual analysis

Causality

The study of enculturation of social norms has focused on questions of "how" and "when". Social norms are internalized in the process of direct instruction, participatory learning, and/or observational learning. Social norms can also be created by children when they interact with each other while solving tasks that require cooperation.

However, why does enculturation of social norms happen? Perhaps because social norms serve a very important function in communal living and collaboration. Without them civilization would have been impossible.

Another aspect of causality is modelling the influence of one variable on another in an experiment. This is very limited in research of enculturation because it is often impossible to manipulate enculturation. We can only study it as it happens in real life.

Perspective

There is no doubt that enculturation of social norms is a sociocultural phenomenon.

However, there is a debate regarding what the main mechanisms of enculturation are. Different theoretical approaches to explaining the same phenomenon are also relevant to the concept "perspective". The main perspectives existing in research are that the main mechanism of enculturation is either active (participatory) learning or passive (observational) learning. The answer to this question appears to depend on the cultural context.

Bias

Research in this area heavily relies on observation. This is probably the weakest point of this research area from the point of view of bias.

Bias becomes a real possibility if either the researcher or the participants who are being observed (children) become aware of what is being investigated and/or the hypothesis of the study. This is why it is important to ensure that the study is blind (that is, the investigation and hypothesis are kept hidden) as much as possible.

One could also question if studies such as Odden and Rochat (2004) suffer from imposed etic. After all, the distinction between active (participatory) learning and observational learning is a Western construct.

Measurement

Enculturation of social norms is a broad, directly unobservable construct. If we want to measure it, we need to come up with some specific operationalizations. This is where researchers demonstrate their creativity. In the examples considered in this section, we have seen that some of the operationalizations included the proportion of "normative utterances" in a child's speech while they are playing a specially designed game, the behaviour of criticizing other children who violate existing social norms, and so on.

Change

Can enculturation of social norms be influenced on purpose? For example, can we bring children up in a way that will ensure that they internalize certain social norms effectively? Yes, we can, provided we understand the factors that influence enculturation of norms. For example, the study of Göckeritz, Schmidt, and Tomasello (2014) suggests that children develop social norms when they have to work together to solve tasks. We can use problem-based approaches in education and let children exercise these skills in a controlled environment under the supervision of adults.

Responsibility

Researchers need to observe all ethical considerations when conducting research. Although studies are not invasive, attention should be paid to informed consent (from children and parents) and the right to withdraw. Children sometimes lack the agency to withdraw from participation even if the study makes them uncomfortable.

There are ethical considerations in applying the results of research. For example, if we are not sensitive about the cultural context, we may overlook the importance of observational learning in enculturation of social norms in some societies.

3.5 Peer influence

Inquiry questions

- To what extent do peers influence child development?
- Is the influence of peers stronger than the influence of family?
- Do peers influence both social and cognitive development?

What you will learn in this section

Key learning:

- Peer influences on development are a special type of sociocultural influence.
- From the behavioural perspective, adolescents are influenced by their peers' behaviour due to mechanisms of social learning.
- From the identity perspective, the driving force behind peer influence in adolescence is the feeling of belonging which contributes to self-concept and the development of identity.
- Research of peer influence on development is very diverse: studies cover different types of behaviour, different peer groups, different age periods, and so on. This makes meta-analyses especially important.
- There are some key challenges that research in this area has to deal with. Peer influence needs to be separated from selection effects. True experiments are necessary to demonstrate cause and effect.
- There is a debate about the relative contributions of peers and family in a child's social development. It is more widely accepted today that the peer system and the family system are interdependent.
- Although most research in this area has focused on social development, peers have been shown to affect cognitive development as well. For example, engaging in rule-based games with peers may promote the development of executive functions such as working memory and inhibitory control.

Key terms: more knowledgeable other, behavioural perspective, identity perspective, internalizing behaviours, externalizing behaviours, meta-analysis, selection effects, cross-sectional study, longitudinal study, false consensus effect, parenting style, executive functions

In a wider context

Peer influence falls under sociocultural influences on behaviour. Interaction with adults is an important source of psychological development of a child, but it is not the only one.

Vygotsky believed that in order for social interaction to have a developmental effect, the partner in a given situation should be a "**more knowledgeable other**" (MKO). It could be a friend or a classmate who has made a bit more progress in a certain area. Piaget believed that peers provide an important resource for

cognitive and moral development because they are a source of alternative perspectives that allow children to challenge and develop their own schemas.

Overview of effects of peer influence

In their development, children are heavily influenced by other children. At some stages in their lives, such as in adolescence, peer influences can even outweigh all other influences (teachers, parents, etc.). Peers can influence children in a variety of ways, through a variety of pathways. These influences may be both positive and negative. However, traditionally research has emphasized the negative side of this dynamic—for example, the effects of peer pressure on smoking.

What is the mechanism behind peer influence? There have been two major perspectives on this: the behavioural perspective and the identity perspective.

1. From the **behavioural perspective**, adolescents are influenced by their peers' behaviour due to mechanisms of social learning. This perspective heavily relies on social learning theory and the idea of observational learning.
2. From the **identity perspective**, the driving force behind peer influence in adolescence is the feeling of belonging, which contributes to self-concept. This perspective is close to the principles of social identity theory. Adolescents identify with a group. This group membership contributes to their positive self-esteem, and it becomes important to them to behave in ways that are consistent with the group's identity.

Exam tip

If you study in the order of the sections as they are presented in this book, you are already familiar with social learning theory. You will learn about social identity theory in Chapter 5: Human relationships.

Peer influence effects have been discovered across the lifespan of humans—from toddlers and early school-aged children (e.g., Haun and Tomasello, 2011) to older people. However, converging evidence suggests that the effects of peer influence on behaviour are the strongest in adolescents.

Research of peer influence on development is very diverse. It differs in terms of the type of behaviour that is investigated (e.g., health behaviours, academic behaviours, violence), the type of peers (e.g., close friends, classmates, teammates on an online video game), the age period (e.g., early childhood, adolescence), and so on.

When there is a large variety of studies, it makes sense to try to combine them in a large **meta-analysis**. This is exactly the kind of task that was accomplished by Giletta et al. (2021) in "A meta-analysis of longitudinal peer influence effects in childhood and adolescence".

These researchers only considered studies that were **longitudinal** and included data from external reporters (i.e., they did not rely solely on self-reporting). A longitudinal study means that measurements were carried out repeatedly on the same group of participants over years. Using these selection criteria, they ended up with 233 effect sizes from 60 different studies conducted in 10 different countries (mostly European countries, the USA, and Canada, but also Asian countries such as China and Indonesia). The combined number of participants in these studies was over 47,000.

Overall, their results revealed that the effect of peer influence on the behaviour of children and adolescents was small in magnitude (0.08), but statistically significant and robust. "Robust" in this context means that the effect was stable across different sub-groups. For example, it did not show much variance depending on the behavioural outcome. It also remained stable across age and across countries.

Discussion

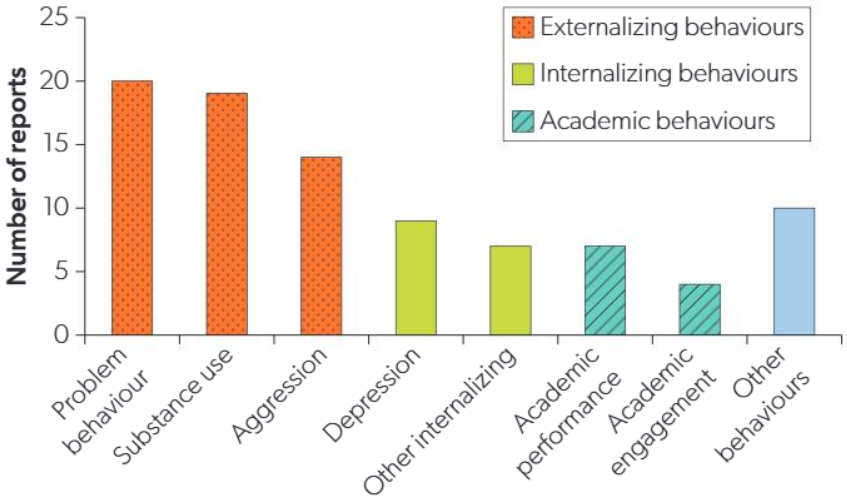
Why is a meta-analysis useful to understand peer influences on development? Why is it necessary in the first place?

If you were to formulate a conclusion from Giletta et al.'s (2021) meta-analysis in just one sentence, what would it be?

Although the effects of peer influence discovered in this study were statistically significant, the average effect size is very small. Recall what you know about effect sizes (see Chapter 1: Research methods and data interpretation). How would you interpret it?

Note: the analysis only included longitudinal studies and only those that did not rely on self-reporting. Could it be that self-report studies that are not longitudinal inflate the significance of peers in child development?

ATL Communication, Thinking, Research



▲ Figure 3.19 Types of behaviours examined in Giletta et al. (2021)

The only prominent difference was revealed in the time lag (peer influence had more effect on behaviour in the short term than in the long term).

However, even this meta-analysis is very limited in scope and does not exhaust the entire research question. Giletta et al. (2021) focused on the following behavioural outcomes (dependent variables):

- **Externalizing** behaviours (e.g., aggression, substance use, and risk-taking).
- **Internalizing behaviours** (e.g., depression and anxiety).
- Academic attitudes and behaviours (e.g., homework completion, school achievement).

There are other important behaviours that are not included in above list, but are important in the context of child development. Examples include: moral development, cognitive development, personal values and beliefs, affective states.

Research challenges

There are various challenges that a research of peer influence typically encounters. These challenges are universal, in the sense that they are relevant to all research in this area.

1. Peer influence versus selection effects

One of the research challenges in this area is separating effects of peer influence from "**selection effects**". Suppose we observe in a study that adolescents who have friends engaging in risky behaviours also engage in risky behaviours. There can be (at least) two alternative explanations here:

1. Friends have influenced this person to engage in risky behaviours.
2. This person was personally inclined to engage in risky behaviours, so they have selected friends who have the same inclination.

The first explanation would be related to peer influence, but the second explanation is a selection effect. It is indeed established that similarity is a reliable predictor of friendship formation. People who are more similar are more likely to become friends. This principle was initially proposed by Byrne (1961) under

the name of "similarity-attraction hypothesis" and it was confirmed in a series of studies in various real-life settings.

2. Experiments versus correlational studies

Another research challenge is conducting true experiments to establish cause–effect relationships.

Such studies have been conducted. An example is Cohen and Prinstein (2006). In this study, 16- to 17-year-old male participants from the same school were led to believe that they were interacting in a "focus group" with three other students in an online chatroom. In reality, responses of these other students were pre-programmed. Responses included statements that endorsed aggressive and risky behaviours. The manipulated variable was the perceived social status (popularity and reputation) of these other students in the chatroom. If the perceived peer status of these other students was high, the real participant was more likely to evaluate such behaviour positively, and more likely to engage in such behaviours after the study in real life. Researchers concluded that perceived social status of peers can influence participants' attitudes.

An advantage of this study (as well as all the true experiments conducted in this area) is that we can conclude with certainty that a high peer status causes the power of peer influence to increase. A disadvantage of this and other similar studies is ecological validity: it is debatable that studies like this approximate real-life situations. They create refined artificial conditions to control potential confounding variables, but in doing so they become artificial themselves, and there is no guarantee that the way participants behave in an experiment will generalize to real-life situations.

A more common approach is to use surveys and the correlational design. The advantage of surveys is that they allow us to investigate a broader range of behaviours than would have been possible in a laboratory. Examples include violence, substance use, self-injury. Additionally, correlational studies based on surveys can be longitudinal. This allows us to trace children's behaviour over time as peer influence continues to affect them in real-life contexts. Longitudinal studies (as opposed to **cross-sectional studies**) also allow us to separate peer influence from the effects of friendship selection.

3. Self-report measures

Another challenge is self-report measures. When we rely on self-report measures in investigating peer influence, we ask participants to rate their own behaviour as well as the behaviour of their peers (e.g., their friends). For example, if we assess peer influence on moral development, we might ask participants two kinds of questions: (1) How acceptable do you think it is to cheat on a test? (2) How acceptable do your friends think it is to cheat on a test?

The problem here is that we are not measuring behaviours and attitudes of peers directly but, instead, we are measuring perceived peer behaviours and attitudes. Participants can misrepresent real-life peer attitudes and behaviours for a variety of reasons, including cognitive biases. One of such biases is the **false consensus effect**—the tendency to perceive our own beliefs as relatively common to other people, especially our friends. Therefore, it is important to conduct studies in which peer behaviours are assessed by surveys given to peers themselves. Unfortunately, such studies are more difficult to conduct and for that reason much less frequent.

Chat with AI

Peer influence is a very broad topic and generative AI could be a good tool to help you explore it further without sifting through multiple research articles. You just need to know what to ask.

Consider the following prompt fragments to support your understanding of this topic:

- Could you please give me three or four examples of how peers can positively influence the social development of children? Then three or four examples of negative influences.
- Building on these examples, would it be possible for you to outline how each of them would have been explained by Piaget versus Vygotsky?
- Could you help me understand how the idea of peer influences on psychological development is related to the idea of sociocultural factors in development? Is peer influence just an example of sociocultural factors or is there anything more to it?

Remember to not take the output of AI at face value. If your generative AI has the function of providing links to the sources it used, it would be a good idea to explore the ones that you find most interesting or surprising.

Activity

Suppose you have reasons to believe that there is a strong negative peer influence in your school community that affects the health behaviours of adolescents (e.g., smoking).

Can you design some strategies that will minimize the negative effects of peers on health behaviours? Note that it is the peer influence that you need to target, or the individual's susceptibility to peer influence. It is not enough to simply promote a healthy lifestyle.

Exchange your ideas with other groups and discuss.

Ask generative AI for some suggestions once you have an initial idea.

ATL Social, Thinking, Self-management

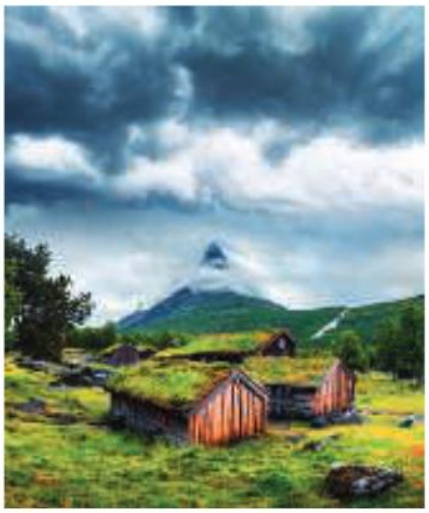
Relative contributions of peers and family

There is a debate about the relative contributions of peers and family in a child's social development. One position claims that the influences are independent of each other: peer and family systems develop separately and perform different functions. The other position views the peer system and the family system as interdependent. The latter position is more widely accepted today.

A variety of research studies have demonstrated that the **parenting style** and the quality of parent–child relationships are associated with success in establishing positive peer relationships. For example, Pettit and Harrist (1993) observed how mothers interact with children during family meals and correlated this with the extent to which children were accepted by peers and were successful at establishing peer relationships. It was found that positive interactional styles during family meals (i.e., being supportive and proactive, involving children in the conversation) were associated with greater social competence of the children in the classroom and in the playground (such as cooperative play). In contrast, children whose parents were more controlling or intrusive during dinner were also rated as more aggressive by their teachers, engaged in less cooperative play and more solitary play, and avoided contact with peers during free time.

A possible explanation is that parents provide the skills that will be necessary in the future to effectively deal with friendship formation. For example, managing your emotions is very important in peer relationships. If you are angry, you need to be able to control yourself and make sure that the anger does not translate into acts of aggression. If you are frustrated, you need to know how to express the frustration in socially acceptable ways. All these skills may be the essential building blocks provided by parents as they establish their patterns of interaction with the child.

Peer influence on cognitive development



▲ Figure 3.20 A farm in Norway

Perhaps not surprisingly, it has been argued that the main agents of developing cognitive skills are school and family, whereas peers may be the main agents in developing social skills. For example, Hollos and Cowan (1973) studied children who grew up in isolated farms in Norway. Naturally, these children had rare encounters with same-age peers. When they were compared to controls, it was observed that social skills were impaired but cognitive skills (logical operations) were not. This suggests that a lack of peer interaction in childhood has detrimental effects on social development, but not so much on cognitive development.

Although studies in this area are not frequent, they do show that peer interactions are impactful not only for social, but for cognitive development as well. One of these studies is Lecce, Bianco, and Ronchi (2020) who investigated the role of peer relationships in the development of **executive functions** (working memory and inhibitory control) in school children. Executive functions are cognitive processes of control and self-regulation. Working memory has a limited capacity and therefore has to distribute resources effectively to handle information. Inhibitory control is a self-regulation skill of suppressing one's impulses or habitual reactions to stimuli. For example, a child needs conflict inhibition to not start eating another child's candy.

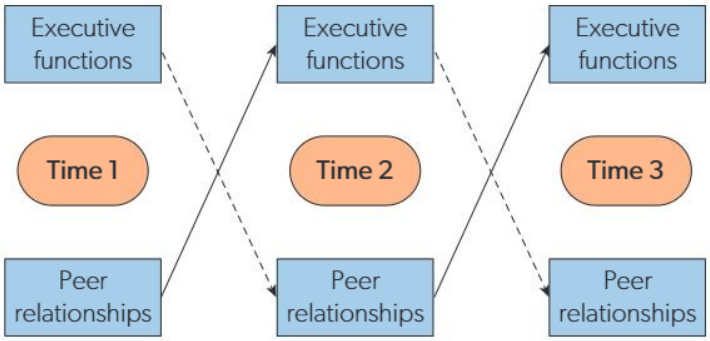
Theoretically, it is plausible that both directions of causality take place:

- Executive functions influence the development of social skills because successful social exchanges require that you inhibit your impulses and regulate your own behaviour.

- Social skills influence the development of executive functions because self-control itself is developed in situations of peer interaction, such as pretend play and play with rules. For example, a child may be required to play the role of a castle guard and stand motionless for a long time while others are acting out a battle.

Lecce, Bianco, and Ronchi (2020) conducted a three-wave longitudinal study, with three measurements separated by an interval of six months.

The study involved 245 children aged 8–12 years old (at Time 1) attending a public school in Italy. Time 1 took place in December, Time 2 at the end of the school year, and Time 3 at the start of the following school year.



▲ Figure 3.21 Two types of longitudinal relationships investigated by Lecce, Bianco, and Ronchi (2020)

Results of the study revealed no longitudinal association between executive functions and peer relationships: working memory and inhibitory control scores at Time 1 did not predict peer acceptance and peer rejection scores at Time 2. The same was true for the transition between Time 2 and Time 3. This goes against the theory that executive functions are important for developing peer relationships in middle childhood.

On the other hand, peer relationships significantly predicted executive functions over time: those children who were accepted by peers at Time 1 also had higher scores on working memory at Time 2. The same was true for the transition between Time 2 and Time 3. Those children who were rejected by peers showed poorer development of inhibitory control over time. These results support the theory that cognitive functions form and develop in the process of social interactions, including interaction with peers. Therefore, they support Vygotsky's theory of cognitive development.

Conceptual analysis

Causality

We know that peers have a great amount of influence on the development of children and adolescents, but the flow of causality in this process is complicated. There is ambiguity between peer influence and selection effects. It is also possible that relationships within the family affect peer relationships which in turn affect development. In other words, peer effects in human development may be indirect, delayed, or mediated. Some effects can also be self-reinforcing. For example, the development of executive functions can lay the foundation for engaging in peer interactions, which in turn promotes the development of executive functions through play.

TOK

Why is it so important for researchers to uncover directions of causality in the complex knot of interrelated variables? Why can we not just accept that human development is a complex phenomenon in which causality cannot be reliably established?

Compare this to other areas of knowledge, such as the natural sciences. These seem to be built on the principle of determinism—the belief that every event has a preceding identifiable cause.

Are there any limits to determinism? Consider examples, such as statistical physics, quantum mechanics, and chaos theory. All of these developments in natural science deviate from the belief that the world is a strictly determined place.

Perspective

Peer influence is a sociocultural factor of development. It deserves a special mention because the mechanism of influence may be quite different from how development is affected by interaction with others. Moreover, the focus in this research traditionally has been social development rather than cognitive development. Parents and other adults cannot provide children with sufficient experiences to fully develop their social skills.

There are two perspectives on how exactly peer influence works: the behavioural perspective and the identity perspective. Additionally, some researchers used to believe that family influences and peer influences acted separately, but there is a growing recognition today that both factors work in interaction with each other.

Bias and measurement

Investigation of peer influences on child development is complicated by several typical challenges that we discussed in this section. True experiments are necessary to investigate the "influence", but there is a trade-off between internal and ecological validity. True experiments may be too artificial, limiting their generalizability to real-life situations. Correlational studies are preferable when we want to keep the research more natural and investigate the role of multiple variables at the same time. However, there is the problem with self-report measures. We cannot be sure that participants will be able to impartially judge the extent to which they are affected by the behaviour of their peers.

Additionally, since it is the process of development we are talking about, longitudinal studies are preferable. They are also more difficult to conduct and there are not so many of them. It is especially interesting that a recent meta-analysis of longitudinal studies (Giletta et al., 2021) revealed an average effect of peers that was statistically significant, but quite small.

Change

Any investigation of human development is an investigation of change. Unpacking the flows of causality in peer interaction sometimes allows us to design strategies that could potentially influence child development in a desirable direction. For example, the longitudinal study of Lecce, Bianco, and Ronchi (2020) suggests that peer interactions affect executive functions but not the other way around. This hints at a possible intervention strategy where we can attempt to support the development of executive functions by purposefully including the child into controlled situations of peer interactions where they would have to play games with rules in teams.

Responsibility

Peers can have a large influence over the life of a growing person. It is our responsibility to understand exactly how large this influence is, as well as the main points of vulnerability. Peer influences can be both positive and negative. We want to minimize the negative ones and maximize the positive ones. The appropriate way to do so is probably to prepare the child to be able to pick their own "niche" in peer relationships and to surround themselves with peers and relationships that are beneficial for their own growth.

3.6 Role of childhood experiences

Inquiry questions

- Can the effects of adverse childhood experiences be reversed?
- What makes children capable or incapable of overcoming trauma?
- Is there a biological explanation for why traumatic experiences negatively affect development?

What you will learn in this section

Key learning:

- It is widely established that adverse childhood experiences create a lasting negative impact on development. For example, poverty experienced in childhood leads to slower development of cognitive abilities, deprivation during a critical period can lead to irreversible loss of cognitive function, early trauma results in issues with mental health and social adjustment.
- There are individual differences in the way people react to adverse childhood experiences. These differences are related to resilience—an individual's ability to recover from harmful effects of adversity.
- The resilience of children to a large extent depends on the resilience of parents.
- Resilience may be culturally specific.
- A key unanswered question is whether effects of adverse childhood experiences on development are reversible. One possibility is that such experiences keep affecting human development over the long term. Another possibility is that effects are strong, but they are gradually replaced by effects of more recent experiences.
- Another key question is the mechanism underlying the influence of adverse childhood experiences on further development. Some promising results have been achieved in epigenetic research that has demonstrated how adverse events can result in epigenetic changes which persist throughout lifetime unless reversed.

Key terms: adverse childhood experiences, poverty, deprivation, trauma, cortisol, resilience, reflexivity, domino causality, epigenetic changes, gene expression, DNA methylation, glucocorticoid receptor (GR) gene

In a wider context

When something is broken, we can understand more easily what it is made of and how it works. In a similar way, when the normal process of development is damaged, we can see more clearly what role various factors normally play in it.

In this section, we will discuss how early adverse experiences—such as poverty, deprivation, and childhood trauma—affect human development as well as the extent to which people can be resistant to such negative events.

Overview of effects of early childhood experiences on development

It has been well established by now that early childhood experiences are highly consequential for human development and have long-term effects on health, cognitive functioning, social development, and interpersonal relationships. **Adverse childhood experiences** create a lasting impact.

Here are a few examples:

- Growing up in **poverty** negatively affects cognitive development. Poverty episodes experienced earlier in life have a more detrimental and long-lasting effect than poverty episodes experienced later. If poverty episodes in early childhood were long and persistent, then the chances of correcting the negative impact later in life are very low. For more on this, see our discussion of the effects of poverty on cognitive functions in Unit 2.8: Environmental influences on cognitive processes.
- **Deprivation** of experiences related to a cognitive function during a critical period (the window of time when the brain is especially prepared for developing this function) may lead to a situation where the cognitive function is irreversibly lost and cannot be developed later, no matter how much training is provided. See more on this in Unit 3.1: Brain development.
- **Trauma** during early childhood can cause long-lasting issues with mental health and social adjustment. It is especially well documented that early traumatic experiences often result in long-lasting post-traumatic stress disorder (PTSD).

Activity

It is estimated that the Rwandan Genocide, which took place in 1994 in the context of the Rwandan Civil War, took the lives of more than 800,000 Rwandans. Another two million people were displaced and became refugees. The core of the conflict was a mass slaughter of the Tutsi people by members of the Hutu majority government. Civilians were killed rapidly (within three months) and mostly by hand-to-hand combat, primarily with machetes. The horrors of this genocide are well documented online.

The genocide gave rise to a PTSD epidemic. A survey carried out shortly after those events estimated that 22% of the country's population met the diagnostic criteria for depression and another 20% for PTSD. Many were children (Rieder and Elbert, 2013).

In one of its outreach programs, the United Nations (UN) has collected first-hand testimonies of survivors of the Rwandan Genocide. These can be accessed on the UN website.

Review these testimonies. Discuss how you would plan a humanitarian mission to reduce the PTSD crisis in Rwanda at that time if you were made responsible for doing so.

ATL Communication, Research, Self-management

Resilience

A phenomenon related to adverse childhood experiences is **resilience**. Resilience can be defined as the capacity to adapt to stressful situations and recover, or "bounce back", from harmful effects of severe adversity. It has been noted that some children get affected by adverse circumstances less than others, that is, those children are more resilient. Obviously, there is a lot of practical value in studying what exactly makes some children more resilient than others. If we cannot prevent all children from experiencing adverse circumstances, perhaps we can do something to promote their resilience.

One of the early findings regarding resilience was that the resilience of children depends on the resilience of parents. To illustrate this, we will consider the study of McFarlane (1987).

On 16 February 1983, a series of bushfires in Australia resulted in a loss of 75 lives and the destruction of more than 3,000 buildings. In the 26 months after the fire, McFarlane (1987) studied 808 children. The study relied on reports from parents and teachers. No children were directly interviewed.

There were some counterintuitive findings. For example, it was found that the most powerful predictors of post-traumatic phenomena experienced 26 months after the fire were:

- separation from parents in the days immediately after the fire
- continuing maternal preoccupation with the disaster
- changed family functioning.

In other words, it was important that parents stayed with the children after the fire, that the mother "bounced back" and started thinking about possible ways forward, and that the family remained intact. These factors predicted post-traumatic phenomena even more strongly than degree of exposure to the fire and losses suffered by the family (McFarlane, 1987).

It is important to bear in mind that resilience may be culturally specific. Methods of coping with adverse childhood experiences that "work" in one culture will not necessarily "work" in a different culture. An illustration of this point is the research of deVries (1984) who reviewed a number of Western studies and concluded that infants with a "difficult" temperament (who were difficult to raise, intense, inflexible, and irregular) cause more family stress—this makes it more likely for them to experience abuse and developmental problems, as compared to infants who are more docile.

With this in mind, deVries (1984) conducted a field study to investigate if the same would be true for Maasai children overcoming a natural disaster. The sub-Saharan droughts in 1974 disrupted the life of the Maasai people of east Africa and resulted in higher rates of infant mortality. After analysing questionnaire responses from mothers of 45 infants, the researcher selected 10 infants who were identified as easy to manage and 10 infants who were identified as difficult to manage. They were scheduled for follow-up two or three months later to assess and compare the level of nutrition.

Discussion

Why were no children directly interviewed in McFarlane's (1987) study? After all, it was the resilience of the children that was being studied. Wouldn't it be more reliable to get information from the primary source? Think about the interplay between methodological (ensuring credibility) and ethical (protecting children from harm) considerations involved in this research. It can be traumatic for children to relive the experiences during an interview. All things considered, do you think it was the right decision not to interview the children directly?

ATL Thinking, Research

Discussion

Reflexivity is an important part of qualitative research.

How can reflexivity be used in interviews and observations? What is the scope of using reflexivity in deVries's (1984) study?

Are there any special ethical considerations involved in the study of Maasai children undergoing a very harsh period of life in which survival is at stake?

ATL Thinking, Research



▲ Figure 3.22 Maasai people

Unfortunately, during these months the drought worsened, and many families were forced to relocate to distant areas. Researchers were able to find 13 of those 20 families anyway (seven "easy" and six "difficult" infants). Although the number of cases was small, observation showed, to the investigator's surprise, that more demanding infants were more likely to get nutrition and survive during periods of severe ecological stress. During the research period itself, five of the "easy" infants and two of the "difficult" infants died. It was concluded that infants who are more aggressive and demanding are more likely to be fed and have their needs satisfied. This indicates that the cultural context may be important for determining factors of resilience (deVries, 1984).

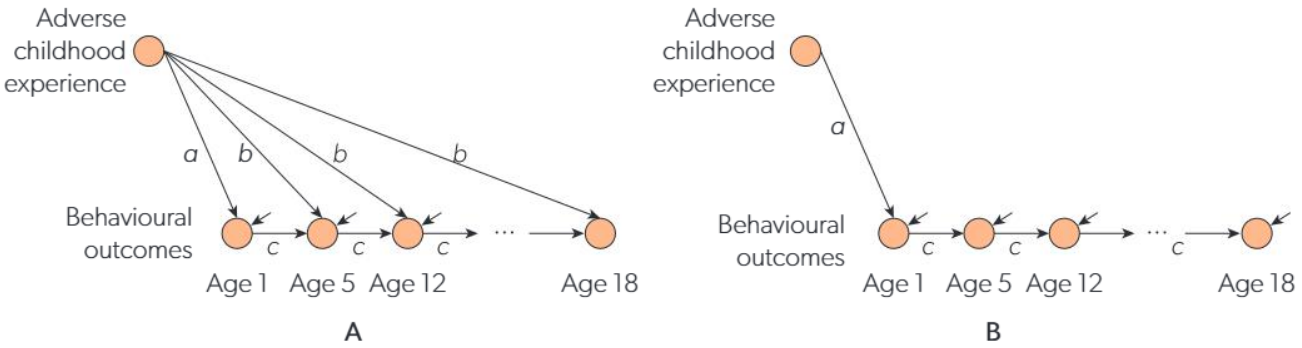
Unanswered questions

Despite extensive studies, some important questions remained unanswered. This could be because it is difficult to conduct a study to find out the answer, or because research yields inconsistent results. In any case, it is important to understand the key questions that currently drive research efforts in this area.

Question 1: Are the effects of adverse childhood experiences irreversible?

There is no doubt that adverse childhood experiences produce a negative impact on development. The debate here has been focused on how reversible these effects are and how inevitable it is that they will remain over a lifetime. Some researchers have argued that, as children grow older, they acquire new cognitive skills, establish new relationships, and encounter new experiences that are different from those they had in early childhood. These new experiences may give them an opportunity to weaken the association between adverse experiences in early childhood and negative outcomes in later life. Other researchers have argued otherwise—that the effect is persistent over the long term, stable, and unlikely to change.

Roisman and Fraley (2013) represented the difference between these two perspectives mathematically and graphically as shown (in a simplified form) in Figure 3.23.



▲ Figure 3.23 Two perspectives on the influence of adverse experiences on child development: persistent effect (A) versus one-off effect (B)

If an adverse childhood experience influences development in the long term, this is probably because it continues its causal effect at any point in the future. Therefore, we can model a behavioural outcome at a time point (e.g., at age 12) as being influenced by:

- this behavioural outcome at a previous time point
- the continued effect of the adverse childhood event
- any other factors.

This is plausible if we imagine, for example, that the adverse event has caused a permanent change in the structure or function of the brain. The abnormality remains as the child is growing up, so it will continue affecting behavioural outcomes directly. In this case, we should expect to see an effect that is consistent and does not diminish over time (A in Figure 3.23).

On the other hand, if an adverse childhood experience only affects us in the short term, then at age 12 it will still affect us due to "**domino causality**", but there will be no direct influence. In this case, we might expect the negative effect to diminish over time (B in Figure 3.23).

The problem is, testing which of the two hypotheses is correct requires longitudinal studies and longitudinal statistical analysis, which unfortunately have rarely been done. A typical longitudinal study simply measures an adverse childhood experience at some point during childhood and correlates this with a measure of behavioural outcomes at a later point of time. Roisman and Fraley (2013) argue that a single correlation is not enough as it may be consistent with both perspectives outlined above. Instead, we should be looking at how patterns of correlations are changing over time as we are following participants longitudinally.

TOK

Domino causality, or the domino effect, is a kind of causality where one event sets off a chain reaction. It affects another event, which in turn affects the third event, which in turn affects the fourth event, and so on. An analogy here is a falling row of dominoes where each piece knocks the other one over. Domino causality is interesting because each of the events (apart from the very first one) can be viewed simultaneously as a cause of the next event and a consequence of the previous one.

Discussion

Is there any objective way to measure exactly how traumatic the experience has been for a given individual?

Some attempts to do so have been made. One example is the study of Luo et al. (2012) who investigated the level of **cortisol** in the hair of 64 adolescent girls who experienced the Wenchuan earthquake in China in 2008. Cortisol is a stress hormone. It seems logical to assume that chronic stress caused by trauma will be reflected in the levels of cortisol in the body.

Luo et al.'s investigation was carried out seven months after the earthquake. Hair cortisol concentrations were measured in each three-cm segment of hair sample from the scalp. This allowed the researchers to "see" back in time and "observe" how the levels of cortisol in the person had been changing. Results showed that hair cortisol can indeed be used as a biomarker of stress reactions in traumatic experiences.

Do you think this is a reliable method of measuring trauma objectively?

What potential practical applications of this discovery can you think of?

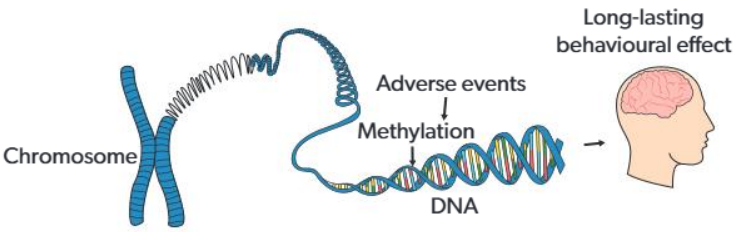
ATL Communication, Thinking, Research

Question 2: What is the mechanism underlying the effect of early childhood experiences on further development?

Another key problem in this area of research is identifying the mechanisms by which early childhood experiences translate into behavioural outcomes. We know that adverse experiences result in negative cognitive, health, and social outcomes. But why? How exactly does this effect occur?

Based on aggregated evidence from both animal and human studies, a likely mechanism underlying the effect of adverse childhood experiences on later development is **epigenetic changes**. Environmental effects induce methylation of certain sections of the DNA which suppresses the expression of genes.

A number of studies with rats have demonstrated the effects of early "mother nurturing" (mothers licking and grooming their pups in the first week of life). For example, Weaver et al.(2004) found that rats who grew up with a low-nurturing mother revealed increased methylation of the **glucocorticoid receptor (GR) gene** in the hippocampus. As a result, they showed lower expression of this gene. The GR gene is responsible for creating receptors in the brain that detect stress hormones. It was also found that these rats had increased vulnerability and stronger responses to stress when they became adults. In other words, adverse childhood experiences (not enough licking and grooming from their mother in the first week of life) permanently changed their gene expression, which continued affecting their vulnerability to stress throughout their life. For these rats, experiences in the first week have set up a life-long behavioural trajectory.



▲ Figure 3.24 Epigenetic changes

Attempts to examine the generalizability of these results from animals to humans are limited. It was originally believed that methylation of the GR gene can only be measured directly in brain tissue, and this is an invasive measurement. A notable attempt to resolve this issue is the study of McGowan et al.(2009), who analysed the brains of people who committed suicide. It turned out that those who had experienced abuse in early childhood had more methylation in the GR gene than those who had not experienced abuse. The differences in methylation were also specific to a brain region: they were observed in the hippocampus but not in other brain areas. This was similar to earlier results with rats, suggesting that the underlying physiological mechanism may be the same.

Some subsequent research showed that the biological response to early life adversity is:

- genome-wide: that is, it is not limited to methylation of one particular gene but rather creates a whole pattern of methylation spread out across multiple genes

Exam tip

Epigenetic changes are explained in more detail in Unit 4.6: Environmental factors in mental disorders.

Here are a few key definitions:

DNA is the molecule that contains the organism's genetic code. Every protein in our body (including neurotransmitters, hormones, and their receptors in the brain) is synthesized based on the code in DNA.

Stressful environmental events cannot change the DNA itself, but they can affect **gene expression**. This is the process of translating the genetic code into the actual expressed characteristics. One of the forms this may take is **DNA methylation**: when certain molecules bind onto sections of the DNA, preventing it from being expressed.

Changes in the organism's appearance and behaviour that result from gene expression without affecting the DNA itself are called "epigenetic changes".

- system-wide: it is not limited to epigenetic changes in the brain but also occurs in other parts of the body—for example, T-cells in the blood.

The discovery that changes in DNA methylation that occur in response to adverse childhood experiences can be detected in the blood (T-cells) is "good news". Taking a blood sample is a lot easier and a lot more ethical than extracting brain tissue for analysis!

Borghol et al. (2012) (as cited in Szyf and Bick, 2013) investigated 40 adult males taken from a British nationwide survey of all infants born in the first week of March 1958. These participants were selected from opposite extremes of the socioeconomic status (SES) in early childhood. Researchers performed a genome-wide methylation analysis of DNA taken from blood samples when participants were 45 years old. Methylation status of 20,000 genes was mapped in this way. They discovered a significantly different pattern of gene methylation across the genome in individuals who experienced extreme poverty in early childhood, as compared to individuals who came from financially well-off environments. Interestingly, this pattern of methylation correlated much more with SES at birth than with SES in adulthood. This shows that adverse SES circumstances in early childhood indeed leave a molecular trace in the form of DNA methylation measurable in blood samples.

Taken together, these studies provide sufficient triangulation of evidence to suggest that various adverse childhood experiences produce epigenetic effects. Once the expression of genes is changed, this may continue influencing behavioural outcomes throughout lifespan, unless the methylation of genes responsible for such changes is somehow reversed. Whether or not such reversal is possible is still a question that requires further research. It is a very exciting research area.

Activity

Conduct some research online on the existing methods and intervention programmes that have been designed to promote resilience of children and help them recover from adverse childhood experiences.

In your group make a graphical representation (a mind map or an infographic) of the ideas and approaches you were able to find.

Use generative AI to help you to make your search more efficient.

ATL Social, Research, Self-management

Conceptual analysis

Causality

The direction of causality in this area of research is clear: adverse childhood experiences negatively impact behaviour (not the other way around). However, the exact parameters of this causal relationship are discussed and debated. Is this a one-off factor or does its influence persist throughout a lifetime? What is the mechanism underlying this influence? Can these effects be reversed?

We should also consider the role of resilience which serves as a mediating variable between adverse childhood experience and the extent to which further development gets affected.

Perspective

At first sight, it may seem like we are dealing with a purely environmental, sociocultural influence: poverty, deprivation, and traumatic experiences occur due to environmental circumstances. However, it is becoming obvious that environmental events can influence biological variables through epigenetic mechanisms (expression of genes). If this is the case, the environmental influences and biological ones are closely intertwined. We could even say that the distinction between them is somewhat artificial. Additionally, cognitive variables play a big role because adverse experiences are interpreted in a certain way. There are individual differences in how children react to the same experiences. These differences are linked to resilience, which may have a cognitive nature.

Measurement and bias

The same experience will not be equally traumatic to different individuals. Measuring how traumatic the experience has been for a particular individual may be tricky. Research studies either investigate the impact of experiences that were obviously traumatic (such as experiencing an earthquake and loss of a home) or rely on self-reporting or try to find objective proxy measures such as the amount of cortisol in the system.

Longitudinal studies are important to understand how persistent the effects are. Unfortunately, there is a lack of such studies, which leaves some questions unanswered.

Experimental research in this area is limited for obvious reasons: we cannot create traumatic experiences on purpose for the sake of research.

Change

Once the child has suffered from the effects of adverse childhood circumstances (i.e., poverty, deprivation, or trauma), is there anything we can do to help the child bounce back and recover? If we provide new positive experiences, can we expect them to reverse the effects of earlier negative ones?

The study of resilience in this area is particularly important. If we find out what factors support the resilience of a child, we can design interventions to help them.

Research has made some advancements in this area. We know that the resilience of children depends on the resilience of adults. For example, children will cope better if adults are emotionally available to them during adversity. A network of social support serves as a protective factor, and so on. In addition, epigenetic research bears the promise of identifying biochemical solutions.

Responsibility

Adverse childhood experiences are a sensitive research topic. We can do harm by making children relive those experiences. There is a reason why some researchers choose not to interview children and instead interview their parents.

However, understanding the mechanisms of early adverse experiences and the role of resilience can help us design programmes that will help future generations of children. This necessitates a careful cost–benefit analysis before any such study is conducted. Ethics committees will consider a range of factors to decide if the research proposal can be approved.

3.7 Theory of mind

Inquiry questions

- When do children acquire the ability to understand other people's beliefs and intentions?
- Do animals have a similar ability?
- How can this ability be measured?

What you will learn in this section

Key learning:

- Theory of mind is the ability to attribute mental states (beliefs, intentions, knowledge) to others. It starts developing in humans around four to five years old.
- It has been debated whether or not higher primates (such as chimpanzees) have theory of mind.
- It was observed that chimpanzees can point to an object that is appropriate to the situation (e.g., a key for a person struggling to open the door). However, the alternative explanation would be a simple association between the object and the situation in the primate's experience.
- With a specially designed rational imitation paradigm, researchers confirmed that chimpanzees can indeed understand someone else's intentions.
- It was not until the invention of eye-tracking procedures that chimpanzees were confirmed to understand false beliefs.
- Similar experimental procedures were used with children. It was shown that children as young as 18 months old can decode another person's intentions.
- The understanding of false belief is formed in children by the age of four to five, as established by using a specially designed Sally-Anne task. However, children with autism lack this ability.
- Theory of mind abilities may have a specific biological basis. In particular, fMRI studies have shown that a region in the temporo-parietal junction (TPJ) is activated when people think about the contents of other people's minds. Additionally, it has been suggested that mirror neurons play an important role in theory of mind abilities.

Key terms: theory of mind, mental states (beliefs, intentions, knowledge), false belief, rational imitation paradigm, higher primates, Sally-Anne task, autism, eye-tracking technology, biological basis, localization of function, temporo-parietal junction, mirror neurons

In a wider context

Theory of mind is a cognitive ability that allows people to reason about the contents of other people's minds. In this sense, it is an important prerequisite of social functioning.

Activity

Make a timeline of the development of theory of mind in children. As you read on, add important milestones on this timeline. You may want to use one of the many infographic tools available online.

ATL Self-management

What is theory of mind?

Theory of mind is the ability to attribute **mental states (beliefs, intentions, knowledge)** to others. In other words, it is the ability to understand another person's beliefs, intentions, and perspectives. Theory of mind is therefore a cognitive phenomenon. The name is a bit misleading perhaps due to the word "theory". It is not a scientific theory created by a scientist, but a cognitive ability that every individual has. Theory of mind is a "theory" simply because you cannot see directly into another person's mind. Therefore, whatever ideas you have about that other person's beliefs, intentions, or knowledge are your (theoretical) inferences. Theory of mind guides our expectations from other people and our prediction of their behaviour.

Theory of mind usually starts developing at pre-school age, specifically four to five years old.

Activity

Can you invent an experimental procedure that will allow you to see whether primates have a theory of mind? How will you test it? Come up with ideas in small groups and discuss them in class.

As you read on, you will come across several theory of mind studies with animals. Check whether the researchers used approaches similar to what you have suggested.

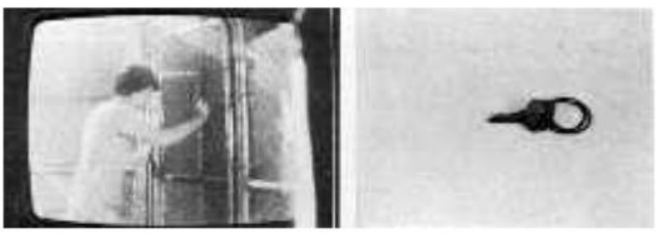
ATL Thinking, Research, Self-management

Theory of mind in higher primates

Animal studies can be useful in this area of research because the formation of a theory of mind in **higher primates** may be more easily observable. If apes do have a theory of mind and we learn how to register it, we may then apply these methods to study the formation of theory of mind in human children.

Premack and Woodruff (1978) in their seminal paper "Does the chimpanzee have a theory of mind?" argued that the answer was positive.

In this study, an adult female chimpanzee (Sarah, 14 years old) was shown videotaped scenes of a human actor struggling with a variety of problems, such as trying to reach a bundle of bananas or trying to get out of a locked cage. With each videotape, the chimpanzee was given several photographs, one of which represented a solution to the problem (such as a stick for the inaccessible bananas, a key for the lock). Sarah consistently chose the correct photograph. This led the authors to believe that she recognized the videotape as representing a problem, understood the actor's intention, and chose alternatives in accordance with that intention.



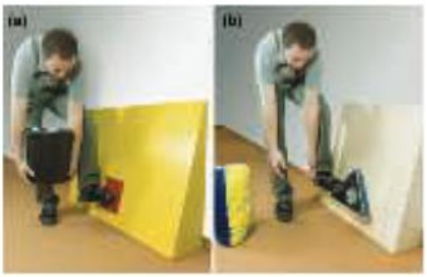
▲ Figure 3.25 One of the problems presented in Premack and Woodruff's study: a key needed to open a lock

However, these results were brought into question. Some scientists claimed that chimpanzees could predict the actions of others based on past experiences but could not understand the psychological states of others (i.e., they do not have a theory of mind). The chimpanzee (in Premack and Woodruff's study) might have chosen a key in response to a video where the actor was struggling to open the lock, not because she understood the intention of the actor to open the door but because in her past experience,

keys and locks often appeared together in the same situation. Indeed, some other evidence also supported the view that chimpanzees were unable to understand mental states of others. For example, Povinelli and Eddy (1996) showed that chimpanzees begged with equal probability from humans facing them and humans with buckets over their heads.

A way to resolve the problem is to use a research paradigm that involves unsuccessful attempts of the actor in novel situations (that the primate has no prior experience with). Call and Tomasello (2008) reviewed 10 studies that used this research paradigm. They came to the conclusion that chimpanzees, like humans, understand the goals and intentions of others.

For example, in one of the studies (Buttelmann et al., 2007) six chimpanzees were tested in the so-called **rational imitation paradigm**. In this paradigm, there is an "interesting" apparatus that produces light or sound when turned on. The human first demonstrates how to operate the apparatus and then the chimpanzee is given a turn. The human never uses their hands, though, and switches the apparatus on with an unusual body part such as their leg or head. In some cases the human has to use the unusual body part due to a constraint—for example, they are holding a heavy bucket with both hands. However, in some cases there are no constraints, and the human simply chooses to use the unusual body part. Results showed that chimpanzees use their hands when the role model's behaviour is constrained, which means that they understand the intention behind the action and simply perform the same action in a more accessible way. When there was no rational explanation why, for example, the human turned on the switch with his leg, chimpanzees imitated the behaviour and used their leg, too.



▲ Figure 3.26 The experimenter turning on the light with his foot because his hands were full, then because he chose to

The use of the rational imitation paradigm revealed that higher primates can indeed understand others' intentions. On the other hand, there was no experimental evidence in the studies reviewed by Call and Tomasello (2008) that chimpanzees understand false beliefs (which are covered later in this section). They concluded that chimpanzees do have a theory of mind, but it is not quite as developed as that of human children, mainly due to a lack of understanding of false beliefs.

Activity

Does a chimpanzee have a theory of mind? Write an answer to this question, using no more than 100 words.

ATL Communication, Self-management

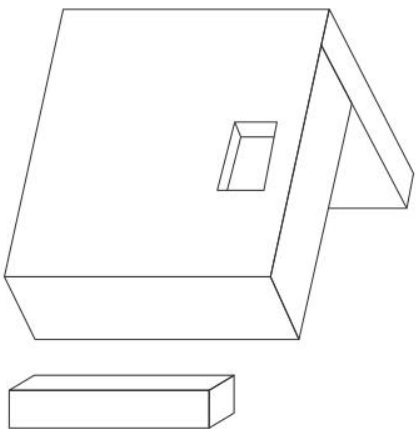
Theory of mind in human children

Understanding intentions

Research with chimpanzees and pre-verbal human children is similar in the way that it is organized, and it yields similar results.

To establish the earliest pre-verbal age when the ability to understand the intentions of others is formed, Meltzoff (1995) investigated whether children would re-enact what an adult actually did or what the adult intended to do—much like Buttelmann et al. (2007) in their chimpanzee study.

- Children in the control group were shown an adult who successfully completed a target act. For example, the adult demonstrated how they used a small rectangular block of wood to push a slightly recessed rectangular button in a wooden box. Pushing the button activated a buzzer.
- Children in the experimental group were shown an adult who tried but failed to perform a target act. All movements and actions were exactly the same, with the exception that the experimenter would try to push the button but miss, and there would be no activation of the buzzer.



▲ Figure 3.27 The equipment used in Meltzoff (1995)

Results showed that children as young as 18 months old could infer the adult's intention. There was no difference in the number of successfully accomplished imitation acts between the two groups. Infants in the experimental group did not go through a period of trial-and-error with the test objects but directly produced the target act, just as those who observed the complete act. This shows that infants were able to "read" the adult's intentions.

An interesting question arises: Is this a response to the physics of the situation (pattern of movements) or is there a psychological interpretation involved?

The experiment was replicated, but this time the same movements were produced by an inanimate device. The device closely mimicked the movements of the actor—both the completed action and the failed action condition. Results showed that infants do not read an intention behind the movements of a mechanical device. They were six times more likely to produce the target act after seeing the human attempt than after seeing the same demonstration by a machine. Therefore, there is a psychological interpretation involved, after all.

Understanding false beliefs

It is very common for people to hold **false beliefs**—beliefs that are not true. When you know the truth, it requires a theory of mind ability to understand that others may not know the truth as you know it, and that they may act on their (incorrect) belief rather than the true information.

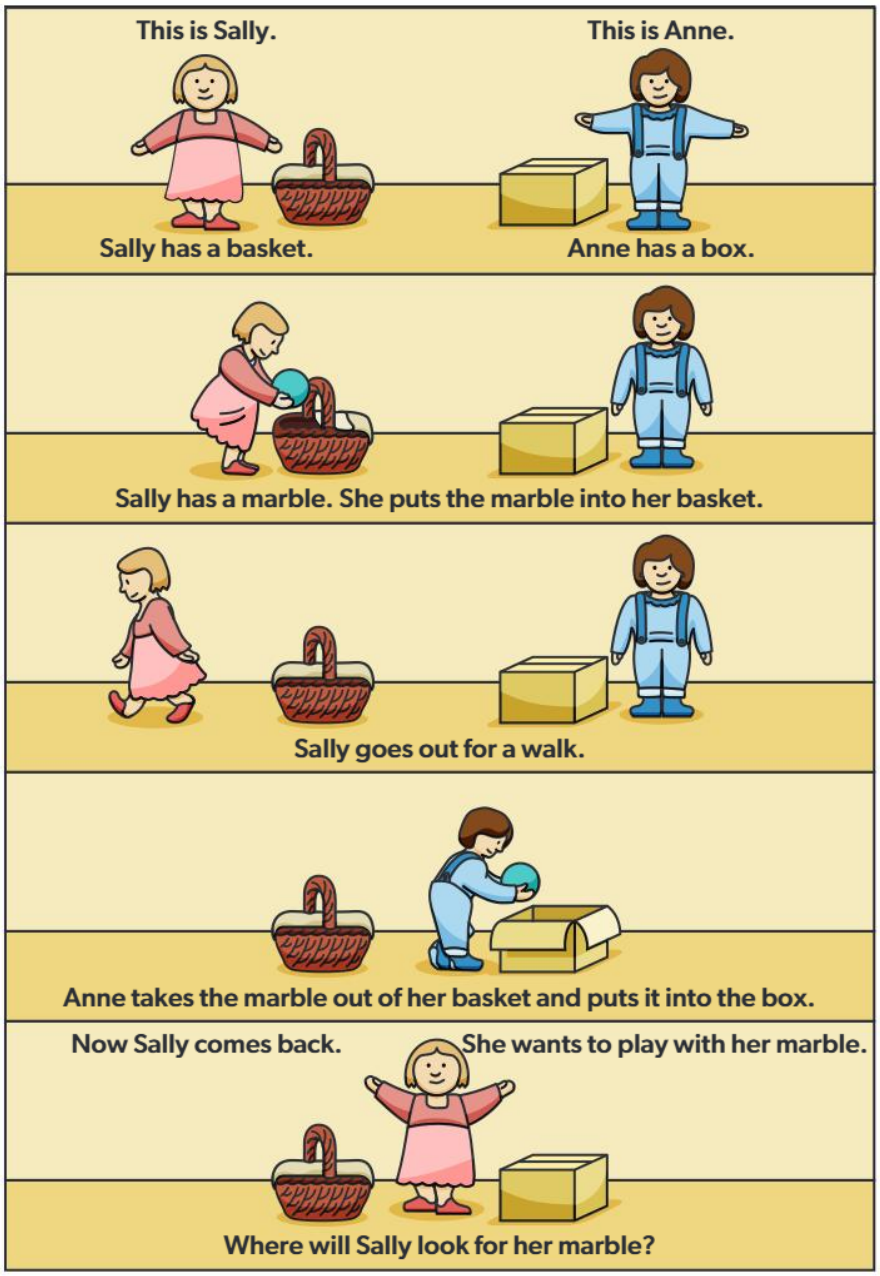
One of the most commonly used tasks designed for false-belief experiments is the **Sally-Anne task**. In this experiment, the child is shown two dolls, Sally and Anne.

- Sally has a basket and Anne has a box.
- Sally takes a marble, places it in her basket and leaves the room.
- Anne takes the marble from the basket and puts it in her box.
- When Sally returns, the child is asked where Sally will be looking for her marble.

To pass the task, it is necessary to understand that another person can have beliefs different from your own, that the person will base their behaviour on those beliefs, and that those beliefs might be false. We conclude that the false-belief task is passed if the child says that Sally will be looking for her marble in the basket, and we conclude that the task is failed if the child points to the box.

Baron-Cohen, Leslie, and Frith (1985) were the first researchers to use this procedure. Their study involved clinically normal pre-school children (mean age 4.5 years) as well as children with Down syndrome and children with **autism**. In the Sally-Anne task, 85% of Down syndrome children and 86% of clinically normal children passed the test, whereas 80% of those with autism failed it. The researchers explain the result in terms of children with autism being unable to represent mental states, which is a disadvantage when trying to predict the behaviour of other people.

It is clear that performance on the false-belief test is a critical milestone in the development of theory of mind, and children normally acquire this ability by the age of four to five years.



▲ Figure 3.28 The Sally-Anne task

Recent research with the use of eye-tracking technology

We should always remember that research findings are inseparable from the methods used to obtain them, and there is always a chance that conclusions are contaminated by confounding variables inherent in the method. This is especially true when our repertoire of methods is limited, as is the case with young pre-verbal children and primates.

As new technology started to make its way slowly to the study of theory of mind, conflicting findings emerged. For example, Krupenye et al. (2016)

Activity

As Baron-Cohen, Leslie and Frith (1985) demonstrated, children with autism have difficulty understanding other people's beliefs and intentions. As they do not fully understand that other people's perceptions and thoughts are different from their own, autistic individuals experience problems with social relationships and communication. For example, they may not be able to predict correctly what others will say or do in a particular situation. From the practical viewpoint, the important question is: how can we teach theory of mind skills to children with autism?

One of the methods used in this area is "Social stories", developed by Carol Gray. "Social stories" are scenarios that allow people with autism to understand others better and motivate them to start asking questions about others.

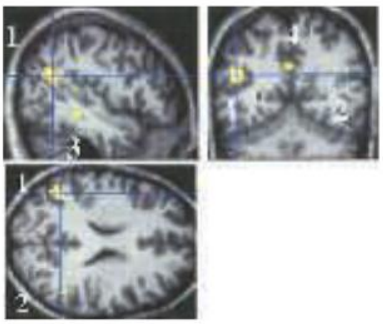
To find out more, explore Carol Gray's website. In particular, watch her introductory presentation.

ATL Research, Self-management

Chat with AI

Use your favourite generative AI to help you explore this topic from the conceptual lens. Here are some examples of prompt fragments that you might find useful:

- I am an introductory psychology student studying the development of theory of mind. I am trying to analyse this phenomenon through the lens of the concept of causality. Could you give me some pointers about how I can approach answering the following questions: "What is the development of theory of mind caused by?" and "What does theory of mind cause?" Are there any other questions I could ask that potentially connect to the concept of causality?
- To continue the same theme, I am now trying to look at theory of mind through the conceptual lens of "perspective". The question I came up with is: "Are there any competing theories or perspectives in psychology on the nature of theory of mind?" Could you give me some starting information that would help me answer this question? Are there any other relevant questions related to the concept of perspective that you suggest asking?



▲ Figure 3.29 Example fMRI scan from the study by Saxe and Kanwisher (2003)

obtained evidence that seems to suggest that great apes do understand false beliefs, after all. In this study, apes were shown videos of an actor dressed in a King Kong suit. This was to increase the ecological validity of the experiment by making the procedure more interesting or engaging for the apes. In the video, the King Kong actor hits a man and then darts under one of two haystacks while the man is looking. The human then disappears behind a door and the ape either switches haystacks or stays under the same haystack (depending on the experimental condition). When the man returns holding a long pole, he smacks the haystack he thinks his opponent is under. Using **eye-tracking technology**, researchers showed that apes anticipated the human's behaviour based on his beliefs. You can find the video of this study by searching "Krupenye 2016".

Prior to the next example, remember that the findings of Baron-Cohen, Leslie and Frith (1985) showed that 80% of autistic children failed to cope with the Sally-Anne task. This makes the 20% who passed the false-belief test an interesting group of subjects—if they have autism, why did they answer correctly?

To investigate this further, Ruffman, Garnham, and Rideout (2001) used the Sally-Anne task coupled with a measure of eye gaze (where children looked when anticipating the return of Sally) as well as a traditional verbal measure (a direct question). They found that eye gaze was better than verbal performance at differentiating children with autism from other children. Children with autism did not look to the correct location in anticipation of the story character's return, even if they answered the false-belief question correctly verbally. This clarifies the unexplained 20%. The occasional correct responses to the false-belief question have something to do with general linguistic ability, whereas eye tracking taps into more fundamental abilities implicated in theory of mind. Even if they give the correct response, their eyes are looking in the incorrect direction.

Localization of theory of mind

Does theory of mind have a **biological basis**? As we know, many cognitive processes and abilities are **localized** in specific brain areas, some more obviously than others. Being a highly specific ability, theory of mind could be expected to reside in a well-defined area of the cortex.

To find out whether this was true, Saxe and Kanwisher (2003) used false-belief stories to compare reasoning about true and false beliefs to reasoning about non-social control situations. Participants had to read short stories, some of which involved false beliefs, and some did not. The researchers carried out fMRI scans and these revealed that one particular region in the **temporo-parietal junction** (TPJ) was involved specifically in reasoning about the contents of another person's mind. The researchers concluded that TPJ was responsible for theory of mind at least in terms of one of its components—understanding the beliefs of others.

Empathy and theory of mind have also been explained with reference to **mirror neurons**. Mirror neurons were originally discovered in the motor cortex. These are neurons that are activated both when you perform an action and when you watch someone else perform the same action.

Discussion

V. Ramachandran, a famous neuroscientist, claimed that the existence of mirror neurons means that all humans are fundamentally connected. To learn more, watch and discuss his TED Talk. Search online for "The neurons that shaped civilization".

ATL Social, Self-management

Mirror neurons were discovered by Giacomo Rizzolatti and his colleagues at the University of Parma (Italy) in the 1980s. Later Iacoboni et al. (2005) investigated the role of mirror neurons in performing tasks that involve a theory of mind. They found greater activation in mirror neuron areas in the inferior frontal cortex when we observe situations where another person performs an action with a clear intention (as opposed to simply observing actions taken out of context).

Conceptual analysis

Causality

Theory of mind is a specific cognitive ability that is unique to humans (although higher primates also have some elements of it). It has quite a specific localization in the brain. We could probably explain the emergence of theory of mind with an evolutionary line of reasoning: it provided a survival value in the life of early humans.

On the other hand, theory of mind affects a range of other social skills and functions. This is clearly seen in people diagnosed with autism spectrum disorder (ASD).

Perspective

Theory of mind is a cognitive ability, but it enables effective social interaction. It also has a specific localization in the brain. Therefore, once again, we are dealing with a phenomenon that exists in all three dimensions: biological, cognitive, sociocultural. If theory of mind has a biological basis, then animal studies may inform our understanding of theory of mind in humans.

Bias

We have seen from some of the research presented in this section that sometimes limitations in research procedures lead to incorrect or biased conclusions. For example, Premack and Woodruff (1978) jumped to the conclusion that chimpanzees could understand mental states (intentions) of others, but there existed a possible alternative explanation.

Carefully controlled experiments, the use of creative research procedures (e.g., the rational imitation paradigm), and a series of manipulations (e.g., a human actor versus a robot) allow us to eliminate potential competing explanations and gradually establish the truth.

Exam tip

Questions in Paper 1 Section C will combine one of the concepts (e.g., causality) with one of the content points within one of the contexts (e.g., theory of mind within human development). Remember that there are many ways to ask these questions. Trying to predict all possibilities and preparing answers in advance is *not* a productive approach. At the same time, there are multiple ways in which concepts could be applied to content, and there are no prescribed "correct" ways to do so.

The "Conceptual analysis" section at the end of this and all other units gives you some suggested ways in which concepts can be linked to content. However, this is not a mandated (required) way of making the link. Other links will be accepted as long as they are relevant and justified. As you study the material, you will eventually achieve a level of understanding which should allow you to "improvise" on exam day, revealing some new links between concepts and content that you had not considered before.

Measurement

In this section, we have seen yet another example of how our findings and conclusions may depend on the methods of measurement that we use. When experiments were repeated with the use of eye-tracking technology instead of verbal responses, they yielded different results. Measurements may be affected by all sorts of confounding factors, and we need to remember it when interpreting the results of a study.

Additionally, we have seen here how the measurement of some psychological phenomena (e.g., theory of mind) depends on the creation of a special measurement procedure—for example, the rational imitation paradigm or the Sally-Anne task.

Change

There are many possible links between the content in this section and the concept of change. First, one could talk about how this ability develops with age (children normally acquire theory of mind around the age of four to five years old). Second, one could look at evolutionary development across species: higher primates seem to have at least some elements of theory of mind, but it is not fully formed. Third, we could ask ourselves if theory of mind abilities can be influenced: trained for faster development or restored in children with autism.

Responsibility

Theory of mind is a high-order cognitive function. Most of the studies considered in this section were not invasive and required participants (children or chimpanzees) to make some simple decisions in game-like scenarios. It does not mean that there are no ethical considerations to be observed. For example, chimpanzees bred in a laboratory should be treated in a special way to ensure that their distress is minimized.

Additionally, one could ask if it is the responsibility of researchers to direct efforts and resources towards studies that can potentially have practical applications. For example, you could try to find out if chimpanzees understand false beliefs, or you could try to find a remedy for autism in children. Is it more ethical to choose the latter?

3.8 Attachment

Inquiry questions

- What is the nature of children's attachment to their parents?
- Is attachment to caretakers vital for further development?
- What happens if the parents do not reciprocate a child's attachment?

What you will learn in this section

Key learning:

- Attachment is an important stage that provides a basis for further development. Its disruption leads to negative long-term consequences.
- Initially, it was believed that attachment can be reduced to satisfaction of basic needs such as hunger. However, Harlow's research with rhesus monkeys has demonstrated that contact comfort is more important for attachment than basic physiological needs. It also showed that attachment provides a "secure base" and enables normal cognitive development.
- It was later suggested by John Bowlby that attachment is a system consisting of both biological factors (the attachment behavioural system) and cognitive factors (the internal working model).
- There are important individual differences in attachment that manifest themselves in attachment styles: secure, avoidant, ambivalent.
- Attachment styles demonstrate cross-cultural variability. However, within-culture variability depending on social factors may be even larger.
- Attachment practices are not culturally universal, which is demonstrated by emic research.
- Attachment patterns developed in early childhood continue influencing our behaviour throughout the course of life and may even affect patterns of romantic relationships in adulthood.

Key terms: attachment, separation distress, contact comfort hypothesis, secure base hypothesis, attachment behavioural system, internal working model, attachment styles, Type A (insecure avoidant) attachment, Type B (secure) attachment, Type C (insecure ambivalent/resistant) attachment, strange situation paradigm, intracultural variation, caretaking practices, romantic relationships

In a wider context

Attachment is a special complex of behaviours that emerges early in childhood and has a profound effect on subsequent development. In some ways, attachment in humans is similar to attachment behaviours observed in many animals. However, this cannot be reduced to a simple instinct. There are large cognitive and cultural components to it.

Activity

Before we go any further, make sure that you remember these concepts and can clearly explain their meaning: environment, behaviourism, learning.

ATL Communication, Self-management

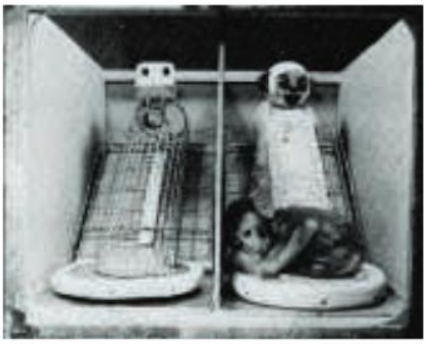
Discussion

We know that animal research may inform our understanding of human behaviour. The rationale for this is that we share the same biological structure and genes with many animals, so presumably we are quite similar.

However, how do we establish a "threshold" of similarity? To what extent exactly, and in what aspects, should an animal be similar to human beings for us to assume that a study of this animal would inform our understanding of human behaviour?

Was Harlow justified in his selection of research subjects?

ATL Thinking, Research



▲ Figure 3.30 Wire and cloth mother surrogates

What is attachment?

Attachment as a concept refers to an emotional connection between individuals that is most clearly seen in one individual's reaction to separation from the other (**separation distress**). Studies on animals and on humans suggest that attachment of infants to their caregivers is an important stage that provides a basis for further development.

Harry Harlow: contact comfort and secure base

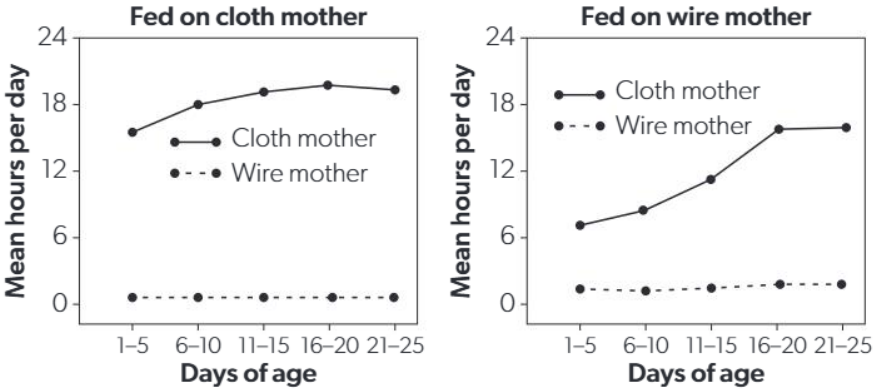
In his article "The nature of love" (1958), Harlow set out to explore the effects of early childhood experiences with one's mother (or a primary caregiver) on later development. At that time, it was widely believed that attachment was secondary to satisfaction of basic needs such as thirst and hunger. This went well together with the behaviourist stance: since the mother is the source of food, attachment is just a behavioural response to anticipation of food. Harlow criticized this belief.

Harlow used rhesus monkeys in his research. Compared to human babies, rhesus monkeys are more mature at birth and grow more rapidly, but the basic behaviours, such as nursing and clinging, as well as basic cognitive phenomena, such as perception, memory, and learning capability, are very similar.

An artificial surrogate monkey mother was built from a block of wood and cotton cloth. A light bulb that radiated heat was placed behind her, and a milk-dispensing mechanism was installed in the breast area. This meant that this surrogate mother provided food, warmth, and comfort whenever the baby needed it—a mother with limitless patience that never scolded her infant. In Harlow's opinion, they "engineered a very superior monkey mother, although this position is not held universally by the monkey fathers" (Harlow, 1958).

The second type of surrogate mother that he constructed was much less comforting. She had the same milk-dispensing device and a heat radiator, but she was made of wire—in other words, similar to the first condition in every way except for "**contact comfort**". She could satisfy all the basic needs, but an infant monkey could not cling to her. Harlow called this surrogate mother monkey "biologically adequate, but psychologically inept" (Harlow, 1958).

In the first experiment, infant rhesus monkeys were placed in a cage adjacent to both the cloth and the wire mother. They could freely choose who they wanted to "spend time with". For four infant monkeys, the cloth mother lactated, and the wire mother did not. For the other four monkeys, the reverse was true. Results showed that baby monkeys preferred to spend time with the cloth mother (maximizing contact comfort rather than satisfaction of basic needs). Even in the condition where the wire mother was the one providing food, by the age of 20 days, monkeys learned to use the wire mother only to get milk and spent increasing amounts of time with the cloth mother. This was completely contrary to the traditional belief that attachment was contingent on hunger and thirst reduction.



▲ Figure 3.31 Time spent on the cloth and the wire mother surrogates

In another experiment with the same subjects, Harlow also tested the "**secure base**" hypothesis. Harlow placed the monkeys in a small room containing multiple stimuli that monkeys usually like to play with (e.g., wooden blocks and blankets). This was called the "open-field test". On some trials, monkeys were in the room alone. On other trials, either a wire or a cloth mother was placed with them. Results showed that baby rhesus monkeys did indeed use the cloth mother as a secure base—when the cloth mother was present, they would rush to her, clutching her and manipulating her body and face. After a while, they would leave her to explore and manipulate one of the toys, then rush back to her again, and so on. In contrast, in the conditions with a wire mother, and no mother at all, monkeys' behaviour was much more anxious—frequently they would freeze and stay motionless in the corner in a crouching position, sometimes screaming or crying, and nothing could calm them down.

These results suggest that properly formed attachment may serve as the basis of exploration of the environment and so cognitive growth.

An obvious point of criticism for this study is its applicability to humans. It has been noted that, despite all the similarities between rhesus monkeys and human babies, attachment processes develop much more slowly in humans.

The ethical considerations of the study are also very clear. Infant monkeys were intentionally raised in conditions that deprived them of appropriate mother–infant contact and potentially disrupted their cognitive development. At the same time, one needs to take into consideration the potential benefits of such studies to people and society. For example, it has changed the way we understand "basic care" provided at an orphanage. It has been acknowledged that human contact in infancy is an essential component of successful development.

Psychological theories of attachment

In a series of publications from 1958–1960, John Bowlby proposed a theory of attachment that combined biological and cognitive factors (Bowlby, 1958; Bowlby, 1959). Attachment in this theory includes two aspects: the **attachment behavioural system** and the **internal working model**.

The attachment behavioural system is an instinctive sequence of behaviours triggered by certain environmental cues. Therefore, it is biologically based. When an individual senses danger, the corresponding activation of the

Activity

What are the ethical considerations involved in Harlow's study? If you were on an ethics committee and Harlow presented his research proposal to you, would you approve it? What questions would you ask?

Role-play an interview of Harry Harlow by an ethics panel.

ATL Communication, Social, Self-management



▲ Figure 3.32 Response to the cloth mother in the open-field test

attachment behavioural system is called "alarm". When an individual experiences an anticipation of fear or discomfort, it triggers another mode of the system: "anxiety". Both alarm and anxiety are characterized by the infant's seeking and maintaining physical proximity with the caregiver.

The internal working model includes beliefs about attachment figures, beliefs about the self, and beliefs about one's relationships with others (i.e., various mental representations). It comprises the cognitive aspect of Bowlby's theory. Suppose a child is being neglected and denied attention, even when they seek proximity and express distress. If this neglect is consistent, they may develop an internal working model in which they are not "worthy" of attention. On the other hand, excessive attention and catering to the child's needs (when parents anticipate distress even before it is expressed) may lead to an internal working model of relationships that includes unrealistic expectations about other people's responsiveness.

Attachment styles

Bowlby did not conduct any experimental studies of his own.

Mary Ainsworth, his student, developed a method for assessing attachment in infants known as the "**strange situation paradigm**" and used it in her research first in Uganda and later in the USA. In this laboratory procedure, the mother and the infant are placed in an unfamiliar playroom with toys. What follows is a scripted interaction of the infant with the mother as well as a stranger (unfamiliar person). This interaction is observed by the researchers through a one-way mirror. The script usually includes eight steps lasting three minutes each.

1. Mother, baby, and the experimenter. The experimenter gives all necessary instructions.
2. Experimenter leaves. Mother with baby.
3. Mother, baby, and a stranger. The stranger enters the room, talks to the mother and approaches the baby.
4. Mother leaves. Stranger and baby (first separation).
5. Stranger leaves. Mother and baby (first reunion).
6. Mother leaves. Baby alone (second separation).
7. Stranger enters. Stranger and baby.
8. Stranger leaves. Mother enters (second reunion).

Ainsworth observed some differences in children's behaviour. The main differences were in the patterns of separation anxiety (reaction of the infant when the mother walked out of the playroom), stranger anxiety (reaction of the infant to a stranger in the playroom), and reunion behaviours (reaction to the mother when she returned to the playroom after a while). Analysis of these behaviour patterns resulted in the identification of three **attachment styles** (Ainsworth et al., 1978).

In **Type A (avoidant attachment)**, observed in 20% of infants, the infant did not show any signs of distress when the mother left the playroom. Neither did the infant show any anxiety in the presence of the stranger—the child would continue playing as usual. These children's reunion behaviour could be characterized as indifferent.

Activity

There are several videos on the internet that you can watch to get a better understanding of the procedures used in Ainsworth's "strange situation paradigm". Some of them document classical studies, and some document more modern re-enactments.

ATL Self-management

In **Type B (secure attachment)**, observed in 70% of infants, the infant showed separation anxiety every time the mother left. The infant avoided the stranger when alone, but acted in a friendly way towards the stranger when the mother was in close proximity. The infant showed a clear positive reaction to the reunion when the mother returned after a short absence.

In **Type C (ambivalent/resistant attachment)**, observed in 10% of infants, the infant showed very intense separation anxiety, but at the same time when the mother returned the infant would resist contact and even push her away. The stranger was avoided at all times, whether the mother was in close proximity or not.

Cultural variations in attachment

Cultural variations are especially interesting because they allow us to understand whether attachment is a predominantly biological or predominantly social phenomenon (which links to the nature–nurture debate).

Van Ijzendoorn and Kroonenberg (1988) carried out a cross-cultural meta-analysis of studies using the "strange situation paradigm" in eight countries. The key focus of the study was the "strange situation" classifications—that is, what percentage of children in this or that country tend to be classified as having secure, avoidant, or ambivalent attachment.

A total of 32 studies from eight countries were selected, with a total of 1,990 participants. Of those samples, 18 were from the USA.

An aggregated average across all 18 US samples did broadly match Ainsworth's findings (A = 20%, B = 70%, C = 10%), but this was only true for aggregated data. There was considerable variation between the samples even within the same country. The most striking finding of the study was that **intracultural variation** (that is, differences in "strange situation" classifications in different samples in the same country) was nearly 1.5 times as large as cross-cultural variation (the difference between countries). This variation could possibly be attributed to the socio-economic status of the family. For example, middle-class professional families seemed to have higher rates of secure attachment as compared to lower-class families, in which higher rates of avoidant and especially ambivalent attachment styles were registered.

While most cultures demonstrate similarities rather than differences in terms of attachment styles, some cultures show striking deviations from the average "standards". Anthropological research of such cultures sheds light on the role of social factors in the development of attachment behaviour. An example of this is the case study of **caretaking practices** among the Efe conducted by Tronick, Morelli, and Winn (1987).

The Efe (pronounced as *Ef-fay*) are an ethnic group in the Democratic Republic of Congo. They inhabit part of the Ituri Forest and are a society of hunters and gatherers. They are semi-nomadic, changing location once a month. It is not likely to see an Efe engaged in a solitary task as most of their activities are shared because of their way of life. When the camp is moved to a new place, they clear a small area in the forest and build huts. Due to the temporary nature of the dwelling, as well as limited space when it comes to food storage or protection from rain, there are practically no physical barriers for this group of (usually) 6–50 residents, so they lead very public lives.

Activity

If you were to replicate a study using the "strange situation paradigm", what would you do to ensure that the study was conducted ethically? Think about the fact that you are creating distress in very young children— first by separating them from their mothers and then by exposing them to a stranger.

ATL Research

Discussion

Based on the studies reviewed so far, would you say that attachment is predominantly biological, psychological, or sociocultural? What are your arguments?

ATL Communication, Thinking



▲ Figure 3.33 Children of the Ituri Forest

Tronick, Morelli, and Winn (1987) collected information using interviews and observation. Rich data collected in the course of this study revealed a lot of interesting facts, some of which are summarized here.

- Most female camp members attend the birth.
- Immediately after birth, the infant is passed among the group of women and sometimes suckled by them whether or not they are lactating.
- The Efe believe that a mother's milk lacks nutritional value. This is why other women suckle the baby. If no female in the camp is lactating, a woman from another camp is recruited to suckle the baby.
- The percentage of time the baby spends in physical contact with people other than the mother gradually increases and reaches 60% by the time the baby is 18 weeks old.
- At 18 weeks old, infants are transferred on average eight times per hour.
- Each infant is cared for by an average of 14 people.
- The Efe are sensitive to their infants and quickly respond to fussing or crying. Observations showed that infants were usually comforted within 10 seconds from the onset of any fuss or crying.

All these findings suggest that caretaking practices are culturally variable and so attachment behaviour is, to some extent, culturally shaped.

Discussion

Do you think a relationship with a romantic partner is determined to a large extent by the patterns of interaction with your primary caregiver in infancy? Can you change these patterns of attachment later in life through conscious effort?

ATL Communication, Thinking, Self-management

Attachment in later life

Attachment in adults manifests itself in the form of **romantic relationships**. Shaver and Hazan (1988) found that adult patterns of attachment (e.g., to romantic partners) closely resemble those that they had with their caregivers in the past. In a self-report measure, adults who described their attachment style in romantic relationships as secure also reported a secure style of attachment to their parents in childhood. The same correspondence between adult and child attachment patterns was established for avoidant and ambivalent styles. This shows how attachment in early childhood may have an influence on patterns of adult romantic relationships.

Conceptual analysis

Perspective

Like many other phenomena in psychology, attachment can be considered from the biological, cognitive, and sociocultural perspective. It is a combination of all three of them that provides the deepest understanding of its nature and its role in human development. Early theories considered attachment as a purely biological phenomenon driven by satisfaction of basic needs. It was then demonstrated that there is more to attachment than that. To explain observed attachment behaviours, psychologists proposed cognitive constructs such as secure base, the internal working model, and so on. Later still, cultural variations in attachment behaviours and attachment styles were investigated, showing that attachment is not entirely universal across cultures and social groups.

Causality

Related to causality is the discussion of the key factors that cause attachment behaviours: satisfaction of basic needs, contact comfort, secure base, the internal working model. We have shown in this section how scientists' views on this have gradually transformed.

It is also important to note that there are individual differences in attachment that manifest themselves in attachment styles. This means that the influence of key factors (such as contact comfort) on attachment behaviours is mediated by some other variables. For example, the style of parenting may determine how exactly attachment will develop in a child.

Finally, attachment itself is a causal factor in a variety of animal and human behaviours. We have considered such direct and indirect consequences of attachment as cognitive development, anxiety, and even romantic relationships in adulthood.

Bias

Attachment has been explored using a variety of different methods and there are specific sources of bias associated with using each of them. For example, Harlow's research with rhesus monkeys may be questioned on the basis of generalizability to humans. Ainsworth heavily relied on an observational procedure called the strange situation paradigm. Any deviation from the scripted procedure may result in changes in the child's behaviour. Cross-cultural research has shown that there may be more variation within one culture than among different cultures. Not considering potential mediators (e.g., SES) could result in a biased understanding of the role of culture in attachment.

Culture was a relatively recent addition to the study of attachment. When this research began, it was focused solely on Western societies, so the classic theories of attachment may have an inherent sampling bias.

Measurement

Attachment is unobservable. We infer it from such behaviours as separation anxiety. Of course, separation anxiety may be caused by some factors other than attachment. For example, if the child becomes upset as the mother leaves the room in the strange situation paradigm, this may be due to a general anxiety in situations of uncertainty, and not necessarily being attached to the mother.

Cross-cultural studies by definition cannot be experimental because culture cannot be manipulated. Animal studies provide much better control over variables and allow hypotheses to be tested, but generalizability to humans and ethics are two big concerns in such studies.

Observational studies such as the ones based on the strange situation paradigm depend on a rigorous scoring procedure. Several independent observers are required to make sure that they converge in the way they interpret children's behaviour in the observed situations.

Change

We could look at the concept of change in relation to attachment by asking about the role attachment plays in human development. This is especially obvious when we consider individuals who have been unable to develop a healthy attachment to a parental figure (e.g., monkeys in Harlow's studies, or children who grew up in harsh conditions in orphanages). Studies of attachment and romantic relationships suggest that attachment patterns formed in early years can have a very long-term impact on our lives.

We could also look at it from the point of view of cultural variation (how attachment behaviours differ from culture to culture). This includes such cultural "outliers" as attachment practices among the Efe.

Another angle in this relationship is individual differences and attachment styles. Depending on individual circumstances, the attachment system may develop in very different ways.

Responsibility

Harlow's research clearly brings out ethical issues associated with the study of attachment (and animal studies). If factors of attachment are experimentally manipulated, it can lead to long-lasting irreversible deeply negative effects on the individual's life. Arguably, the treatment of animals in these studies violated ethical standards. However, Harlow's studies were instrumental in changing a whole range of social practices. For example, there used to be no requirement in orphanages to maintain human contact—infants were given milk and left to their own devices. After Harlow, institutionalized care was changed, so one may say that this research has positively affected millions of human lives.

On the one hand, we are responsible for conducting research ethically. On the other hand, we are responsible for ensuring that the way we bring up our children provides the best possible outcomes for them, in an evidence-based way.

Exam-style practice questions

Paper 1 Section A (4 marks)

Explain neuroplasticity with reference to one example.

Paper 1 Section B (6 marks)

At a scientific conference on cross-cultural psychology, a presenter observes that smiling at strangers and making small talk with them is viewed differently in different cultural contexts. For example, suppose you are in a supermarket buying something and the shopkeeper smiles at you and asks how your day has been. In North America, this may be considered polite and may even be an expected way to behave in a social situation. However, in some places in Eastern Europe, the same behaviour would be perceived as strange or even inappropriate.

Assuming that this observation is true, use the idea of enculturation to explain this difference in behaviour.

Paper 1 Section C (15 marks)

To what extent can causality be established in peer influence on human development?

(Concept: Causality. Content: Peer influence. Context: Human development)



4 Health and well-being

Introduction

Of all areas of research in psychology, mental health is the one that probably has the most immediate practical applications. Mental health problems can incapacitate us and make it impossible for us to function normally, and mental health professionals can help. For many students, mental health also serves as an entry point into psychology.

Historically, ideas of mental health have undergone some considerable transformations:

- Initially, the focus of study was on mental disorders. By analogy with medicine that studied diseases of the body, psychology studied diseases of the mind. Sigmund Freud (1856–1939) believed that the human psyche is like a crystal. Its structure only becomes apparent when it is broken. Therefore, we will only understand the human mind if we study the different ways in which it can be broken.
- With the course of time, the focus gradually shifted towards the idea that health cannot be reduced to a simple absence of disease. This is how the World Health Organization (WHO) defines mental health today: "Mental health is a state of mental well-being that enables people to cope with the stresses of life, realize their abilities, learn well and work well, and contribute to their community" (World Health Organization, 2022). Health psychology emerged as a more holistic counterpart of abnormal psychology. It has been recognized that the study of "normality" is as important as, or even more important than, the study of abnormality.
- It was recognized that mental health is determined by an interplay of a great variety of factors. Recent theories are more holistic and consider the contribution of biological, cognitive, and sociocultural variables into mental health and related behaviours.

For mental health disorders, this chapter will focus on one example: Major Depressive Disorder (MDD), also commonly known as depression.

Depression is very common, and its prevalence is growing. Depression is also hugely misunderstood. For example, it is common to hear statements such as "oh, I'm depressed" or "this makes me depressed", where to be depressed means to feel sad. MDD is a psychiatric diagnosis with a clear set of criteria and it has nothing to do with such colloquial misuse of the concept.

To diagnose mental disorders, psychiatrists over the world rely on diagnostic criteria formulated in psychiatric manuals. The one that is commonly used in the USA and Canada is called the "Diagnostic and Statistical Manual of Mental Disorders", which is currently in its fifth edition and is commonly referred to as DSM-5. It has been developed by the American Psychiatric Association. Other manuals are used in other parts of the world. For example, ICD-10 (International Classification of Diseases) was developed by the WHO and is often preferred in European countries. The third edition of the Chinese Classification of Mental Disorders (CCMD-3) is currently used in China.

There is considerable overlap between diagnostic systems used in different countries.

Depression is categorized as a mood disorder. DSM-5 lists the following symptoms of MDD:

1. Depressed mood most of the day, nearly every day.
2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day.
3. Significant weight loss when not dieting or weight gain, or a decrease or increase in appetite nearly every day.
4. A slowing down of thought and a reduction of physical movement (observable by others, not merely subjective feelings of restlessness or being slowed down).
5. Fatigue or loss of energy nearly every day.
6. Feelings of worthlessness or excessive or inappropriate guilt nearly every day.
7. Diminished ability to think or concentrate, or indecisiveness, nearly every day.
8. Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

Importantly, to be diagnosed with depression the individual must be experiencing five or more of these symptoms during the same two-week period and at least one of the symptoms should be either (1) depressed mood or (2) loss of interest or pleasure.

The prevalence rate of depression is on the rise. In the USA, the percentage of adults who report having been diagnosed with depression at any time in their lifetime reached 29% in 2023. This is 10 percentage points higher than in 2015 (Witters, 2023).

For health problems, our focus will mainly be on social media addiction, although sometimes we will use examples related to other health problems, especially obesity. Social media addiction is not currently recognized as a mental disorder and not included in the DSM-5. However, there is another disorder that is recognized by the DSM called "behavioural addiction".

The prevalence rates of social media addiction appear to be on the rise. According to a recent study (Cheng et al., 2021), the average prevalence of social media addiction in the world is reaching 25%. The diagnosis of obesity may rest on different criteria, but the most commonly used one is the body mass index (BMI). The rates of obesity are on the rise, too.

The outline of the context "Health and well-being", as presented in the IB DP Psychology Guide, is shown below. Note: in order to make the presentation of material more efficient, the sequence of topics in this chapter is slightly reordered from the IB Guide. Remember that the headings in the Guide (such as "Health problems", "Prevention and treatment") are used for convenience *only* and will not be used in the formulation of exam questions.

Mental health disorders	Health problems	Prevention and treatment
• Biological explanations • Cognitive models • Cultural differences • Environmental factors	• Stress and health • Social learning and health • Prevalence of health problems	• Biological treatment for one disorder • Psychological treatment for one disorder • Prevention and/or treatment for one health problem

4.1 Biological explanations of mental disorders

Inquiry questions

- Are mental disorders genetically pre-determined?
- Can we fully explain mental disorders with brain chemistry?
- How do you design a study to establish if depression is genetically inherited?

What you will learn in this section

Key learning:

- To some extent, mental disorders are genetically inherited.
- Genetic inheritance of mental disorders can be investigated in studies (1) based on the principle of genetic relatedness (family studies, twin studies) or (2) using molecular genetics.
- Mental disorders can also be caused by a disruption of neurotransmission in the brain (e.g., the serotonin hypothesis).
- Research in this area is complicated by the fact that effects of neurotransmitters can be complex and indirect. They can also depend on various factors such as severity of the disorder.
- Biological factors of mental disorders do not act in isolation. Biological factors may create a predisposition that will make people more vulnerable to environmental influences.

Key terms: neurotransmission, genetic inheritance, genotype, phenotype, genetic relatedness, concordance rate, twin studies, molecular genetics, neurotransmitter, diathesis–stress model

In a wider context

When we speak about biological factors causing mental disorders, we typically mean one of the following:

- imbalance of neurotransmitters
- genetic inheritance.

The two factors are not unrelated: genetic factors may determine the number of neurotransmitter receptors and the levels of neurotransmitters in the brain.

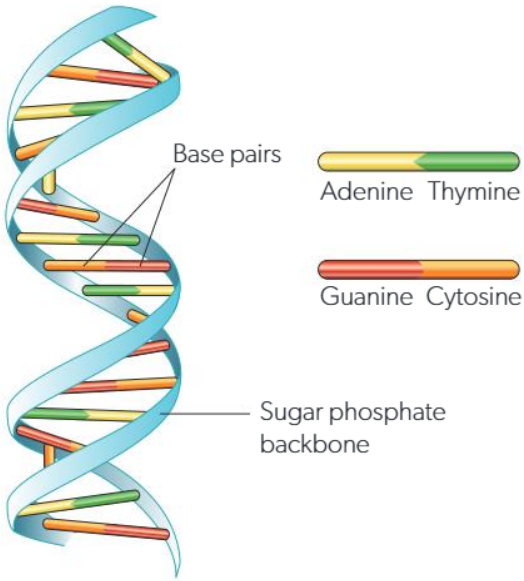
Genetic inheritance

SAQ Genetic inheritance

Genetic inheritance is the process by which genes are passed from parents to children. It also refers to the extent to which the child's behaviour is inherited from parents.

All cells in the human body that have a nucleus contain a set of chromosomes. The long DNA molecule of a chromosome is tightly coiled many times around supporting proteins, so a chromosome is a "package" that contains folded DNA.

DNA (deoxyribonucleic acid) stores information. It is a code made up of a long sequence of four chemical bases (A=adenine, G=guanine, C=cytosine, T=thymine). The bases are paired up, making a sequence of base pairs. DNA has the structure of a double helix which looks a bit like a ladder, where base pairs are the ladder's rungs. Information is coded in this sequence of bases like letters in a sentence (change the order of letters and you get a different sentence). This is a very long sentence: human DNA consists of about three billion bases.

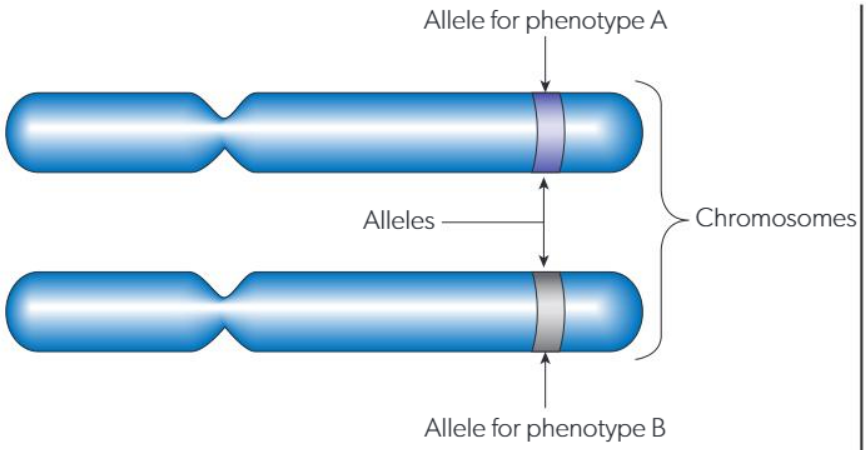


▲ Figure 4.1 DNA

If DNA is one extremely long sentence, and base pairs are letters, then genes are probably words. A gene is a unit of heredity, a region of DNA that encodes information that determines a specific trait or function. The total number of genes in the human organism is currently estimated to be around 20,000.

Humans have 23 pairs of chromosomes. One of the chromosomes in each pair is from your mother and the other one from your father. Both the chromosomes in the pair have a code for identical characteristics, but the chromosomes themselves might not be identical.

Alleles are different forms of a gene. They can be dominant or recessive. The trait controlled by the recessive allele only develops if the allele is present in both chromosomes in the pair, whereas the trait controlled by the dominant allele will develop if at least one of the chromosomes in the pair contains it. For example, in the gene that codes for eye colour, the allele for brown eyes is dominant and the allele for blue eyes is recessive. Therefore, you will have blue eyes only if both the alleles in your chromosome pair are recessive. In the other three combinations, your eyes will be brown.



▲ Figure 4.2 Gene alleles

The complete genetic sequence of alleles in an organism's DNA is called **genotype**. The set of traits that actually manifest in an individual's body, appearance, or behaviour is called **phenotype**. Phenotype comprises observable characteristics (e.g., eye colour, height, outgoing personality, vulnerability to depression, etc.) and unobservable characteristics (blood type, immune system, etc.), as well as behaviour. Genotype is the "plan" and phenotype is its implementation.

Generally speaking, there are two approaches we can use to investigate genetic heritability of mental disorders (such as Major Depressive Disorder—see Introduction to Chapter 4): studies of **genetic relatedness** and **molecular genetics**.

Studies of genetic relatedness involve comparing individuals with similar genotypes to individuals with less similar genotypes. The idea is that, if individuals who have more similar genotypes are also more similar in terms of depression, then depression is genetically inherited. We know that various groups of relatives share the following proportions of their genotypes, on average. See Table 4.1.

Relative	Amount of genotype shared
Identical (monozygotic) twins	100%
Fraternal (dizygotic) twins	50%
Siblings	50%
Child and parent	50%
Child and grandparent	25%
Children and cousins	12.5%
Parents and adopted children	0%

▲ Table 4.1 Average amount of genotype shared by various groups of relatives

We can also quantify the degree of similarity in depressive symptoms by means of a concordance study. **Concordance rate** is the probability that one person in the pair has a trait if the other person has it. For example, think about a pair of twins. If one of the twins has depression and so does the other one, they are concordant.

If one of the twins is not diagnosed with depression, but neither is the other one, they are concordant as well. However, if one of the twins is diagnosed with depression and the other one is not, they are discordant.

Twin studies are convenient for the study of genetic inheritance of behaviour because we can analyse the differences in concordance rates between monozygotic and dizygotic twins. Monozygotic twins develop from the same egg, so they share 100% of genotype. Dizygotic twins are born at the same time, but they develop from separate eggs, so they share 50% of genotype on average, just like regular siblings.

By comparing the degree of genetic relatedness to concordance rates, we can make quantitative judgements of the following type: "genetic heritability of the Major Depressive Disorder is *x* per cent".

This is exactly what early researchers attempted to do. Here are some examples:

- In a meta-analysis of twin studies by Sullivan, Neale, and Kendler (2000), genetic heritability of depression was estimated at 37%.
- Kendler et al. (2006) conducted the Swedish national twin study of major depression. The aims of the study were to compare genetic effects on MDD first in males and females and second across different generations. Depression was assessed in 42,000 twins by DSM-5 criteria in a computer-assisted telephone interview. All twins were found in the Swedish Twin Registry. Results showed that the heritability of MDD was significantly higher in women (42%) than men (29%). No evidence was found that genetic heritability of depression is different across different generations.

Another approach is molecular genetics. This became possible with the development of methods of DNA mapping and technology that allows us to create a full description of a person's genes and their variants. The technology is not cheap to use in research, but it has one clear advantage: it allows researchers to pinpoint specific genes that may be responsible for specific behaviours or symptoms.

Chat with AI

Consider using generative AI to carry out additional inquiry into the topic. Here are some sample questions and prompts that you could use:

- What are the methods used to establish genetic heritability of depression? What are the strengths and weaknesses of these methods?
- How can we calculate the coefficient of genetic heritability of depression from the results of a twin study?

Activity

You have the following three options to choose from.

1. A family study of a large representative sample of individuals where you will assess depressive symptoms together with the degree of genetic relatedness between relatives (siblings, cousins, etc.).
2. A twin study where you will compare concordance rate in monozygotic twins (who share 100% of genotype) and dizygotic twins (who share 50% of genotype).
3. A molecular genetics study where you will compare full genetic maps between a sample of people who never experienced a depressive episode and a sample of people who experienced at least four depressive episodes by the age of 35.

Your goal is to conduct a study that will provide an accurate estimate of the extent to which depression depends on genetic factors in a given population.

Which of the options will you choose and why? What considerations will you explore? Defend your choice.

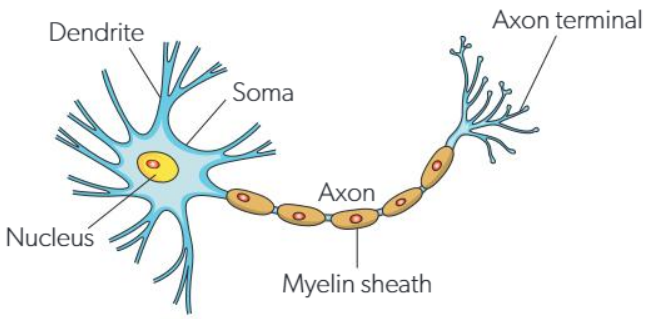
ATL Research, Thinking, Communication

Imbalance of neurotransmitters

Another popular biological explanation for mental disorders is the theory that mental disorders are caused by disruptions in the process of **neurotransmission**—for example, an imbalance of a certain **neurotransmitter** in the brain (too much or too little).

SAQ Neurotransmission

The nervous system is a system of neurons—the nerve cells. A neuron consists of three parts: the body (soma), dendrites, and the axon. Dendrites and the axon are filaments that extrude from the soma. Typically there are multiple dendrites but always a single axon. The function of dendrites (and soma) is to receive signals from other neurons, while the function of the axon is to transmit signals onwards. Where the axon of one neuron approaches a dendrite or soma of another neuron, a synapse is formed. A synapse (or a synaptic gap) is a structure that connects two neurons: the word "synapse" comes from the Greek *synapsis* meaning "conjunction". Each neuron on average has about 15,000 connections with other neurons, so it is a very elaborate network.



▲ Figure 4.3 A neuron

The nature of information transmission in the nervous system is partly electrical and partly chemical. Every neuron has a certain threshold of excitation received from the other neurons, and if the sum excitation exceeds this threshold, the neuron "fires". This means it generates a brief pulse called an action potential that travels along the axon to other neurons, passing the excitation further.

The pulse reaches the end of the axon and there, at the synaptic gap, the mechanism of transmission becomes chemical. This is how it happens:

- When the action potential reaches the end of the axon, a neurotransmitter is released from the axon terminal into the synaptic gap.
- Neurotransmitters are chemical messengers which are constantly synthesized in the neuron and moved to the axon terminal to be stored there.
- A released neurotransmitter is available in the synaptic gap for a short period.
- During this period of time, it may be destroyed (metabolized), pulled back into the pre-synaptic axon terminal through reuptake, or reach the post-synaptic membrane and bind to one of the receptors on its surface.

If the neurotransmitter binds to a receptor in the post-synaptic membrane, this process changes the membrane potential and so contributes to activating an electric pulse in the post-synaptic neuron. Here the transmission mechanism becomes electrical again.

A class of drugs known as SSRIs (selective serotonin reuptake inhibitors) do exactly what their name suggests. They selectively inhibit (block) the reuptake of the neurotransmitter serotonin from the synaptic gap. This increases the concentration of serotonin in the synapse. SSRIs have been shown to be effective against depression.

Activity

Draw a schematic representation of the process of neurotransmission.

ATL Self-management

Serotonin is perhaps the most widely discussed neurotransmitter in the context of depression studies. The "serotonin hypothesis" of clinical depression is about 60 years old.

The study of neurotransmitters as factors in developing a mental disorder is closely related to treatment. This is because the logic behind the research studies is as follows:

If we give patients a drug that is known to modify the level of a given neurotransmitter in their brain, and if this results in a reduction of symptoms, then that neurotransmitter was responsible for these symptoms.

Therefore, showing that a drug is effective for the treatment of depression is at the same time showing that depression is caused by (or at least associated with) the neurotransmitter that the drug affects. See Unit 4.2: Biological treatment for some examples of supporting studies.

What makes research in this area challenging is the complexity of the processes involved. To put it simply, there are so many things changing in the brain at the same time that it is hard to tell which of them exactly are responsible for which symptoms.

Here is a hypothetical example to illustrate this complexity.

Suppose you have a theory that serotonin (the neurotransmitter) is responsible for symptoms of depression: when the level of serotonin is low, depression sets in. To test this theory, you take a sample of mildly depressed participants and split them randomly into two groups:

1. one group receives a dose of a drug that is known to increase amounts of serotonin in the brain
2. another group receives a placebo (a similar looking harmless substance which does not have any effect).

Discussion

What might be some problems with this hypothetical study? Three such potential problems are discussed in this text, but there are more you can think of.

ATL Research, Communication

Participants take the drug or placebo once a day and fill out a questionnaire at the end of the week. Suppose you have observed that participants who took the drug demonstrated a depression score 20% lower than those who took a placebo.

Although this may seem like a straightforward experimental design, there are multiple complications and biases that may arise on the way and here are just a few of them.

- What if the effects of serotonin are mediated by something else? For example, what if the presence of serotonin in the brain triggers the production of some other neurotransmitters, or even a gradual reshaping of some neural pathways, and it is these processes, rather than serotonin itself, that influence depression? It has been observed that changes in the level of chemical messengers in the brain occur within minutes and even seconds after the intake of a drug. However, it may take weeks for symptoms of depression to begin to subside. So there clearly is something that is happening in between.
- What if the effects of serotonin are different for different levels of depression? In our hypothetical example, participants had a mild depression and the effect of serotonin intake on them appeared to be positive. But would it be different for participants with moderate or severe depression?
- What if the effects of serotonin are different in the short term and in the long term? How do we know that the improvement observed in our hypothetical study over the course of a week will not get reversed in one year's time?

Research in focus: Caspi et al. (2003)

As an example of a study of depression that used methods of molecular genetics, we will consider Caspi et al. (2003). In this study, a representative birth cohort of more than 1,000 children from New Zealand was followed longitudinally. The sample was divided into three pre-existing groups based on the variant of their 5-HTT gene:

1. Both short alleles of the 5-HTT gene
2. One short allele and one long allele
3. Both long alleles.

5-HTT is a serotonin transporter. It is a protein that transports the neurotransmitter serotonin from the synaptic gap back to the presynaptic neuron. Therefore, it enables serotonin reuptake. Many antidepressants that function as SSRIs function by targeting exactly this protein. They bind to 5-HTT, thus preventing it from "taking" serotonin from the synaptic gap and transporting it back to the neuron that released it.

The gene that encodes this protein in humans is called the SLC6A4 gene. It contains a region known as 5-HTTLPR which comes in two versions: a short allele or a long allele. That is, there are different variants in different people.

In Caspi et al.'s (2003) study, stressful life events were assessed at age 21 and 26. The measurement was performed with the "life history calendar", which focused on 14 major stressful life events in such fields as employment, finance, housing, health, and relationships.

Results showed no differences between the three groups in the number of stressful life events they experienced. However, individuals who had the short allele of 5-HTT exhibited more depressive symptoms in relation to stressful life events. More specifically, individuals who carried a short allele whose life events occurred after their 21st birthday experienced increases in their depressive symptoms from the age of 21–26 years old. In contrast, individuals carrying the long/long alleles did not (even though they experienced the same events at the same time). Among participants suffering four or more stressful life events, 33% of

Chat with AI

Here are some sample prompts you could use with generative AI to carry out additional inquiry into the topic:

- What evidence suggests that mental disorders are a result of neurotransmitter imbalance in the brain?
- Is depression caused by too much serotonin or too little serotonin?
- Could you explain the process of neurotransmission to me in simple words as if I were a 10-year-old?

individuals with a short allele of 5-HTT developed depression, compared to 17% of those having the long/long variant.

The study demonstrated that genetic set up can moderate a person's sensitivity to adverse environmental effects (life stress). This study also allowed researchers to pinpoint the specific alleles responsible for this increased vulnerability to stressful events.

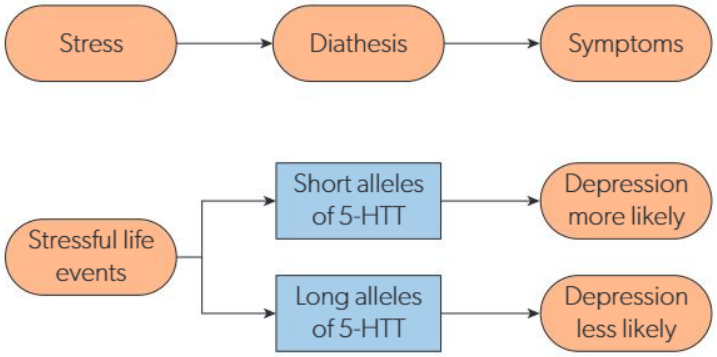
This study is interesting because it illustrates three points at the same time:

1. The method of molecular genetics.
2. The effects of neurotransmission (in this case serotonin) on mental disorders.
3. The point that biological and environmental factors of mental disorders are often combined in the so-called diathesis–stress model.

SAQ Diathesis–stress model

The **diathesis–stress model** states that biological factors create a predisposition (diathesis) for developing a mental disorder but by themselves do not directly cause it. Instead they make people more vulnerable to adverse environmental influences (stress) that trigger symptoms of the disorder. According to the model, neither of the two factors alone is sufficient for causing a mental disorder, but it is the combination of the two that makes us susceptible.

The study of Caspi et al. (2003) is an illustration of the diathesis–stress model because both genetic predisposition (the variant of the 5-HTT gene) and environmental factors (stressful life events) were measured. It was shown that stressful life events are associated with higher likelihood of developing depression, but especially so for people with vulnerable genotypes.



▲ Figure 4.4 The diathesis–stress model, as illustrated by Caspi et al. (2003)

Conceptual analysis

Causality and perspective

Causal relationships between biological factors and symptoms of mental disorders are complicated. The two main biological factors commonly considered in relation to mental health are imbalances in neurotransmission and genetic inheritance.

The problem with researching neurotransmitters is that drug intake does not just change one parameter of brain activity. It triggers a complex chain of reactions, some immediate and some delayed, some direct and some mediated. There are side effects, too.

The problem with researching genetic inheritance of mental disorders is that genetic influences and environmental influences are not independent of each other. The diathesis–stress model suggests that genetic factors may create a susceptibility to depression, but whether or not this will actually be manifested in depressive symptoms depends on environmental variables.

Measurement

Problems with causality in this area are directly translated into problems with measurement. If reaction chains involving neurotransmitters in the brain are so complex, how do we reliably measure the effect of one neurotransmitter on one behaviour? If genetic and environmental variables interact with each other, how do we separate effects of genetics from effects of the environment? See the next unit for some of the solutions proposed by researchers.

Bias

Of course this whole area of investigation may be viewed as biased because it attempts to reduce mental disorders to their biological causes (biological determinism). But as you know, this kind of reductionism is a necessary step if we want to understand the variables that cause mental health problems. It is not because researchers believe that non-biological factors play no role, it is because in an experiment, they have to isolate all variables except the one they are studying.

There are plenty of opportunities to introduce bias in research studies. Multiple things could go wrong, and failure to control even one potential confounding variable may skew the results of the whole study.

Change

The concept of change is related to treatment: by manipulating the patient's intake of chemical substances, we are trying to induce changes in their mental health symptoms. This is closely linked to measurement: we need to know if change occurs, as well as what kind of change it is—is it stable or transient, deep or superficial, quick or slow? However, it is also linked to our knowledge of what causes depression in the first place. If we know the cause, we can potentially control it. The ability to change the course of the illness is probably the force that has been driving most research attempts in this area.

Responsibility

A curious ethical conundrum results from our ability to identify genetic susceptibility to some mental disorders. We know, based on studies that employed methods of molecular genetics, that certain genes or combinations of genes are associated with a higher risk of developing mental disorders (including depression). So the question is, what do you do with that information? Suppose you are a parent and the results of the genetic test are for your newborn child. Suppose also that the disorder in question is more serious, such as schizophrenia with cognitive impairment. If you know that your child is "60% likely" to develop this disorder later in life, what will you do with that information?

4.2 Biological treatment

Inquiry questions

- Are antidepressants more effective than talk therapy?
- How do you decide to prescribe an antidepressant to a person?
- Should we trust the evidence we obtain from clinical trials of drugs?

What you will learn in this section

Key learning:

- Biological treatments are based on the assumption of biological aetiology of mental illness.
- There are different ways to measure effectiveness of antidepressants: response, remission, relapse.
- Studies show that effectiveness of medical and non-medical treatments of depression are similar over the long term, but medication takes effect more quickly.
- Randomized controlled trials (RCTs) have been criticized for their lack of ecological validity and the lengthy drug approval procedure.
- There are many potential sources of bias in RCTs (e.g., poor allocation concealment). Standards such as the CONSORT statement have been created to ensure that all research is high in quality.

Key terms: aetiology, treatment, selective serotonin reuptake inhibitors (SSRIs), randomized controlled trial (RCT), placebo group, waitlist control, response rate, remission rate, relapse rate, chemical imbalance theory, trade-off between internal and external validity, allocation concealment, unblinding, vested interests, publication bias, Phase 0, I, II, and III trials, CONSORT statement

Exam tip

This section is closely linked to the previous section (Biological explanations) because treatment relies on assumptions about aetiology. You may want to use material from both sections together to answer some exam questions.

In a wider context

The term "**aetiology**" refers to causes of a disease, or the study of such causes. When we say "aetiology of depression", we mean all the factors that cause it. **Treatment** refers to the purposeful activity of trying to eliminate the disease or at least reduce its symptoms.

Any treatment is based on assumptions about aetiology. In this section, we will consider biological treatment for mental disorders (using depression as an example).

Biological treatment of depression

Biological treatment of depression assumes that imbalance of neurotransmitters in the brain is a major factor in the development of depression. That is, it assumes that the aetiology of depression is biological. Antidepressants are drugs that target the key neurotransmitters involved in depressive symptoms.

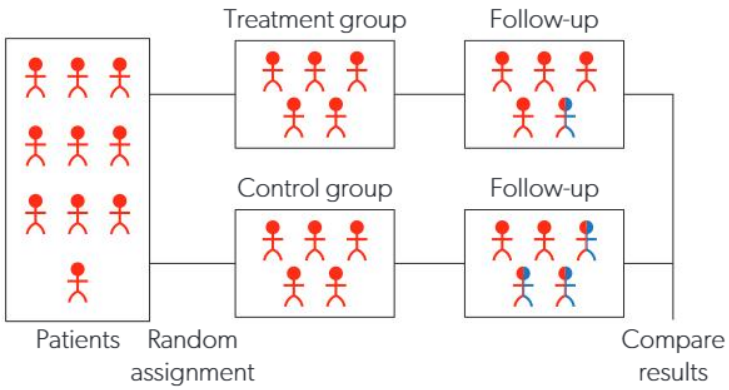
There are different classes of antidepressant drugs, such as tricyclic antidepressants (TCAs), MAO (monoamine oxidase) inhibitors, and **selective serotonin reuptake inhibitors (SSRIs)**. What differs from one class of drugs to another is which neurotransmitter(s) are affected and the mechanism of action. For example, SSRIs only act on serotonin and not any other neurotransmitter

(hence the name "selective"). The way they work is by preventing the reuptake of excess serotonin from the synaptic gap. Side effects of drugs may also differ.

Perhaps the most popular SSRI is fluoxetine. The most common trade name for fluoxetine is Prozac.

Randomized controlled trials

Clinical trials are studies conducted with the aim of investigating the effectiveness of treatment. **Randomized controlled trials** (RCTs) are one of the types of clinical trials and are the gold standard in testing drugs. An RCT is a true experiment with an experimental group and one or several control groups where group allocation is performed randomly and all potential confounding variables are carefully controlled. Normally participants do not know which group they have been allocated to.



▲ Figure 4.5 A randomized controlled trial

In testing the effectiveness of a drug, it is important to consider the placebo effect. The placebo effect is when participants change their behaviour simply because they think they are undergoing treatment. For example, when they take a pill that they think is a drug but is really just sugar. Placebo effects are actually very strong: it was shown that placebo effects may account for up to 30% of effectiveness in psychotherapy on average (Prochaska and Norcross, 2018).

Surely if we want to show that a drug is effective, we need to demonstrate that it causes a significant reduction of symptoms over and above what could be expected with placebo. For this reason, clinical trials typically include a **placebo group**. Sometimes the second control group is included—the waitlist group. In this case, there are three experimental conditions:

1. Experimental group: participants receive the drug
2. Placebo control: participants think that they receive a drug, but actually they receive a placebo.
3. **Waitlist control**: participants are enrolled in the study but told to wait until it is their turn to undergo treatment. While on the waitlist, they receive no medication (or placebo).

The idea is that you would probably observe some minimal improvement in the waitlisted condition, a larger improvement in the placebo condition, and an even larger improvement in the experimental condition provided that the drug is actually effective.

Discussion

Why do you think we would expect participants in the waitlist condition to show an improvement, however minimal?

ATL Research, Thinking

Measuring the dependent variable

We want to demonstrate that taking a medication leads to a reduction in symptoms. But what exactly do we consider to be a reduction in symptoms? One of the measures that was commonly used in research to assess the symptoms of depression is called the Hamilton Depression Rating Scale (HAM-D).

The original scale (Hamilton, 1960) consisted of 17 questions and was filled out by a clinician or a researcher in the process of an interview. So you might say that the scale served as a script for a semi-structured clinical interview. An example of an item on the scale is:

Feelings of guilt:

- Absent = 0
- Self-reproach, feels he or she has let people down = 1
- Ideas of guilt or rumination over past errors or sinful deeds = 2
- Present illness is a punishment. Delusions of guilt = 3
- Hears accusatory or denunciatory voices and/or experiences threatening visual hallucinations = 4.

There are several measures that are typically used to measure the effectiveness of treatment:

- **Response rate**: percentage of patients who reported at least a 50% decrease in scores on a standardized scale (before treatment versus after).
- **Remission rate**: percentage of patients who reported few or no symptoms of depression at the end of the treatment.
- **Relapse rate**: percentage of patients who developed symptoms again (after the treatment was discontinued).

It is possible for a medication to have good response and remission rates, but a bad relapse rate. This would happen if a drug is effective at alleviating symptoms, but all symptoms immediately return once the drug is discontinued. So the question is, what should we consider an indicator of effectiveness of treatment: response rates, remission rates, relapse rates, all of them combined, or perhaps something else entirely? Different research studies used different indicators, which makes comparing results of these studies difficult sometimes.

Research in focus: treatment of adolescents with depression (TADS) study

Treatment of adolescents with depression (TADS) was a multi-site clinical research study in the USA. It examined the short-term and long-term effectiveness of drug treatment, non-medical psychotherapy, or a combination of the two for treating depression in adolescents aged 12–17 (The Treatment for Adolescents with Depression Study, 2007).

The treatments used in the study included fluoxetine (Prozac), which is a commonly used antidepressant, cognitive behavioural therapy (CBT), which is a non-medical intervention, or a combination. We will consider CBT in detail in Unit 4.4: Psychological treatment.

The study was conducted in three stages.

- Stage 1 (the first 12 weeks), acute treatment—participants were randomly assigned to one of four groups:
  i. fluoxetine alone
  ii. placebo + clinical management
  iii. cognitive behavioural therapy (CBT) alone
  iv. combination of fluoxetine and CBT.

  Clinical management is a mode of treatment where clients and therapists meet and talk, but no specific psychotherapeutic technique is used. It was used here so that participants could not easily work out that they were in the placebo group. At the end of the 12-week period, participants who were taking the placebo were informed about it. If they showed no improvement during the study period, they were allowed to choose any of the other three modes of treatment (but they no longer participated in the study).
- Stage 2, consolidation treatment—participants from the three active treatment groups (fluoxetine, CBT, and the combination) continued with their treatment for six more weeks.
- Stage 3, continuation treatment—participants who improved earlier were observed for another 18 weeks. So, the study lasted 36 weeks in total (The Treatment for Adolescents with Depression Study, 2007).

The main indicator used in this study to measure the effectiveness of treatment was response rates. Results are summarized in Table 4.2.

Response rates	12 weeks	18 weeks	36 weeks
Placebo + clinical management	35%	Discontinued	Discontinued
Fluoxetine (Prozac)	61%	69%	81%
CBT	44%	65%	81%
Fluoxetine + CBT	71%	85%	86%

▲ Table 4.2 Treatment of adolescents with depression study (TADS): results

What conclusions can be derived from these results?

1. All three active treatments of depression outperform placebo. Combination treatment outperforms the other treatments in the short term.
2. CBT gradually "catches up" with medication and even with a combination treatment. Given this finding, one may be tempted to ask if drugs should even be used—they may have unpleasant side effects and they do not allow the patient to learn new behaviours. However, in many cases quick action is needed, especially if depression at the initial stages is severe. Combination treatment seems to be the best choice in a variety of situations. Cognitive treatment also depends a lot on the patient's involvement and motivation, something that depressed patients may not have a lot of, especially at the initial stages.
3. No therapy is effective for 100% of patients. Note: one in every five adolescents did not demonstrate response to treatment even after 36 weeks of continued medication.

However, the overall result of the study suggests that antidepressants (either alone or in combination with cognitive therapy) are effective for the treatment of

depression both in the short term and in the long term. This seems to support the **chemical imbalance theory**—the proposition that depressive symptoms are caused by neurotransmitter imbalances in the brain and that restoring balance will lead to a reduction in symptoms.

Criticism of randomized controlled trials

Although RCTs are widely recognized as the gold standard in investigating the effectiveness of biological treatments, they are not without challenges.

One problem is that trial participants typically do not represent the population on the whole, so results from RCTs may not apply more generally. Participation in RCTs is voluntary. Think about people who signed up for such trials—are they likely to be different in some ways from the general population? Furthermore, researchers usually eliminate participants with comorbid conditions (e.g., anxiety in addition to depression) to enable a "cleaner" comparison. These considerations negatively affect the representativeness of the research sample, which in turn affects the extent to which results apply to real-life patients. It is a well-known problem of a **trade-off between internal and external validity** in true experiments. RCTs are designed to maximize internal, but not external validity.

Another problem is the social and ethical issue surrounding RCTs. They are lengthy, time-consuming, and difficult to conduct. Administration bodies such as the FDA (Food and Drug Administration) in the USA base their approval of a drug on evidence from clinical trials. When evidence is not sufficient, approval will not be given, and it will be illegal to manufacture and sell the drug. This process may take years.

Activity

Here is the typical process of conducting clinical trials that any new drug developer must follow in order for the drug to be approved for use by the FDA.

Phase 0 trials are designed to develop the drug and conduct preclinical research with animals.

Phase I trials test the safety, side effects, and the best dose of a drug. They include a small group of volunteers, usually 20–100 people. To determine the best dose, different subjects get different doses of the drug (dose-ranging) to find the point where the dosage becomes too harmful or poisonous. Volunteers should be healthy and they are paid for their participation.

Phase II trials typically include a larger group of participants (50–300). Researchers measure the effectiveness of the drug. Most drugs that fail to be approved will fail Phase II. The most common reasons are that the drug's effectiveness was lower than anticipated or that its toxicity turned out to be too high.

Phase III trials assess the expected effectiveness of the drug in real-life clinical practice. They include large patient groups (usually thousands). Phase III trials are always designed as RCTs. It is usually expected that there will be at least two successful Phase III trials before a regulation body such as the FDA in the USA gives approval for putting a drug on the market.

The entire procedure can be very costly and time consuming. It was estimated (Holland, 2013) that the full cycle of new drug development from preclinical research to marketing can take 12–18 years and often costs over one billion USD. Clinical trials can also be unsuccessful, so they may need to be restarted from the beginning or discontinued. It is a procedure designed in order to ensure that no harm is done.

Imagine that there is an unexpected epidemic of a mental disorder that incapacitates people (a pandemic similar to what happened with COVID-19). The question is, if you are pressed for time and need to work faster, which of these phases would you be willing to compromise? How do you think it will affect the ethics of new drug development?

ATL Thinking, Communication

Typical biases in randomized controlled trials

Not every study that labels itself as a RCT has been performed adequately. Just because a study is randomized does not mean it is unbiased. Here are four examples of typical pitfalls (Lewis and Warlow, 2004):

1. Poor **allocation concealment**. Allocation is which of the groups (treatment or control) the patient will be assigned to. Allocation is said to be concealed when it is not possible for a clinician to opt out of randomizing after they have seen what group the patient will be in if randomized. However, this standard is often violated. For example, if the clinician alternates between treatment and control groups (this patient, treatment; next patient, control; next patient, treatment; and so on), then the clinician will know prior to making the randomization decision which group the client will be assigned to. This is an obvious source of bias because such knowledge may affect the clinician's decision not to randomize (as in "the next one should be placebo, but this patient's symptoms look too severe, so I will not offer them participation in the trial").
2. **Unblinding**. In a good trial, everyone involved should be blind to the allocated condition. Participants should not suspect what group they are in, nurses administering treatment should not know whether they are prescribing real medication or placebo, and even the person conducting results analysis should not know what "Group A" and "Group B" stand for. Unblinding occurs when someone acquires that knowledge. Complete blinding should be the goal, but it is not always possible. For example, medications can have side effects, so a patient experiencing side effects may suspect that they are in the treatment condition.
3. Groups that are not entirely equivalent from the start. Randomization is a powerful equalizer. However, it works best with large sample sizes. When the sample is small (as is often the case in some clinical trials), there exists a chance that randomization will not produce groups that are absolutely equivalent. To address this problem researchers typically use stratified sampling.
4. **Vested interests** and **publication bias**. New medication is often developed by private pharmaceutical companies. If clinical trials are funded by these companies, there could be publication bias—results of successful trials could be more likely to get published than results of unsuccessful trials. Publication

Discussion

What other approaches to ensuring adequate concealment can you think of?

ATL Research, Thinking



▲ Figure 4.6 Food and Drug Administration (FDA)

itself is an expense, so why would a company incur this cost to publish something that damages its commercial interests? It has been observed that research funded by pharmaceutical companies appears to produce "better" results than research funded from other sources.

CONSORT statement

An important milestone in ensuring the quality of RCTs worldwide is the **CONSORT statement**. It stands for "Consolidated Standards of Reporting Trials" and it is a set of recommendations for reporting randomized trials. It offers a standard way for authors to write reports of their findings, increasing transparency, and ensuring unbiased reporting.

The most recent version (CONSORT 2010 Statement) is a 25-item checklist. You can find it online. The statement has been embraced by many publishers, with hundreds of peer-reviewed journals now requiring that authors' submissions be presented in accordance with the CONSORT format.

Chat with AI

Here are a few suggested prompts for generative AI for you to carry out a further inquiry into the topic:

- How could the process of getting FDA approval for COVID-19 vaccines be accelerated?
- Apart from RCTs, what other methods can be used to test medical treatments for depression? What are the strengths and limitations of these methods compared to RCTs?
- Could you provide an example of a hypothetical report of a clinical trial investigating the effectiveness of a new antidepressant? This report should use the following criteria of the CONSORT statement [include the criteria].

Activity

You can study the full checklist in Table 1 of Schulz et al. (2010), "CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials" published in *BMC Medicine*. The article is freely available online. The article also contains a flowchart to help a researcher with making the important decisions along the way.

Select any RCT (you can use the TADS study) and apply the CONSORT criteria to analyse it. How difficult do you think is it to report a study in full accordance with all these criteria?

ATL Research, Self-management

Conceptual analysis

Causality and perspective

Biological treatment of mental disorders is based on the assumptions and findings of the biological approach to understanding behaviour. For example, it is believed that the main factor that causes the disorder is an imbalance in the level of a specific neurotransmitter in the brain.

Bias

As you have seen, clinical trials are designed to enable researchers to make conclusions about the effectiveness of a drug that are as unbiased as possible. To help with this, initiatives such as the CONSORT statement have been put in place. One of the reasons why clinical trials are so long and expensive is that we try to ensure a rigorous investigation and accurate results. However, bias does still creep in. In this section, we have considered some typical sources of bias in RCTs.

Measurement

Any clinical trial needs to measure the target (dependent) variable, but this in itself is not straightforward. A variety of indicators can be used—response rates, remission rates, relapse rates. Short-term effects are compared to long-term effects. Effects may be different for mild, moderate, and severe symptoms and different for various populations. There are also side effects and their severity that need to be considered, as the decision to prescribe a drug is always a comparison of expected risks and benefits. There is also the problem of measuring symptoms of depression in a research setting, where the full clinical diagnosis is not possible.

Change

The concept of change is related to treatment: by manipulating the patient's intake of chemical substances, we are trying to induce changes in their mental health symptoms. This is closely linked to measurement: we need to know if change occurs, as well as what kind of change it is—stable or transient, deep or superficial, quick or slow. One of the strongest arguments in favour of treating severe depression with antidepressants is that they have a very quick effect on dangerous symptoms such as suicidal ideation. One of the strongest arguments against antidepressants is that over the long term, the effectiveness of talk therapy is the same as medication. However, the effects of talk therapy are more long-lasting whereas antidepressants only work if you continue taking them.

Responsibility

In the process of developing a new drug, we have to juggle many conflicting considerations. For example:

- commercial interests of pharmaceutical companies
- ethical considerations to avoid adverse effects
- methodological considerations to maximize study validity and accuracy of results.

These considerations are often at odds with one another. For example, if all we ever cared about was methodology, we would fast-track to Phase III trials and recruit large groups of participants for an RCT. However, this would be unethical because some of the participants may experience adverse effects. If all we ever cared about were commercial interests, we would not have any trial process at all. Regulatory bodies (such as the FDA) and standards (such as the CONSORT statement) exist to ensure the delicate balance between all these considerations. Doing minimal harm (ethics) is always a priority of such organizations and standards.

4.3 Cognitive models of mental disorders

Inquiry questions

- What role do our thoughts play in the development of mental disorders?
- Can we treat a disorder by learning to control our thoughts?
- How can we establish the main factors that cause depression?

What you will learn in this section

Key learning:

- Cognitive appraisal mediates the influence of environmental stress on the development of symptoms of depression.
- Aaron Beck's cognitive theory of depression is based on the idea that automatic negative thoughts are the leading cause of depression.
- Evidence for this comes from studies that show an association between cognitive styles and depressive symptoms. Longitudinal evidence also supports this claim.
- Indirectly, effectiveness of CBT also serves as evidence of validity of the cognitive theory: CBT is effective; therefore, the cognitive theory is valid.
- The cognitive theory of depression has limitations:
  - the causal link between automatic thoughts and depressive symptoms is debated
  - it cannot explain depression-like behaviour in animals
  - It ignores non-cognitive factors of depression.

Key terms: cognitive appraisal, automatic thoughts, Beck's cognitive theory of depression, cognitive triad, negative thinking, cognitive biases, cognitive style, cognitive behavioural therapy (CBT), treatment–aetiology fallacy

In a wider context

According to the diathesis–stress model, biological factors create a susceptibility to a disorder, but it will only manifest itself in symptoms if environmental variables are "favourable". Where do cognitive variables come into the picture?

Cognitive factors may serve the function of being important mediating variables between the influences of the environment and genetics and symptoms of mental disorders. Humans interpret environmental events, and it is this interpretation that has an immediate effect on them, rather than the event itself.

Cognitive appraisal

You might have noticed how many individual differences there are in the way people react to adverse circumstances. Some people panic even when the tiniest thing goes wrong. Others remain calm even in the face of a crisis. This may also be evident in the development of mental disorders, such as depression. Some people interpret events of their life in a way that makes them more vulnerable to depression. For example:

- they see signs of hostility in others when in fact there are none
- they fail to notice their successes and exaggerate their failures
- they see their future as dark and catastrophic when in fact there is no reason for that
- they think that other people notice flaws in their appearance when in fact other people are not even looking.

Cognitive appraisal is an individual's interpretation of their experiences. It is this interpretation that plays an important role in developing symptoms of mental disorders. It is not stress but perceived stress, not lack of control but perceived lack of control, not weakness of social ties but perceived weakness of social ties. Of course, this is not to claim that environmental and biological factors themselves are not important. Indeed, it is hard to imagine being in an abusive relationship and being able to simply re-interpret it into something positive, or living in poverty and not letting it influence you. However, cognitive variables do function as important mediators in a variety of circumstances. Additionally, cognitive factors are sometimes the only thing we can change: it is not always possible to change your environment and it is surely impossible to change your genetic inheritance.

Discussion

To what extent do you agree with the notion that, if you cannot change the situation, you need to change the way you think about the situation?

Taking this one step further, do you agree that teaching people how to change the way they think about their situations should be the primary focus of psychotherapy?

ATL Self-management, Thinking

Beck's cognitive theory of depression

A great example of the cognitive approach in explaining and treating MDD is the cognitive theory of depression, initially proposed by Aaron Beck.

Beck's cognitive theory of depression (Beck, 1967) suggests that negative thoughts and a biased interpretation of events are the factors influencing depression. At the heart of the theory is the co-called **cognitive triad**:

- negative views about oneself
- negative views about the world
- negative views about the future.

All these "views" are essentially schemas. They are formed by early life experiences and they may remain latent until they are activated by adverse environmental events. They manifest themselves as **automatic thoughts**—thoughts that occur involuntarily in response to everyday events.

Imagine a student submitted draft coursework in mathematics and got a preliminary mark lower than they expected. An example of the cognitive triad in action would be the following thoughts: "I am bad at maths" (about oneself), "The teacher despises me and will be unwilling to help" (about the world), "I will probably fail this course" (about the future).



▲ Figure 4.7 Aaron Beck

Exam tip

The concept of schemas is explained in detail in Unit 2.2: Schema theory.

Exam tip

The concept of cognitive biases is explained in detail in Unit 2.5: Cognitive biases and the dual processing model of thinking and decision-making.

Having these negative schemas influences the way people perceive reality and results in a number of **cognitive biases**, such as:

- overgeneralization (making a general conclusion based only on a small number of events: "I didn't do well on this test, so I am bad at maths")
- selective abstraction (focusing on the negative aspects of the situation and ignoring the positives: "The teacher said my essay is not logically structured", although the teacher also pointed out that the essay was excellent in all other ways)
- personalization (blaming oneself for causing the situation while downplaying the significance of other factors: "I am to blame, I didn't do well because I didn't put in enough effort", although the school gave too many assignments with a due date on the same week).

This pessimistic lens creates an ongoing, self-reinforcing cycle: the person only sees the negatives, this changes their behaviour and elicits more negative responses from the environment, this reinforces their negative perceptions, and so on.

Beck's original theory has evolved over time. Most notably, Beck used elements of behaviourism to incorporate the behavioural component into his cognitive theory. This led to the creation of **cognitive behavioural therapy (CBT)**—the most widely accepted method of depression therapy today.

Activity

Role-play a mock client–therapist session. One person will play the role of a depressed patient. Find a case study online, or make one up, or use generative AI to make one. Remember to present various symptoms (cognitive, emotional, behavioural) and make it as specific as possible. The other person will play the role of a therapist. Using principles of Beck's theory, design a plan for the session and carry out the intervention.

ATL Communication, Self-management, Social

Chat with AI

Here are some suggested prompts to get more in-depth knowledge about Beck's cognitive theory of depression:

- Could you give me examples of cognitive biases used by depressive patients according to Beck's cognitive theory of depression?
- Could you outline the principles of Beck's theory in five to seven key statements?
- What other cognitive explanations of depression are there apart from Beck's theory?
- Please give me an example of a case study of a depressed patient (describe their problem). Then imagine this patient came to a cognitive-behavioural therapist and the therapist planned 10 therapy sessions based on the principles and ideas of Aaron Beck's theory of depression. Could you give me an outline of each of these 10 sessions, including the exercises and home assignments used by the therapist and a brief explanation of their intended purpose?

Beck and his followers also put a lot of effort into gathering evidence and investigating the empirical basis of the theory. To this end, they developed such measurement tools as the Beck Depression Inventory (Beck and Steer, 1984) and the Dysfunctional Attitude Scale.

Later on, the theory was enriched by physiological elements, with a demonstration of how negative schemas can affect biological factors such as brain structure and the release of stress-related hormones such as cortisol (Beck, 2008). Later still, additional perspectives were added (such as genetics) and the

theory got transformed into the so-called "unified model of depression" (Beck and Bredemeier, 2016).

Evaluation of Beck's theory: cognitive styles and symptoms of depression

There is some sound empirical evidence behind the claims of Beck's cognitive theory of depression.

For example, Alloy, Abramson, and Francis (1999) aimed to investigate if negative **cognitive styles** affect vulnerability to depression when people confront negative life events. In other words, if you confront a negative life event, to what extent will developing depression depend on your thinking patterns?

The hypothesis here is similar to the diathesis–stress model, only with cognitive variables instead of biological ones. The diathesis–stress model suggests that the influence of stress on depression is mediated by genetic predisposition. Alloy, Abramson, and Francis's (1999) hypothesis was that it is mediated by a "cognitive predisposition".

At the beginning of the study, a group of college students with no mental illness was split into a low-risk and a high-risk group. Groups were based on presence versus absence of negative cognitive styles, which were assessed by a questionnaire. Follow-up assessments were conducted for 5.5 years after that: every six weeks for the first 2.5 years, and every four months for the rest of the time. Each follow-up assessment included both self-report and interview measures of stressful life events, cognitions, and symptoms of depression.

When researchers only focused on students who had no prior history of clinical depression, they found that 17% of those in the high-risk group developed a first onset of MDD during the 2.5 year follow up period, compared to only 1% of those in the low-risk group.

When researchers only focused on participants with past depression, they found a similar pattern: 27% of the high-risk group developed a recurrent episode of MDD, compared to only 6% of the low-risk group.

As part of the assessment procedures, participants were given cognitive tests, one of which included memory for adjectives. It was found that high-risk participants were faster at processing and better at remembering negative adjectives such as "incompetent" and "worthless". Conversely, low-risk participants were better at processing and remembering positive adjectives such as "successful" and "lovable".

The study seems to provide convincing evidence that cognitive factors do indeed play an important role in mediating the onset of depression. An important limitation, of course, is that the study is not a true experiment, so cause-and-effect inferences are not warranted. Alloy, Abramson, and Francis (1999) did not manipulate participants' cognitive styles. They merely identified pre-existing groups of students with positive and negative cognitive styles and compared them (longitudinally). The problem is that these groups might have differed in something else in addition to their cognitive styles, and it is this "something else" that might have created these differences in the follow-up study.

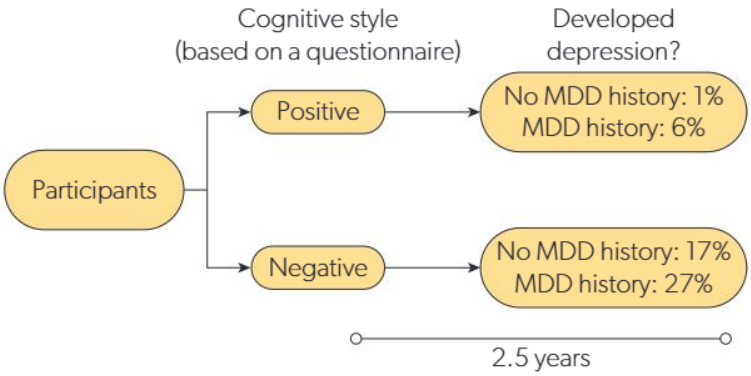
Activity

Alloy et al.'s (1999) study is not a true experiment, which limits our interpretation of causality in it. But how could we conduct a true experiment to test their hypothesis? We would have to force some participants into a negative cognitive style, and that is neither easy nor ethical.

Could you take the original hypothesis of Alloy et al. (1999) and try to design a true experiment to test it? Answer the following questions:

- How will you select your sample of participants?
- How will you allocate them into groups?
- How will you manipulate their thinking style?
- When and how will you measure the dependent variable?

ATL Research, Thinking



▲ Figure 4.8 Results of Alloy, Abramson, and Francis (1999)

Evaluation of Beck's theory: comparison with alternative models

Oei and Kwon (2007) conducted an interesting study in which they compared several models of depression and found that Beck's cognitive model was best at explaining and predicting symptoms of depression and anxiety after negative life events. Participants in this study were 107 Korean recent migrants in Australia. Due to the stress of having recently moved to a new country, these people are more likely than the general population to develop anxiety and depression.

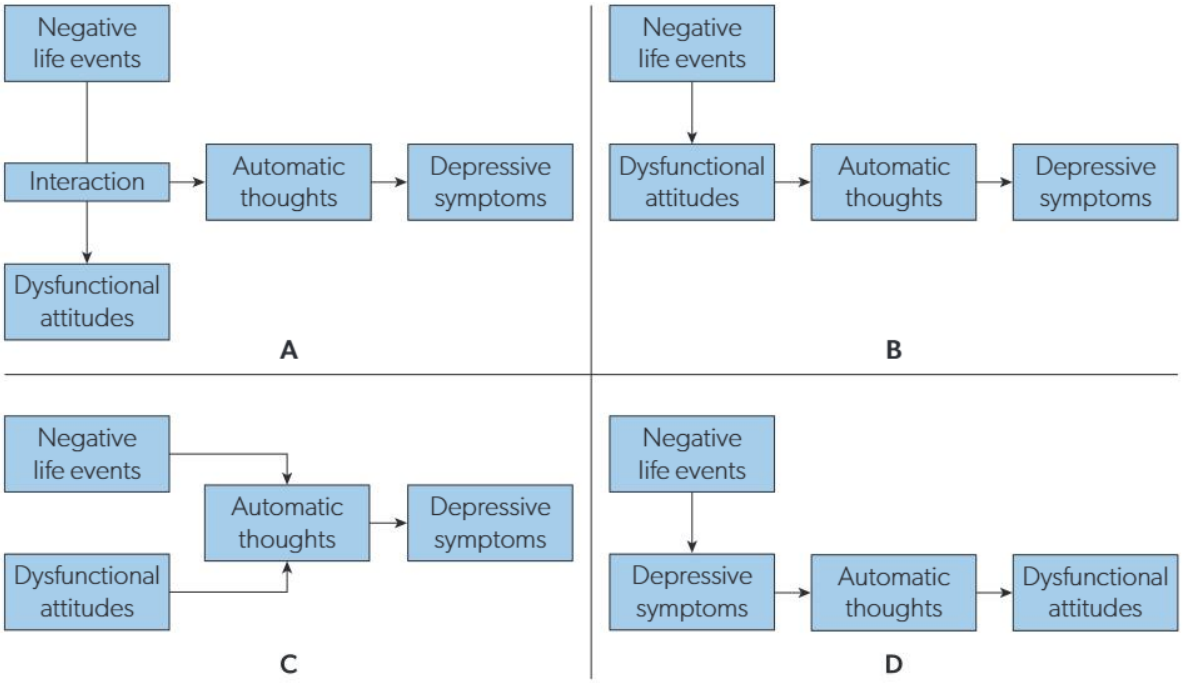
Participants were given a set of questionnaires on two occasions (Time 1 and Time 2), six months apart. The battery of questionnaires included the following:

- Migrant Stress Inventory (MSI). This self-report measure included questions about the stressors that participants encountered in a new culture.
- Dysfunctional Attitudes Scale (DAS). This questionnaire included 40 items such as: "If I do not do well all the time, people will not respect me", "If someone disagrees with me, it probably indicates that he does not like me".
- Automatic Thoughts Questionnaire (ATQ). This included 30 items capturing the typical automatic negative thoughts of people with depression.
- Beck Depression Inventory (BDI). This is the most widely used measure of depression, a questionnaire that includes 21 items capturing the level of severity of various depressive symptoms.

Patterns of correlations between all measures were examined to find which of the alternative models provided the best fit to the data. Researchers compared four different models (See Figure 4.9).

Model A here is in line with the modern version of Beck's theory. It suggests that negative life events interact with dysfunctional attitudes and this interaction leads to the development of automatic thoughts, which in turn leads to the development of depressive symptoms.

Model D is the most contradictory to Beck's theory because it presents cognitive variables as consequences of depression rather than the causes of it.



▲ Figure 4.9 The alternative models used in Oei and Kwon's study

Results showed that model A (Beck's theory) provided the best fit to the available correlational data, model D (the most contradictory one) provided the worst fit, and the other two models (A and C) were in between.

The study used a correlational technique known as "structural equation modelling". Of course, it is not a true experiment and no variable was directly manipulated by the researchers, so strictly speaking no cause–effect inferences are possible. However, remember that the study included two measurements six months apart. If at Time 1, we observe negative thoughts but no depressive symptoms and then at Time 2, we observe both negative thoughts and depressive symptoms, then it is likely (though not certain) that negative thoughts have influenced depressive symptoms. It is certainly unlikely in this scenario that depressive symptoms have influenced negative thoughts.

Effectiveness of cognitive behavioural therapy as support for Beck's model

Apart from studies that seem to support the basic claims of the cognitive theory of depression, there have been a number of studies investigating the effectiveness of CBT for treating depression. We already considered one such study in a previous section (Unit 4.2: Biological treatment). The Treatment of Adolescents with Depression Study (TADS, 2007) demonstrated, among other things, that CBT results in response rates similar to medication over the long term. The onset of effects with CBT is not as quick as with medication, but CBT patients are less likely to relapse. If CBT is effective in treating depression, then this may be taken as indirect evidence that the cognitive theory of depression is valid.

See Unit 4.4: Psychological treatment for more discussion on the effectiveness of CBT.

Limitations of the cognitive theory of depression

It is clear that cognitive factors such as **negative thinking** play a role in developing depression, and the numerous studies under the umbrella of the cognitive model have been successful in demonstrating this. A more difficult question is whether negative thinking does in fact cause depression, as opposed to being something correlated with depression or a by-product of depression or a mediating factor. Some researchers doubt that the evidence collected so far is sufficient to say with confidence that negative thinking is the main causal factor of depression.

Another limitation is that there exist phenomena related to depression that Beck's model cannot explain. Sometimes depressed patients develop cognitive delusions: for example, Cotard's delusion is the condition where patients believe that they do not exist. There is nothing in the cognitive model of depression that would explain why such a delusion may arise.

Animals exhibit depression-like behaviours. Antidepressants are often tested on mice before they are tested on humans because mice may behave in a way that can be described as "depressed" or "non-depressed". Mice cannot have negative thoughts. Therefore, Beck's theory in its traditional form is not applicable to animals.

Finally, the cognitive model is obviously focused on cognitive factors. However, we know that other factors also affect depression—for example, some people are more genetically predisposed to depression than others. We know that antidepressants can be as effective as CBT in treating depression and can outperform cognitive treatment in the short term. The cognitive model alone cannot explain these findings. Therefore it is difficult to view this theory as a universal theory of depression (although the modern version of Beck's theory includes these other variables and is therefore more universal).

TOK

Why is establishing cause–effect relationships so important to us? Do you think it is equally important in various areas of knowledge (such as the natural sciences, the human sciences, history)?

Conceptual analysis

Causality

Perhaps the main ongoing debate in the cognitive aetiology of mental disorders is that of the exact role of cognitive variables in the onset of depression. When we say "cognitive factors play a role in the development of depression", what do we mean?

One position is to claim that cognitive factors are important by-products or aggravating circumstances of depression—once depression develops, it may get worse if the person has negative patterns of thinking. Another position is to claim that cognitive factors have a causal role in depression. If we assume this position, then we believe that depression develops because there are negative patterns of thinking. Beck's cognitive model of depression makes this strong causal claim: negative patterns of thinking cause depressive symptoms.

Measurement

Research in this area heavily relies on self-reporting: one cannot easily think of other ways to measure negative thinking patterns. Most research has used

questionnaires where participants have to self-assess the frequency of given typical thoughts or attitudes. Some researchers became more creative and used more objective measures such as the speed of processing of positive versus negative words. Even depressive symptoms are typically measured with the use of self-report questionnaires (e.g., Beck's Depression Inventory).

Bias

On the one hand, the concept of bias is relevant to Beck's cognitive model of depression because the negative thinking of a depressed patient is characterized by a number of cognitive biases (overgeneralization, black-and-white thinking, selective abstraction, etc.).

On the other hand, one may speak about biases in the way the model is constructed and researched. For example, one of the typical arguments to support Beck's cognitive model is the effectiveness of CBT—if CBT works, then the theory that lies at its basis must be correct. However, this argument represents a fallacy (sometimes called the "**treatment–aetiology fallacy**"). This fallacy is obvious even with simpler explanations: for example, if drinking mint tea makes you cough less, it does not mean that your cough was caused by a lack of mint in your body.

Perspective

Beck's cognitive model of depression represents the cognitive perspective in psychology. It is important to note, however, that the model does not negate the role of all other factors. Instead, it just states that cognitive variables are more important. Pragmatically speaking, the cognitive model may have an advantage over the biological approach. Suppose we have acknowledged that both cognitive factors and genetic inheritance mediate the influence of stressful experiences on the development of depressive symptoms. One cannot change genetic inheritance, but it is possible to change patterns of thinking. The cognitive model may be preferable in this practical sense.

Change

Cognitive behavioural therapy is a systematic intervention targeted at changing automatic negative patterns of thinking. The concept of change is therefore very prominent in this therapy. More on this can be found in Unit 4.4: Psychological treatment.

A related issue is change versus stability in negative patterns of thinking. We need to make sure that research uncovers patterns of thinking that are stable and not just a consequence of transient mood changes on the day the participant filled out the questionnaire. Beck's cognitive theory sees negative patterns of thinking as stable traits rather than transient states. Ideally, the measurement in research studies should reflect that, confirming, for example, that the uncovered patterns of thinking remain the same each day. This is difficult and not always done, as it would require repeat measurements on the same participants over multiple days.

Responsibility

Ethical considerations in studies of CBT are probably more crucial than ethical considerations in studies that test theoretical claims of the cognitive model. The former involve creating groups of participants who receive different treatments, often assigning them randomly into these groups. Inevitably, some groups will

see greater improvement compared to others. A placebo group is also typically included, and these are participants who do not receive any treatment at all although they think they do. The latter does not normally involve treatment and only compares depressed and non-depressed people in terms of their typical patterns of thinking.

Researchers have a great ethical responsibility of reporting results in a balanced and impartial way. If we mistakenly believe that cognitive factors are more important in the aetiology of depression than biological factors, we may end up accepting a belief that does not provide the best possible foundation for successful treatment.

4.4 Psychological treatment

Inquiry questions

- How does talk therapy work?
- Is a self-help app as effective as seeing a real therapist?
- Will we see the emergence of "artificial therapists" in the future?

What you will learn in this section

Key learning:

- Psychological treatment assumes that the root cause of mental disorders is non-medical. CBT assumes that depression is caused by negative automatic patterns of thinking.
- CBT relies on client motivation and agency and is designed as a series of sessions with home assignments and behavioural exercises.
- Computerized CBT has shown promising results: it offers similar effectiveness at a reduced cost.
- Another promising area is CBT apps. They completely remove the therapist and there is an associated drop in effectiveness, but this may be offset by a much wider reach and a much lower cost.
- Assessing effectiveness of treatment is complex: we need to consider indicators of effectiveness along with the dimensions of short-term versus long-term effectiveness, mild versus severe symptoms, and also patient adherence, cost, and accessibility.
- The ethics of the relationship between a client and a therapist is regulated by a set of ethical standards that are distinctly different from the standards regulating the work of a psychologist conducting a scientific study.

Key terms: cognitive behavioural therapy (CBT), automatic thoughts, negative thinking, behavioural exercises, computerized CBT, CBT app, chatbot, conversational AI, bibliotherapy, adherence to therapy, code of ethics

In a wider context

Biomedical treatment is based on research that demonstrates the influence of biological factors (such as concentration of neurotransmitters) on mental disorders. It assumes that biological factors are the root cause of such disorders. Antidepressant drugs are the biological treatment for depression.

Psychological treatment assumes that the root cause of a mental disorder is non-biological. For example, Beck's cognitive theory of depression maintains that depression is caused by automatic patterns of **negative thinking**. Accordingly, cognitive behavioural therapy (CBT) attempts to teach patients how to replace these negative patterns of thinking with more realistic ones.

Exam tip

The content of the previous units (such as 4.1: Biological explanations of mental disorders, 4.2: Biological treatment, and 4.3: Cognitive models of mental disorders) is also highly relevant to this topic.

For example, we have already looked at TADS—the Treatment of Adolescents with Depression Study (2007). This study provided a direct comparison between medical and non-medical treatment of depression, outlining their relative advantages and disadvantages.

Activity

There is a great variety of non-medical (psychological) treatments. Apart from CBT, there is behavioural psychotherapy, psychodynamic psychotherapy, systemic (family) therapy, supportive psychotherapy, interpersonal therapy. All of these are broad types, and there are more specific approaches within each.

Make up a fake problem. Split into groups. Let each group do some research about one non-medical treatment other than CBT. Then role-play a mock client–therapy session with each of these methods: how would these different therapists approach the same problem?

Use generative AI for suggestions about specific interventions that could be used in one of these approaches to work with the hypothetical client. You can provide the AI with a description of the fake case as part of your prompt.

ATL Communication, Self-management, Social

Cognitive behavioural therapy: how it works

In **CBT,** both the client and the therapist take a scientific approach to the problem with a clear focus on measurement and hypothesis testing. Suppose you tell the therapist that lately you have been experiencing prolonged periods of bad mood. The therapist may want to know what kind of **automatic thoughts** you have during such episodes. You probably won't know what they are: by definition, they are automatic and you are not aware of them unless you intentionally focus on them. However, you agree with the therapist that it would be important to find out.

In order to measure your automatic thoughts, you could get homework. The assignment could be to keep a diary and, every time you realize that you have been experiencing a really low mood, you need to find a quiet place, remember everything you had been thinking about for the last five minutes and write it down. You should do this at least once a day.

Suppose that at the next meeting with the therapist you analyse these notes and realize that a prominent theme is a feeling of worthlessness. For example, every time you work on a school assignment, you focus on its flaws and worry that your teacher will not be pleased with it. The therapist then wants to investigate if these thoughts and expectations are unrealistic.

You get another homework assignment where you are required to keep a log of automatic expectations and track the outcomes. For example, for each school assignment, you predict the grade it will receive and the feedback your teacher will give. Then next to these predictions you write down the actual grades and feedback. After a month, you try to look at this log impartially together with the therapist and analyse whether or not the predictions were realistic, and what exactly was unrealistic about them. Once that is established, you will work on replacing unrealistic expectations with more rational ones. Again, this will take the form of discussions with your therapist and subsequently engaging in a set of home assignments and **behavioural exercises**. You work on your own behaviour similarly to how you would practise the piano or how you would follow your gym programme.

CBT and motivation/agency

Unlike some other forms of therapy, CBT places a large importance on the client's own decision-making and agency. A therapeutic alliance between the therapist

and the client is considered to be one of the major factors of success. The client is encouraged to take responsibility for what is happening, to set their own goals, and to take control of the details of the process. For example, imagine the home assignment is to spend one hour a day enjoying pleasurable activities (to learn how to dissociate from the pressures of work), but the client finds it too difficult. They keep thinking about unfinished work, keep ruminating, and fail to relax. The client is encouraged then to modify their goals and to change the assignment to something more manageable, for example 20 minutes a day. It is important to instil a sense of self-efficacy in the client because it is the client and not the therapist who will be the main driver of change in the long term.

Computerized CBT

CBT was developed at a time when psychotherapy was still dominated by "heavy" approaches such as psychoanalysis, which required many hours of analysis and interpretation, often going deep into childhood memories. Today, clients want results quickly. At least, they want to feel that something is changing in a positive way even after one or two sessions. Otherwise there is always a risk that they will stop coming. CBT provided such quick tangible results. It allowed clients to feel empowered and believe in their own agency, their ability to change things for the better.

If face-to-face interaction between client and therapist is replaced by video call sessions or even simple text chat, clients could do CBT without leaving home. If the drop in effectiveness is not very large, then a broader reach would be gained at the expense of a small decrease in outcomes. **Computerized CBT** (CCBT) would be cheaper, and would further improve access.

Andrews et al. (2010) conducted a meta-analysis of research studies of the effectiveness of CCBT. Their meta-analysis aggregated 22 RCTs for participants suffering from one of the following four disorders: Major Depressive Disorder, social phobia, panic disorder, or generalized anxiety disorder (GAD). The major intervention in all these studies was CCBT, and it was compared to either the "regular" face-to-face CBT or to control conditions such as placebo or waitlist.

The results are shown in Figure 4.10.

Here are the results:

- Across all 22 studies and all four disorders, CCBT demonstrated a statistically significant superiority over control conditions (placebo or waitlist). It was also large in terms of the average effect size. Interestingly, the heterogeneity was not significant, which means that on average all studies agreed that CCBT works.
- Follow-up data was reported in 14 of the 22 studies (from 4 to 52 weeks post-treatment). There was no evidence of significant relapse in any of the studies. CCBT also showed good levels of **adherence**: a median of 80% of participants across all studies completed all sessions in the programme.
- There were five studies that compared CCBT to face-to-face CBT. The difference between these groups turned out to be insignificant. However, therapist time in the computerized version was reduced by 50–79%. Reported satisfaction was good in both conditions.

Activity

Find the original Andrews et al. (2010) paper online. The title is "Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis" and the full text is available–for example, from the National Library of Medicine (NLM).

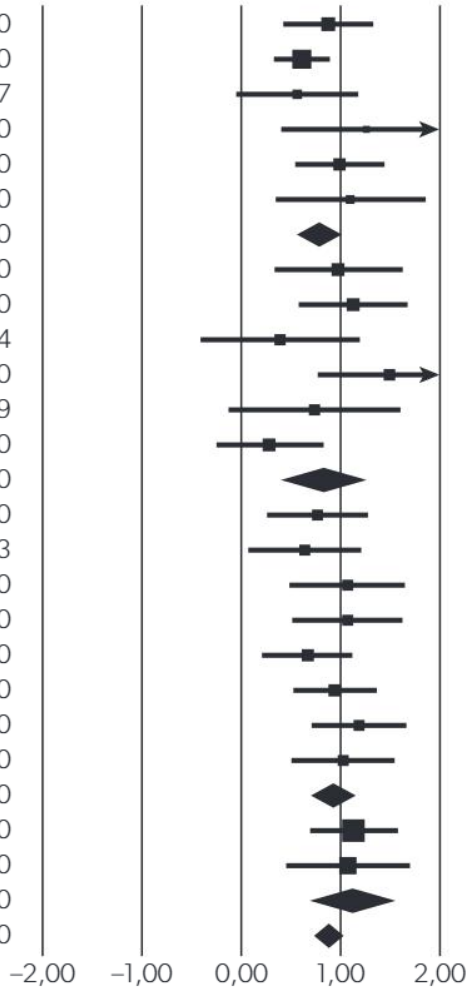
Answer the following questions based on your reading:

- What were the criteria that a study had to meet to be included in the meta-analysis?
- What was the overall superiority of computerized CBT on depression (effect size)?
- How long did the effect last?

If you find it difficult to read and understand the original paper, remember that generative AI can help to simplify, paraphrase, and explain difficult paragraphs.

ATL Self-management

		Hedges' g	Lower limit	Upper limit	*p*-Value
MDD	Andersson, 2005	0,87	0,43	1,32	0,00
	Kessler, 2009	0,61	0,33	0,88	0,00
	Perini, 2009	0,56	–0,05	1,17	0,07
	Selmi, 1990	1,26	0,40	2,11	0,00
	Titov, 2010	0,99	0,55	1,44	0,00
	Wright, 2005	1,10	0,35	1,85	0,00
		0,78	0,59	0,96	0,00
Panic	Carlbring, 2001	0,97	0,33	1,61	0,00
	Carlbring, 2006	1,13	0,59	1,66	0,00
	Klein, 2001	0,39	–0,41	1,18	0,34
	Klein, 2006	1,49	0,77	2,20	0,00
	Richards, 2006	0,74	–0,12	1,60	0,09
	Wims, 2009	0,28	–0,25	0,81	0,30
		0,83	0,45	1,21	0,00
Social phobia	Andersson, 2006	0,76	0,26	1,26	0,00
	Berger, 2009	0,64	0,08	1,20	0,03
	Botella, 2009	1,07	0,50	1,64	0,00
	Carlbring, 2007	1,07	0,52	1,61	0,00
	Furmark, 2009	0,67	0,22	1,11	0,00
	Titov, 2008 I	0,94	0,53	1,36	0,00
	Titov, 2008 II	1,18	0,71	1,65	0,00
	Titov, 2008 III	1,02	0,51	1,53	0,00
		0,92	0,74	1,09	0,00
GAD	Robinson, 2009	1,13	0,70	1,56	0,00
	Titov, 2009	1,08	0,46	1,69	0,00
		1,11	0,76	1,47	0,00
All studies		0,88	0,76	0,99	0,00



▲ Figure 4.10 Overview of effect sizes (Hedges' g) in Andrews et al.'s meta-analysis. Each study resulted in a mean estimate of the effect of therapy on the mental disorder, with bands indicating 95% confidence intervals.

Discussion

Would you prefer face-to-face or computerized therapy? Why?

ATL Thinking, Communication

Results of this meta-analysis seem to suggest that computerized CBT can offer comparable effectiveness (both short term and long term), while at the same time considerably reducing the amount of contact with the clinician. If that is indeed the case, it could make CBT more accessible to a larger group of people.

We still need to be cautious, however. This research is relatively new and available evidence is limited as 22 studies is not a large number for a meta-analysis. Only five of these studies provided a direct comparison between computerized and face-to-face CBT. Participants in almost all studies were volunteers which limits the generalizability of results to real-life patients. Measures of effectiveness in most studies were based on patient self-reporting. Regardless of these limitations, the first evidence in this area is promising.

Activity

The delicate therapist–client relationships are regulated by a specific set of ethical standards. Therapists have a great degree of influence over their clients. Clients are vulnerable and in a stressful time of their life, so abuse is a real possibility.

Find a **code of ethics** for the work of a practical psychologist online and analyse it. Present a brief summary to your class. Which guidelines and principles do you find to be the most crucial? Which guidelines and principles did you find surprising?

How are these codes of ethics different from the ones that regulate the work of psychologists conducting scientific research?

Here are a few options of professional codes of ethics from different countries:

- American Psychological Association (APA): "Ethical Principles of Psychologists and Code of Conduct".
- British Association for Counselling and Psychotherapy (BACP): "Ethical Framework for the Counselling Professions".
- Australian Psychological Society (APS): "Code of Ethics for Psychologists".
- Canadian Psychological Association (CPA): "Canadian Code of Ethics for Psychologists".
- Chinese Psychological Society (CPS): "Code of Ethics for Clinical and Counselling Psychology".

ATL Thinking, Self-management

Cognitive behavioural therapy apps

If it is established that CCBT seems to provide a comparable level of effectiveness at a reduced cost and clinician time, the next logical step is to question what will happen if we remove the therapist entirely and use **CBT apps** only.

There have been a number of smartphone applications enabling the delivery of self-help or guided CBT. With the rise of AI, there has been a rise of text-based **conversational AI** agents (**chatbots**). One wonders what level of effectiveness is provided by such applications and how beneficial (or harmful) their use is for the treatment of real-life patients with depression or other disorders.

Inkster, Sarda, and Subramanian (2018) studied the effectiveness of a chatbot ("empathy-driven conversational artificial intelligence agent") in promoting mental well-being. Participants were anonymous users from around the world who installed the app and self-reported symptoms of depression on a questionnaire. App usage was monitored, which allowed researchers to split 129 participants into two groups (low users and high users).

Results showed a larger improvement in the group of high users compared to low users (a moderate effect size of 0.63, $p = 0.03$). Of the user feedback about the app usage experience, 68% was positive. Researchers recognize that

Activity

Research different existing AI platforms for delivering psychotherapy or mental health counselling. This is a rapidly evolving field, so your best option may be asking your favourite AI chatbot!

Here are a few examples suggested by an AI search engine at the time of writing this book: Woebot, Tess (Therapeutic Emotional Support System), Wysa, Karim, Ellie. Carry out your own research to find some more examples.

What is the possibility that face-to-face therapy work will be gradually replaced by AI? If not, how will therapists learn to work together with AI to provide mental health services?

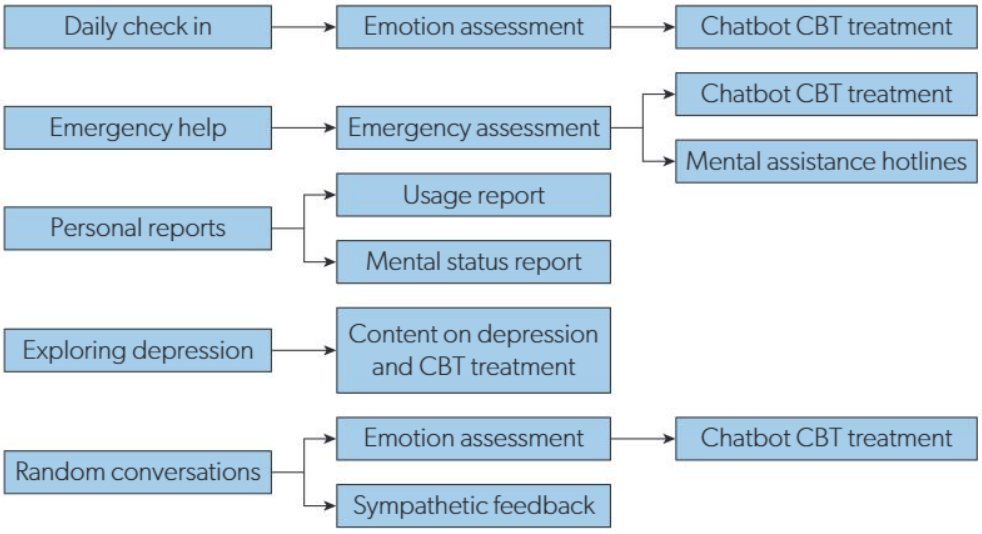
ATL Research, Thinking

these results are only preliminary, but they concluded that the findings "show promise". An app cannot replace a therapist, but using an app could be useful for maintaining the ongoing progress and filling in the gaps where the therapist may not be available.

Note: this study can be seen as a quasi-experiment or even a correlational study, because rather than randomly assigning participants to conditions, researchers used pre-existing differences. Symptoms of depression were established on the basis of an in-app brief self-report questionnaire where participants had to recollect problems over the last two weeks. Moreover, there was no direct comparison between using a CBT app and using a real psychotherapist. Instead those who used the app a lot were compared to those who barely used it.

Liu et al. (2022) conducted a randomized trial of the effectiveness of using AI chatbots to provide self-help depression interventions for Chinese university students. They were especially interested in long-term effects. Participants (83 university students) were randomly assigned to either a chatbot intervention or a **bibliotherapy** control group. Participants were given a set of questionnaires at baseline and every four weeks afterwards for a period of 16 weeks.

A therapy chatbot named XiaoNan was developed specifically for the trial. The chatbot was deployed on WeChat (a popular Chinese social network). Chatbot input could be delivered either by typing or voice. The chatbot was pipeline-based, which means that there existed a pre-programmed sequence of conversational actions that the bot was instructed to follow (see Figure 4.11). It was designed in consultation with professional therapists based on principles of CBT. The main purpose of the chatbot was to alleviate depression. It worked by helping the user separate their emotions from their thoughts and behaviour, and to replace the negative automatic thoughts with new more realistic ones, mimicking a real-life CBT session.



▲ Figure 4.11 Pre-programmed sequence of conversational actions in the XiaoNan chatbot

Bibliotherapy is a form of self-help therapy where patients are prescribed literature to read after a conversation with a professional. Bibliotherapy can be seen as a minimal intervention technique—the therapist only plays a role at the start and the therapy is mostly self-paced and self-driven.

In the study, the experimental group was instructed to use the chatbot for 16 weeks. The control group was given a self-help book to read and to follow its advice for the same period of time.

No significant group differences were found at baseline. However, participants in the chatbot condition demonstrated a significant reduction in self-reported symptoms of depression in 16 weeks. The effectiveness of the chatbot was also found to be superior to that of bibliotherapy.

It should be noted that, although the study is relatively recent (2022), the technology used in this research is already outdated. Conversational AI and chatbots have made tremendous progress since then, making the interaction between AI and the user more believable, more human-like, and more universal. The development of chatbots has also become easier and more accessible. Technology evolves at a pace that far exceeds the pace of conducting research studies and publishing them.

Conceptual analysis

Perspective

Psychological treatment refers to a wide range of therapies, some based purely on talk, some on behavioural exercises with carefully selected reinforcements. What is common for all these approaches is that they are non-medical. They are based on the idea that mental disorders can be caused by factors other than biological ones. However, medical and non-medical treatments are not mutually exclusive, and it is not uncommon to use a combination of, for example, a prescription of antidepressants and CBT.

Causality

You already know that causality in mental disorders can be a matter of intricate interrelationships where some factors may:

- cause a disorder directly
- be mediators in the effects of other variables
- be merely correlates.

It is probably true that any given disorder is "caused" not by one but by a whole range of factors. This is why it can be said that many different treatments "work", that is, produce desirable effects. CBT "works", but so does a simple course of bibliotherapy.

Measurement

The real question is not whether or not a therapy "works". There are many aspects of effectiveness that should be considered:

- effectiveness: is the reduction in symptoms that the therapy produces significantly better than that in control groups (e.g., waitlist, placebo)
- relative effectiveness: does the therapy produce better effects on symptoms than other alternative therapies (e.g., CBT compared to medical treatment)
- dimensions of effectiveness: short term versus long term, different levels of severity of the disorder, different indicators of effectiveness (response rate, remission rate, relapse rate), different populations.

TOK

One of the problems of knowledge in natural sciences is known as "underdetermination of theory by evidence". It means that no matter how much scientific evidence we have, there can always be more than one scientific theory that fits the available evidence equally well. For example, today we accept the Big Bang theory as the predominant scientific theory of the origin of the universe. However, there exist other scientific models that do a decent job explaining the available evidence in their own way. Available empirical evidence alone is not sufficient for rejecting one model and accepting another. There are other factors that are at play when some theories gain dominance.

Can we apply the same reasoning to human sciences, in particular, to various theories of mental disorders? Is it a bad thing that we have multiple theories that emphasize different factors of mental disorders (e.g., biological versus cognitive)? Should we be aiming at creating one single objective theory? Would it even be possible?

In addition, it is important to investigate adherence to therapy. Adherence is the extent to which patients see the therapy through to the end. For example, participants in a CBT programme may find it challenging to attend sessions on a regular basis and difficult to complete demanding home assignments. Adherence in a self-help app assisted by AI may be low because there is no control and patients may perceive it as just another app in their phone. In a practical sense, if the therapy is effective but the level of adherence is low, it may not be the best choice. Beyond adherence, there are also considerations of cost and accessibility.

Bias

As usual, one needs to be careful about treatment–aetiology fallacy. All the usual biases and threats to validity of RCTs also fully apply here.

Change

All considerations of change that we analysed in the previous units also fully apply here. One unique consideration is the fact that the agent of change in psychological treatment is the client himself or herself. Whatever the therapy, it is important that the client takes control and responsibility over their own improvement. CBT is a very effective approach, but it will not work if the client is not willing to change and does not put any effort into the process. This is perhaps in contrast to biomedical treatment where to "comply" with the treatment all you need to do is remember to take the pills. Therefore, the ideas of motivation and agency are much more prominent in psychological treatment. In fact, one of the reasons behind the widespread effectiveness of CBT could be the fact that it gives patients the chance to select their own pace of treatment, set their own goals, and experience small successes already at the start of the process.

Responsibility

The ethics of the working relationship between a therapist and a client is regulated by a special set of standards and guidelines. It is very different from the code of ethics that regulates conducting a psychological research study. In essence, the therapist–client relationship may be seen as a relationship of power—the client is vulnerable and is revealing very sensitive information whereas the therapist is perceived as an authority and sometimes has tremendous influence on the client's decisions. In some cases, further complications emerge. For example, when the client starts having romantic feelings towards the therapist and the therapist may take advantage of this.

4.5 Cultural differences in mental disorders

Inquiry questions

- Are people in some cultures more vulnerable to depression than others?
- Could it be that a person has depression, but due to their cultural values they interpret it as something different?
- Is it possible to eliminate cultural bias from diagnosis and treatment of depression?

What you will learn in this section

Key learning:

- Cultural differences play a role in mental disorders in a variety of ways. For example, aetiology of disorders, cultural expressions of symptoms, or culturally specific responses to treatment.
- Hypothetically cultural values of collectivism may play both a positive and a negative role in depression. The relationship is not obvious.
- According to gene–culture coevolution theory, cultural values of collectivism have evolved in response to the increased prevalence of the short alleles of 5-HTTLPR.
- Reporting bias is when people may be reluctant to report symptoms of mental illness to a professional, due to cultural values. It complicates the study of aetiology of depression.
- When the patient and the psychiatrist belong to different cultural backgrounds, clinician biases may be introduced: it is possible for the clinician to misinterpret the behaviour of the client.
- A diagnosis is often an act of human judgement which may be flawed—for example, due to lack of cultural sensitivity. There have been some attempts to find "objective" markers of depression, such as vocal biomarkers.

Key terms: cultural dimensions, individualism, collectivism, gene–culture coevolution, serotonin transporter gene, reporting bias, patient and clinician variables, vocal biomarkers

In a wider context

Just like any other complex behaviour, mental disorders are caused by a range of biological, cognitive, and sociocultural factors. These factors are not entirely independent from each other, so we might say that mental disorders are influenced by an interaction of multiple factors. For example, genetic inheritance can influence the activity of chemical messengers in the brain, which in turn

can affect attentional biases and patterns of thinking. As you will see, biological factors can also affect the development of cultural values over time.

We have considered the role of biological and cognitive factors in the aetiology of depression. We now shift the focus to investigating the role of culture.

Overview

Cultural differences play a role in mental disorders in a variety of ways:

1. The aetiology of mental disorders: it is possible that people develop symptoms of a mental disorder because they belong to a particular culture and share its norms and values.
2. Expression of symptoms: even when they have one and the same disorder, representatives of different cultures may express its symptoms in different ways.
3. Reporting bias: people in different cultures may differ in their willingness to report symptoms of mental disorders to mental health professionals.
4. Social perception: the perception of mental illness may differ from culture to culture; in some cultures mental illness is stigmatized whereas in others it is not.
5. Diagnosis of mental disorders: cultural differences between the client and the psychiatrist may result in misinterpretation of the client's behaviour and misdiagnosis.

In this unit, we will focus on point 1 above and briefly consider points 2 and 5. Material covering aspects 3 and 4 can be found in Unit 6.8: Culture and mental health.

Discussion

If it is impossible to manipulate culture as an independent variable, does this mean that true experiments in cross-cultural psychology are not possible? Does this mean in turn that we will never be able to make any cause–effect conclusions about culture?

ATL Research, Thinking

Exam tip

Cultural dimensions are discussed in detail in Unit 5.2: Cultural dimensions. See the in-text feature on page 344 for an outline of what you need to know about cultural dimensions for the exams.

Cultural dimensions and the aetiology of mental disorders

The most straightforward way in which culture may influence mental disorders is aetiology: culture can cause the disorders. However, it is also the most difficult relationship to establish. This is because, as you already know, cause–effect inferences require experimental data and true experiments are not always possible in investigating the influence of culture. For example, how do you manipulate culture as an independent variable in an experiment?

It has been observed, however, that there is a significant difference in the prevalence of mental disorders between individualistic and collectivistic cultures. It has also been suggested that the **cultural dimension** of **individualism–collectivism** is implicated in the development of disorders such as depression.

Individualistic cultural values emphasize individual independence, freedom, success, and achievement. In contrast, collectivistic cultural values emphasize social belonging and community networks. They focus on the interdependency of people as opposed to one individual person. Hypothetically, collectivism may be both good and bad for symptoms of depression.

- On the one hand, emphasizing the group over the individual may mean reduced individual identity and sense of self compared to individualistic cultures. If the expression of your unique personality is not encouraged and instead you are being judged by how well you integrate into your group, that may cause an internal conflict and lead to depression.
- On the other hand, members of collectivistic societies have a much better network of social support. Close family ties, cherished family values, extended families, and a sense of belonging to a community may all be very important protective factors in the development of depression. Quite simply, people in individualistic societies frequently have nobody to talk to about their problems.

Gene–culture coevolution in the aetiology of depression

To appreciate the complex interaction between biological and cultural factors in depression, let us consider **gene–culture coevolution** theory (Lumsden and Wilson, 1981).

This theory suggested that human behaviour is a product of two processes: genetic evolution and cultural evolution, and that these processes interact with each other. The theory attempts to arrange a coherent image out of multiple puzzle pieces. Consider the following "puzzle pieces" and think about how you could combine them.

First piece: carriers of the S allele of 5-HTTLPR are more vulnerable to depression

- The **serotonin transporter gene** contains a region known as 5-HTTLPR which may come in two versions: a short allele (S) and a long allele (L). This region regulates the activity of serotonin in the brain. Research has shown that having the S allele is associated with increased risk of depression (remember the study of Caspi et al., 2003).

Second piece: the S allele is more prevalent in East Asia

- Evidence also shows that there is a higher prevalence of the S allele of the 5-HTTLPR in certain geographical regions, most notably in East Asia. A total of 70–80% of the population of East Asia are carriers of the S allele, compared to only 40–45% of the population in Europe. Genetically speaking, this makes people in East Asia more susceptible to depression.

Third piece: prevalence rate of depression in East Asia is low

- Research shows that prevalence rates of depression in East Asian populations are consistently lower than those of Western populations.

Taken together, these three pieces of the puzzle are confusing. East Asians carry more S alleles of the 5-HTTLPR genotype which makes them more vulnerable to depression, and yet there is less depression in East Asian nations compared to the West. Gene–culture coevolution theory attempts to explain this contradiction. The theory suggests that cultural norms themselves are adaptive and have evolved in response to environmental demands.

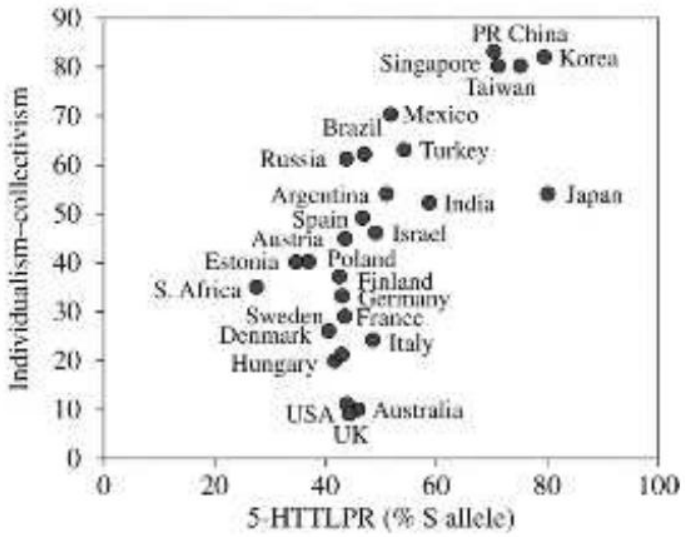
Fourth piece: more pathogens = more collectivism

- Historically, individuals in geographic regions with a larger number of pathogens (e.g., malaria, typhus, leprosy) were more likely to endorse collectivistic values (as demonstrated by Fincher et al., 2008). Collectivistic values themselves might have evolved as an adaptation to a large number of pathogens. Collectivism is characterized by values of tradition as opposed to risk-taking in exploring new things. When you apply this to behaviours such as cooking food, it becomes clear how conforming to tradition may ultimately protect you from disease (if you use well-known ingredients and recipes, you are less likely to get poisoned).

Fifth piece: larger prevalence of the S allele = more collectivism

- Chiao and Blizinsky (2010) suggested that cultural values of collectivism have evolved in response not only to increased pathogens, but also to increased vulnerability to depression in S-allele carriers. According to them, collectivism also serves the function of providing a protective buffer against mental illness. By promoting values of social harmony and group belonging, it reduces the risk of exposure of group members to chronic life stress. When people are part of a close community, they always have the social support necessary to give them protection against depression.

To support their hypothesis, Chiao and Blizinsky (2010) conducted a study of 50,000 individuals living in 29 countries. They demonstrated that collectivistic cultures were significantly more likely to carry the short allele of 5-HTTLPR. Interestingly, the other cultural dimensions (e.g., power distance, uncertainty avoidance, or masculinity–femininity) did not correlate with the frequency of the short allele in the population. See Figure 4.12.



▲ Figure 4.12 Results of correlational analysis between individualism–collectivism and frequency of short allele carriers of the 5-HTTLPR

TOK

Every area of knowledge uses some kind of interplay between deductive reasoning and inductive reasoning in the acquisition of knowledge. They are rarely used exclusively.

It is sometimes said that natural sciences are mainly inductive because they are based on conducting experiments, observing results, and generalizing these results into scientific laws. However, experiments themselves are usually guided by prior theory—we test a certain hypothesis because we expect this test to bring us useful knowledge, based on our prior theoretical understanding. Therefore, the formulation of a hypothesis is often a deductive process. One area of knowledge that is almost purely deductive is mathematics. Do you agree?

Looking at gene–culture coevolution theory, what do you think is the role of deductive reasoning in its creation and testing?

Chat with AI

Gene–culture coevolution theory is a great illustration of the complexity involved in establishing the causality of mental disorders. Although we traditionally separate the causal factors into biological, cognitive, and sociocultural, these factors themselves are not independent from each other. Since they can interact, asking which one of them is "more important" is probably inappropriate.

However, gene–culture coevolution theory may not be easy to understand. Use your favourite generative AI to check your understanding.

- I am a psychology student learning about the role of culture in depression. Here is a summary of my understanding of gene–culture coevolution theory and how it applies to the development of depression: [provide a summary]. Please check my writing, correct my mistakes, and give me suggestions on what can be improved to make the summary more accurate. Am I missing anything?
- One thing I don't understand about this theory is [provide information]. Could you explain?
- Could you check my understanding of the theory by asking me three questions that require short answers? After I provide responses, please give me feedback and explain to me what I need to understand about the theory to fill in gaps in my knowledge.

Reporting bias and prevalence rates

Kleinman (1977) made a distinction between disease (a biological malfunction) and illness (one's personal reaction to the disease). Unlike disease, illness is heavily influenced by culture.

An important problem that complicates our understanding of cultural aetiology of depression is **reporting bias**. Reporting bias is the phenomenon when people may be reluctant to report symptoms of mental illness to a professional, due to cultural values. If this is the case, depression will exist within society, but it will not get reported. Therefore, "officially" it will appear as if prevalence rates of depression in this society are low.

Furnham and Malik (1994) investigated cross-cultural beliefs about depression. Their participants were a group of middle-aged women (aged 35–62) and a group of younger women (aged 17–28). Half of the participants were classified as native British. The other half were of Asian origin and had received education in their countries of origin. All participants filled out questionnaires that targeted both their symptoms of mental illness and their beliefs about depression and anti-depressive behaviours.

It was known from previous observations that British Asians were rarely diagnosed with depression.

The results showed that perception of depression differed among Asian and British participants. For example, Asian (but not British) women tended to agree with the following statements:

- "Having a job outside the home helps keep women from getting depressed."

Activity

Reflect on your own reporting bias. For example, you have a toothache, but you do not report it to the dentist, although you know early intervention is important. Or you have insomnia but decide not to see a doctor about it.

What illnesses would you personally be reluctant to report to a professional? What illnesses go unreported most often in your culture?

For example, how "normal" is it to see a psychologist regarding social anxiety in your society? How many of your friends are currently seeing a psychologist?

ATL Thinking, Self-management

Activity

There are two independent variables in Furnham and Malik's study, with two levels each: culture (native British versus British Asian) and age (middle-aged versus young).

This means that the study followed a 2×2 design. Note, however, that the study is quasi-experimental. Researchers did not randomly allocate participants to the groups, which limits their ability to make cause-and-effect inferences from the study findings.

1. Name the dependent variables and how they were operationalized.
2. Describe the interaction between two independent variables. Use this format: "A report less depression and use more collectivistic explanations of the illness than B, but only if A are _____ (younger/older)".

ATL Research

- "When feeling depressed, it is more helpful to talk it over with a family member than with a friend."
- "Feeling depressed is no different from feeling depressed about something."

Interestingly, the differences were less pronounced in the group of younger women. This shows how globalization gradually influences younger representatives of traditional cultures. Asian middle-aged women reported being depressed significantly less often than the other three samples. According to researchers, this could be explained by reporting bias due to their (culturally influenced) perceptions of depression.

Culture and diagnosis of mental disorders

Differences in prevalence rates may be a result of differences in how symptoms of mental disorders are expressed and presented to psychiatrists. You might call such differences "**patient variables**". However, it is also conceivable that it is the psychiatrists who introduce cultural biases in the interpretation of their patients' behaviour. You might call such differences "**clinician variables**". In reality, both patient variables and clinician variables probably have an effect.

Even if clinicians are culturally competent, a diagnosis is always an act of interpretation. As with any human judgement, it may be flawed by the clinician's own cultural background or a mismatch between their culture and that of the client.

This is why it would be useful to have an impartial, objective measure of depression. Such a measure does not exist but attempts to find one have been made. One of such attempts is by Alghowinem et al. (2016) and their study of **vocal biomarkers** of depression. Researchers used three datasets from different cultures and languages. Each dataset contained video recordings of interviews with patients. Datasets came from different cultures: one Australian, one American, and one German. Interviews in the German dataset were conducted in the German language.

Since researchers wanted to find a universal measure that would be independent of culture and language, they focused on paralinguistic features—acoustic cues such as intonation, frequency, intensity, loudness, shimmer. They extracted 504 features per participant. Data analysis techniques (machine learning) were applied to select the most informative features and estimate how closely they are associated with group membership (depressed or not).

Statistical models constructed in this study were able to achieve an accuracy of classification around 75%. People with severe depression do indeed go through certain changes in the paralinguistic features of their voice, and such changes may be used quite reliably to tell the difference between normal functioning and severe depression. Remember that participants came from different cultures and spoke different languages, so this accuracy was achieved by an algorithm that was applied to a very diverse dataset.

This study shows that paralinguistic features of a person's voice may be one of the "objective" characteristics of depression, the use of which may help us overcome multiple cultural biases existing in this area. This also serves as an example of a breakthrough in research and practice that only becomes possible with the advancement of technology (data analysis and machine learning in this case).

Conceptual analysis

Causality

Whether or not culture influences the development of mental disorders is a question of aetiology. It is difficult to answer. There have been abundant observations confirming that prevalence rates of mental disorders (i.e., depression) are different depending on culture. However, this finding does not mean that culture influences mental disorders.

Cultural values themselves might have evolved in response to environmental demands, as suggested by gene–culture coevolution theory. Cultural values of collectivism could have evolved as a protective factor against increased vulnerability to stress due to a higher prevalence of the S allele of 5-HTTLPR. If that is the case, then cultural values are actually a consequence of depression rather than its cause.

Measurement

Figuring out the role of culture in the development of mental disorders is complicated. The observed difference in prevalence rates from one culture to another can be explained by several reasons other than cultural factors actually causing the disorder. For example, cultural differences in the presentation of symptoms, cultural or racial biases of clinicians who establish the diagnosis, and differences in the way people from various cultures interpret their own symptoms.

It is further complicated by the fuzzy boundaries of the concept "culture" and the fact that this concept is usually approximated by a country or a nation in research. It is assumed that people coming from a particular country all share a particular culture, although this does not necessarily have to be the case. Many studies have revealed that effects of culture appear more pronounced with older people, presumably because senior generations are less globalized and share more of the traditional values of their place of origin.

Bias

Sampling biases is a huge factor contributing to the risk of bias in this field of research. This is because we always compare groups of people and interpret differences between these groups as differences between "cultures". But how do we select these groups? Sometimes samples are recruited from particular countries, based on the assumption that most members of the population in each of these countries share country-specific cultural norms and values. Sometimes samples are recruited in the same location based on country of origin—for example, international university students. This is even trickier because we are now adding the assumption that while these people lived and studied abroad, they did not lose the identity linked to their culture of origin.

Reporting bias is a problem because people may experience depressive symptoms but fail to report them.

Change

In investigating the effects of culture on mental disorders, we must remember that things change. For example, cultural values change with the course of time, and there is an increasing amount of interpenetration of cultures due to migration and globalization. In other words, the effects of culture on mental disorders are not

static. They change over time both on the level of a separate individual and on the level of society as a whole.

The concept of change is also relevant to treatment. Once a person has developed symptoms of a mental disorder, we want to change that. The process of change should be culturally sensitive. Treatment approaches that work with representatives of one culture may not work with representatives of another.

Responsibility

Conclusions derived from research in this area may have very large implications for public health policy and society in general. For example, if we acknowledge reporting bias, we will have to conclude that people in some cultures are more depressed than we used to believe. We may need to develop public health policies that address this.

Cultural sensitivity of the clinician is hugely important. Clients from different cultures may experience and interpret depressive symptoms very differently (disease versus illness). Do we respect the client's cultural interpretations of what's happening, or do we try to impose our own interpretation and try to make them accept it? There are ethical consequences to both of these decisions.

Perspective

What was discussed in this section is most pertinent to the sociocultural approach to understanding behaviour. However, we have also seen that perspectives are only clearly separated in theory but not that clearly separated in reality. For example, the idea of gene–culture coevolution suggests that genetic evolution (biological) and cultural evolution (cultural) may be closely intertwined, and that biological factors may influence sociocultural factors. If we dig even deeper, we will probably discover that the mechanism of such influence is cognitive.

4.6 Environmental factors in mental disorders

Inquiry questions

- Can stress in early childhood make us more likely to become depressed as adults?
- Can parents (nurture) change our biological predisposition to depression (nature)?
- Can symptoms of depression spread from person to person like symptoms of flu?

What you will learn in this section

Key learning:

- "Social" and "cultural" in sociocultural are sometimes separated for more detailed analysis; in this case social factors are also referred to as environmental.
- Environment can mediate the influence of genetic factors of mental disorders through the epigenetic process of regulation of gene expression. Genes themselves cannot be changed by environmental influences, but their expression can be.
- Stress experienced in early childhood can affect the expression of the gene coding for the serotonin transporter. This can affect the amount of serotonin in the brain and increase vulnerability to depression.
- Genetic predisposition and early stress appear to have an additive influence on depression. This is different from an interactive influence.
- There are a number of social vulnerability factors that increase the risk of depression. Among the most important of them are lack of intimate relationships and lack of employment.
- Symptoms of depression can spread across the social network of relationships similar to how a virus spreads in a population.

Key terms: social environment, environmental factors, epigenetics, gene expression, transcription, translation, regulation of gene expression, methylation, epigenetic change, serotonin transporter, additive influence, interaction, vulnerability factors of depression, person-to-person spread of depressive symptoms

In a wider context

The term "sociocultural" is often understood in a broad sense to include any factor that forms a part of the social or cultural environment surrounding the individual.

The key components of the cultural environment are captured in Geert Hofstede's cultural dimensions such as individualism–collectivism and masculinity–femininity (see a detailed discussion of Hofstede's theory in Unit 5.2: Cultural dimensions.

Social environment includes factors such as: socio-economic status (SES), employment, education, lifestyle, the neighbourhood a person lives in, population density, social class, occupation, housing, and religion. One may even argue that such factors as air pollution are also social factors because more polluted areas tend to be inhabited by people with lower income. For example, a negative correlation has been reported between family income and levels of lead in children's blood (Brooks-Gunn and Duncan, 1997). This is because in poorer families children who live in old housing are exposed to deteriorating lead-based paint to a greater extent.

For the sake of more detailed research "cultural" and "social" in sociocultural factors are sometimes separated. When separated this way, what we mean by **environmental factors** is any component of the environment except for those that may be directly attributed to culture.

Environmental factors and epigenetic influences

You know already that serotonin is a neurotransmitter that is implicated in mood regulation and has been associated with depression. An important protein that is involved in the process of reuptake of serotonin from the synaptic gap is the "**serotonin transporter**". Not enough serotonin transporter means not all serotonin is taken back, which means a higher concentration of serotonin in the synapse.

Where does serotonin transporter come from? Like all proteins in our body, it is synthesized according to the programme defined by our genes. There is a section of the DNA that is responsible for that—the serotonin transporter gene. This gene comes in three different versions depending on the combinations of alleles in its structure. You might recall from Caspi et al.'s (2003) research that people with two short alleles in this region (S/S) have been found to be more vulnerable to stressful life events and therefore more susceptible to depression than people with other combinations (S/L or L/L).

But the story does not end here. The serotonin transporter gene is merely a genetic programme, a plan written in the language of DNA. How does this plan get manifested in the actual serotonin transporter molecules? There is a process known as **gene expression**. It consists of two stages:

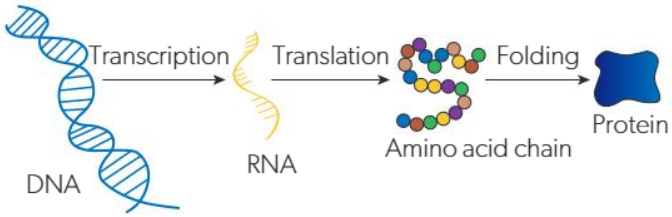
- **Transcription**: here the sequence of the gene is copied to make an RNA (ribonucleic acid) molecule.
- **Translation**: here the RNA molecule is decoded into a sequence of amino acids in a protein.

Both of these stages can be affected by environmental factors. For example, environmental variables can cause some chemicals to bind onto the DNA molecule and suppress the transcription of certain genes. This process when chemicals are added to the DNA molecule is known as **methylation**.

Exam tip

Epigenetics (regulation of gene expression) is related to the topic "genetic inheritance". You can use content related to epigenetics to support your answers about genetic inheritance.

Of course, ideas of epigenetics are also a great illustration of complexities involved in causality. Regulation of gene expression is an example of how the environment can mediate the influence of biological (genetic) factors on behaviour.



▲ Figure 4.13 Gene expression

The process of gene expression is complex, and environmental factors may influence this process. Having a gene does not automatically mean that this gene will be manifested in the phenotype. Although the genotype itself cannot be changed by environmental factors (apart from cases of mutation—for example, caused by radiation), gene expression can be changed. If this happens, the gene is there, but it does not get translated into a protein. Any changes in the expression of a gene induced by environmental factors are known as **epigenetic change**. Here the environment affects behaviour by suppressing the expression of genes.

Epigenetic influences in depression: Wankerl et al. (2014)

Wankerl et al. (2014) studied the effects of genetic and environmental risk factors on the expression of the serotonin transporter gene. They recruited 133 German participants via advertisements. Participants went through a structured clinical interview assessing prevalence of mood disorders based on DSM-IV. They filled out a set of questionnaires on early and recent stress and trauma.

- One of the questionnaires (the Childhood Trauma Questionnaire—CTQ) measured such factors as emotional abuse, physical abuse, sexual abuse, emotional neglect, and physical neglect in early childhood.
- Another questionnaire was used to collect information about recent stressful events (death of a relative, serious accidents, abortion, etc.).
- A survey of prenatal stress was given to participants' mothers. This survey included questions about such factors as divorce, lack of social support, relationship conflicts, high workload, financial difficulties during pregnancy.

At the end of the session, blood samples were taken from the participants and DNA was extracted. Methods of molecular genetics were used to study gene expression and methylation patterns in the region of the DNA known as SERT (this region codes for the production of serotonin transporter).

Here are the results of the study:

- Participants whose mothers reported having at least one major stressful or traumatic life event during pregnancy were found to have lower levels of SERT expression compared to those without prenatal stress. This was further corroborated by a correlation—the larger the number of stressors experienced by the participant's mother during pregnancy, the lower the level of SERT expression. Maternal stress inhibited the translation of the serotonin transporter gene prenatally (and it remained like this throughout the participant's life).

Chat with AI

Consolidate and check your understanding of basic principles of epigenetics in interaction with your favourite conversational AI. Here are some examples of prompts you can use:

- "Here is a paragraph that provides a summary of how I understand the process of gene expression [insert your paragraph]. Could you check it and let me know if there are any mistakes, inaccuracies, or gaps in my understanding? Note that it is for a school psychology course (not biology), so I only need a basic understanding of the process."
- Do the same exercise to check your understanding of the role of environmental factors in regulation of gene expression.
- "Based on our previous conversations, could you ask me three questions to check my understanding of epigenetics? After I provide my responses, please rate them and give me feedback for improvement."

Exam tip

There is a lot of overlap between this topic and another topic from Unit 3.6: Role of childhood experiences. Conceptual understandings from these two sections complement each other.

- Similar reduction in the expression of SERT was found in individuals who reported early life trauma (e.g., a history of abuse). Again, this was corroborated by a correlation: the higher one's score on the CTQ, the lower the level of expression of SERT.
- In contrast to these findings, recently experienced stress had no effect on the expression of SERT.

We can consider short alleles of the 5-HTT gene and adverse early experiences (e.g., prenatal stress and child maltreatment)—two major risk factors in suppressing the expression of SERT. Wankerl et al. (2014) found that there is an additive relationship between these two factors. This means that the most adverse effect is observed when the two factors are combined. For example, participants with both short alleles of 5-HTT who also experienced maltreatment in early childhood had 56% lower SERT expression than individuals with long alleles of 5-HTT and without a history of child maltreatment.

Activity (HL only)

The difference between interaction of factors and additive influence

Wankerl et al.'s (2014) study found evidence for the existence of an additive relationship between genetic predisposition and early stress. However, evidence in this study did not support the existence of interaction between the two factors.

What is the difference between **additive influence** and **interaction**? Consider the following two hypothetical scenarios:

Hypothetical scenario 1: interaction

In this scenario, the effect of childhood maltreatment on depression depends on the genotype. Individuals with long alleles of 5-HTT are more resistant to stress, so childhood maltreatment does not affect them as much. Individuals with short alleles run approximately the same risk of developing depression when there is no history of childhood maltreatment. However, they are less resilient. In other words, the extent to which childhood maltreatment will affect their vulnerability to depression depends on their type of 5-HTT. In this hypothetical example, the two factors interact.

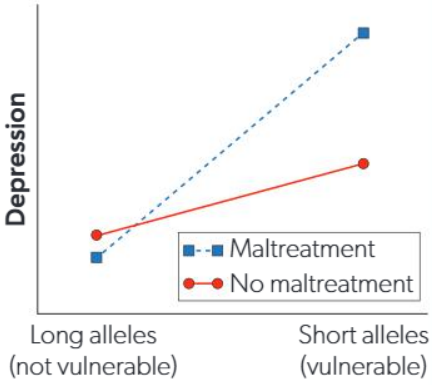
Hypothetical scenario 2: additive influence

In this scenario, the lines are parallel to each other. The more childhood maltreatment there is, the more likelihood of depression, and it is worse if the individual's 5-HTT alleles are short.

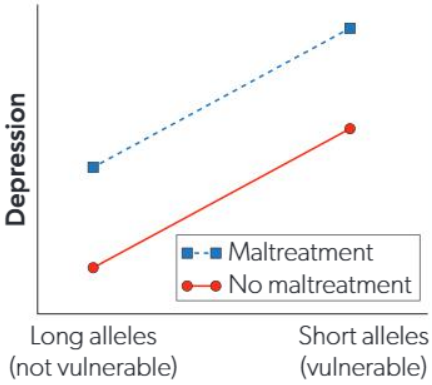
Can you think of other real-life examples of additive influence versus interaction?

ATL Research, Thinking

Interaction



Additive influence



▲ Figure 4.14 Examples of additive influence and interaction

Vulnerability factors in depression

You have seen that epigenetic changes related to the development of depression may be triggered by stressful life experiences, especially if they occur early in life. Many different life experiences may be perceived as stressful, so it raises the question: what are the key factors in a person's environment that make him or her particularly vulnerable to depression?

Brown and Harris (1978) proposed a model of depression which outlines "**vulnerability factors of depression**". They reported results of a community study of 458 women from London who were surveyed on the history of life events and depressive episodes. Semi-structured interviews were used to gather in-depth information. The overall finding was that the following four vulnerability factors, when combined with acute or chronic social stressors, were likely to provoke depression in women:

1. Three or more children under the age of 14.
2. Lack of an intimate relationship with a husband or boyfriend.
3. The lack of employment.
4. The loss of a mother before the age of 11 years.

Of course, the gender-biased sample makes it impossible to generalize the findings to men. However, the study was replicated several times, and the results broadly supported the notion that the four vulnerability social factors are associated with the development of depressive episodes. In particular, Patten (1991) summarized the quantitative results of replications and concluded that the lack of an intimate relationship increases the risk of developing depression 3.7 times, whereas each of the other three factors "only" doubles the risk. This increase of risk is comparable, for example, to smoking as a risk factor for atherosclerosis (Patten, 1991).

Discussion

The sample composition in Brown and Harris's study is a clear limitation. Why do you think the researchers made the decision to limit their sample to working-class women from London?

ATL Research, Thinking

Spreading of symptoms along social networks

There have been claims that some mental disorder symptoms may spread from person to person along the network of interpersonal relationships (much like an infectious disease).

Rosenquist, Fowler, and Christakis (2011) investigated the possibility of **person-to-person spread of depressive symptoms**. Data for this research were obtained from Framingham Heart Study (FHS), a population-based longitudinal study of 12,067 individuals investigating risk factors for heart disease. To keep track of participants, FHS interviewers documented information on the participants' friends (based on their self-reports), neighbours (based on their address), co-workers (based on their job), and relatives.

This information was collected longitudinally. Researchers computerized these records and obtained longitudinal information on the development of social network structure. To assess symptoms of depression, a standardized depression scale was administered three times between 1983 and 2001.

Results showed that there was a significant correlation in depressive symptoms between people up to three degrees of separation (a person's depression

TOK

What is the relationship between knowledge and the methods we use to obtain it? According to one argument, knowledge is determined and limited by the method. Think about the study of microbiology at a time when the microscope was not invented. This entire area of study did not even exist! Creating the new observation method made it possible to study tiny living things. Think about our knowledge of celestial bodies at the time prior to the invention of the telescope, or our knowledge of the history of the universe without the existence of mathematical modelling.

Looking at the various studies considered in this section, do you think the argument holds true here also? Could it be that we cannot reach definitive conclusions about the origin of mental disorders at the moment because the real ground-breaking method of study has not been invented yet?

depends on the depression of their friend, their friend's friend, and their friend's friend's friend). Participants were 93% more likely to be depressed if a person they were directly connected to was also depressed. Participants were 43% more likely to be depressed if depression was observed in the friend of their friend (two degrees of separation), and 37% for the third degree of separation.

Changes in social ties predicted changes in depressive symptoms, but changes in depressive symptoms did not predict changes in social ties. In other words, acquiring new friends can influence you to become depressed, but becoming depressed does not come with acquiring new friends. Interestingly, directionality of friendship also appeared to be important. For example, in a couple where A nominated B as a friend but not vice versa (A→B), if B becomes depressed it doubles the chances of A also becoming depressed in the near future. If A becomes depressed, it has no effect on B. In contrast, in a mutual friendship (A↔B), when B becomes depressed it increases the chances of A becoming depressed by 359%.

Discussion

Rosenquist, Fowler, and Christakis (2011) conducted a correlational study (they did not randomly allocate people to be friends). As we know, this means that cause-and-effect inferences cannot be made. However, their study was longitudinal and they looked at how the variables were changing dynamically. For example, they looked at how acquiring a new friend today correlates with developing depressive symptoms several years later. Does this change anything in terms of our ability to make cause-and-effect inferences?

ATL Research

Conceptual analysis

Perspective

In this section, we separated the sociocultural perspective on human behaviour into environmental (social) and cultural. This was done to underline the unique importance of such causes of mental disorders as adverse childhood experiences (e.g., childhood maltreatment) and social stressors (e.g., unemployment, poverty).

As always, this separation is somewhat artificial. It is done for the purposes of research because to understand something we need to analyse it, and analysis literally means separating into components and looking at them individually. In reality, multiple factors act together. For example, cultural factors probably underpin many of the social variables whereas social factors have an impact on the development of cultural norms. At some point, after we have acquired sufficient knowledge about the component parts, we should assemble them back together (synthesize) and understand how they interact.

Causality

There is no doubt that environmental factors such as adverse childhood experiences and poverty influence the development of mental disorders. Research efforts in this area have mostly concentrated on two questions:

1. Which of the numerous environmental factors are the most significant?
2. What exactly is the pathway through which such influence occurs?

The study of Wankerl et al. (2014) is a great example of how complicated this pathway may be. It is possible that childhood experiences influence the expression of the serotonin transporter genes through epigenetic mechanisms (such as methylation), which in turn influences reuptake of serotonin in synaptic gaps. Here environmental variables serve as a mediator in the influence of genetics on behaviour. However, we also know from such studies as Brown and Harris (1978) that environmental factors could influence depression directly, by exposing a person to chronic stress.

Bias

The usual considerations apply here. Research studies may be biased due to various considerations of research methodology (sampling, control of variables, construct validity, etc.) and our conclusions may be biased due to a misinterpretation of research results.

One prominent consideration is the difference between effect sizes and statistical significance. There is a wide variety of social (environmental) factors that influence mental disorders. Research will show that quite a lot of these factors have a statistically significant impact on mental disorders. However, "statistically significant" does not always mean "large". For more detail on this see Unit 1.5: Correlational studies.

Measurement

The task of measuring the effects of environmental variables on mental disorders is challenging on many levels.

You already know that measuring mental disorders is not easy. Measuring environmental variables presents various challenges also. For example, childhood maltreatment can only be established through self-report. It is tricky to quantify the severity of it. Asking participants to respond to questions that trigger them to relive those adverse experiences could be an ethical consideration. Factors such as poverty are easier to measure (e.g., family income), but the challenge could be timing: poverty experienced at one age may not have the same effect as poverty experienced at another age.

Finally, measuring the effects of environmental variables on mental disorders may be challenging because of the many intricate pathways that these effects may follow. What appears to be a social influence may turn out to be epigenetic. What appears to be genetic may turn out to be cultural.

Change

An investigation of environmental variables in the aetiology of mental disorders is important in the practical sense because environmental variables are something we can have some control over. Arguably, knowledge of environmental factors is more important for policymaking than any other group of factors. For example, if we know that living below the poverty line at a certain age dramatically increases the chances of developing a mental disorder later in life, then we may want to design some policies to address the poverty problem, prioritizing families who have children of that age.

Mental disorders themselves are a social problem because people who suffer from mental disorders cannot fully contribute to society and require expensive

help. Understanding cycles of causality (such as unemployment causing depression, but depression preventing people from finding a job) may help us eventually break these cycles.

We have also considered the process of epigenetic change—changes in gene expression induced by environmental factors.

Responsibility

Regular ethical considerations in conducting a research study apply. Experiments are not common in this area, but there are a lot of correlational studies and self-report measures. Since the study of environmental variables could include sensitive topics such as poverty or childhood abuse, care should be taken about protecting participants from possible harm associated with reliving negative experiences.

Ethical considerations should also be taken into account in the application of research findings. We should:

- interpret results carefully and precisely, understanding their constraints and avoiding unjustified generalizations
- derive our conclusions based on research programmes rather than individual studies
- understand and tolerate an extent of uncertainty that is always present in human sciences.

4.7 Prevalence of health problems

Inquiry questions

- How prevalent in the population are health problems such as social media addiction and obesity?
- How do we measure the prevalence rate of a health problem?
- How do prevalence rates change from population to population and how do they change over time?

What you will learn in this section

Key learning:

- A key challenge in measuring prevalence is that it is a metric that is applied to the entire population, and yet we only have access to a limited sample of people. An added challenge comes from the fact that criteria and thresholds of having a health problem may be defined differently by different researchers.
- Prevalence rates of social media addiction reported for various populations have differed massively (4–40%). This may be caused by genuine variations between populations or by different criteria used for diagnosing social media addiction.
- Meta-analyses are very important in untangling this variation and establishing unbiased prevalence estimates.
- A recent meta-analysis established an average prevalence rate of 24% for moderate-to-severe social media addiction, but also revealed considerable variation depending on: diagnostic criteria, cultural collectivism, cultural individualism, questionnaire quality, and type of sampling used.
- Investigating how obesity changes over time is difficult because it is usually a mixture of three effects: age effects, period effects, and cohort effects.

Key terms: point prevalence, period prevalence, lifetime prevalence, representative sample, social media addiction (salience, tolerance, mood modification, relapse, withdrawal, conflict), BFAS, BSMAS, polythetic classification, monothetic classification, obesity, body mass index (BMI), age effects, period effects, cohort effects

In a wider context

The name of this chapter is "Health and well-being". So far, we have discussed mental disorders, focusing on one example: Major Depressive Disorder (MDD).

However, there are other health problems such as obesity, social media addiction, or smoking that, while they are not mental disorders, play an important role in daily functioning. These problems are not medical. They are determined by a person's health behaviours.

Additionally, it is currently recognized that mental well-being is not simply absence of disease. If a person does not have any mental disorders, it does not mean that this person is well. A study of what it means to be well, and what kind of behaviour contributes to being well, is both theoretically and practically significant.

What are prevalence rates?

Prevalence is defined as the proportion of a population who have a specific health problem in a given time period. The term "prevalence" comes from epidemiology where it is used for any illness or disease. It is usually reported as a percentage (e.g., 5% of the population).

There are three types of prevalence that are commonly used:

1. **Point prevalence** is the percentage of the population that has the health problem at a given point of time.
2. **Period prevalence** is the proportion of the population that has the health problem at any point of time during a given period. A 12-month prevalence is most commonly used, which is the percentage of the population who had the health problem at any point during the last 12 months.
3. **Lifetime prevalence** is the percentage of the population who have had the health problem at any point in their life.

Challenges in measuring prevalence rates

The challenge in measuring prevalence is that it is a metric that is applied to the entire population and yet we usually only have access to a limited sample. To estimate prevalence, we need to simply divide the number of people in the sample who have the health problem by the total number of people in the sample. To ensure that this estimate can be generalized to the entire population, we must ensure that the sample is **representative** of that population. This is especially tricky because the "population" in question is often an entire country.

To ensure the sample is representative of the entire population, researchers usually use a random sampling strategy. When every member of the population has the same chance of being included in the sample, we can expect that the sample will reflect all characteristics of the population, especially if the sample size is sufficiently large.

In psychology, an added challenge of measuring prevalence (and especially comparing it across studies or populations) is that whether or not the person has the health problem depends on how we decide to diagnose the problem. The outcome variable used in prevalence is binary: whether a person has obesity or not, whether a person has social media addiction or not. How we define it depends on the criteria and where we have set the threshold. If different studies use different approaches, then these studies become non-comparable.

Activity

Go back to Unit 1.3: Analysing research (methodological considerations) and refresh your memory on the types of sampling techniques that can be used in research. Why is the random sampling strategy preferred in this case? How can you carry out random sampling in practice?

ATL Research

Diagnosing social media addiction

Social media addiction is not currently recognized as a mental disorder. Therefore, there are no diagnostic criteria for it in the DSM (nor other diagnostic manuals). We refer to it as a health problem.

There exist several measures of **social media addiction**, but the Bergen Facebook Addiction Scale (**BFAS**; Andreassen et al., 2012) is the most commonly used. It is built on the diagnostic criteria of a broader condition recognized by the DSM-5: behavioural addiction. In line with those criteria, the scale measures six components of social media addiction:

1. **Salience**: spending too much time on social media.
2. **Mood modification**: using social media to cope with negative emotions or as a way to distract from problems.
3. **Tolerance**: feeling the urge to spend more time on social media to maintain the same level of pleasure.
4. **Withdrawal**: feeling uncomfortable and restless when unable to use social media for a period of time.
5. **Conflict**: disruptions in work, life, or interpersonal relationships brought about by the use of social media.
6. **Relapse**: being unable to give up or limit the use of social media despite trying to do so (Andreassen et al., 2012).

The scale consists of 18 items, three per component. There also exists a shortened six-item version of the scale, one question per component. Questions are scored on a 5-point Likert scale (Very rarely–Rarely–Sometimes–Often–Very often).

Following the launch of Facebook, the world saw a boom of social media platforms and many people now use more than one. The scale was modified accordingly and renamed the Bergen Social Media Addiction Scale (**BSMAS**; Andreassen, Pallesen, and Griffiths, 2017).

BFAS and BSMAS have been widely used in studies aiming to determine the prevalence of social media addiction in a specific population. However, the result was a very large variance of findings, from 4% to >40% estimated prevalence of social media addiction. This is clearly a problem. Such variation in estimates may be due to some real-life factors (e.g., one population being generally more addicted to social media than people in another population) or it could be due to how the scale was applied in a particular study (e.g., different cut-off points).

Andreassen et al. (2012) suggested two diagnostic criteria based on BFAS:

- "**polythetic classification**" where the cut-off point is 3 (on a 5-point scale) on at least two-thirds of the items
- "**monothetic classification**" where the cut-off point is 3 out of 5 on all items.

This means that, using Andreassen's suggestion, you would be classified as a person having social media addiction if you agreed or selected the neutral option in at least two-thirds of the questions in the questionnaire (polythetic classification) or all the questions (monothetic classification).

To make matters more complicated, some researchers have maintained that the cut-off score of 3 out of 5 is too liberal and used 4 out of 5 instead.

Based on such a variety of diagnostic criteria, one would expect considerable variation in estimates of social media addiction. However, another source of variation could be actual differences in rates of addiction from one sample

Chat with AI

If you want to gain a deeper understanding of what is meant by social media addiction, consider asking your favourite generative AI for clarifications.

- Here are the six components of social media addiction as described by Andreassen [insert the description]. Could you provide a case study for me to determine if the case qualifies as social media addiction or not?
- [After your response] Was I correct? Could you give me feedback?

TOK

All human sciences (including psychology) use statistical reasoning, and one of the key concepts is the level of statistical significance. There is an agreed-upon cut-off score of $p < 0.05$, which means a 5% probability of random error. If the probability of obtaining a result in a study by random chance is less than 5%, we consider this a statistically significant result. If it is over 5%, we accept it as a non-significant result and reject our hypothesis. Is there any good reason this cut-off has been set at 5%? Why not 6% or 7% or 10%?

Similarly, is there any good reason for the cut-off score on a scale of social media addiction to be what it is? Is it absolutely arbitrary?

What other examples of using somewhat arbitrary cut-off scores can you think of? Are different areas of knowledge equally vulnerable to biases related to the arbitrary nature of these cut-off points?

to another. For example, we may expect a difference between collectivistic and individualistic societies (using Geert Hofstede's theory—see Unit 5.2: Cultural dimensions). This is because people in collectivistic societies are more interdependent and place more importance on their social environment. Theoretically, one could therefore expect them to be more susceptible to the fear of missing out and to developing an addiction to social media.

Meta-analysis is the only method that would allow us to separate these two sources of variance: differences in the way diagnostic criteria were applied and actually existing differences between populations.

Prevalence rates of social media addiction

Cheng et al. (2021) conducted a meta-analysis of prevalence of social media addiction in 32 nations. The reason they started this study in the first place is that prevalence rates reported in various studies varied dramatically, most probably for the reasons already outlined. As with any large meta-analysis, this study also allowed researchers to investigate factors that could potentially explain this variance in reported prevalence rates.

The analysis aggregated 63 studies with a total sample of almost 35,000 participants.

To identify potentially relevant studies, Cheng et al. (2021) conducted a systematic search of 28 electronic databases. They also searched through various conference proceedings and approached study authors to obtain unpublished papers and data.

Many studies included in this meta-analysis (41%) used the original English version of BFAS/BSMAS. The remaining studies used a version translated into one of 13 languages. Most of the studies (74%) examined Facebook addiction, while the other studies examined social media addiction in general.

With all results of all studies combined, the prevalence rate of social media addiction was estimated at 25%. However, in different studies these estimates ranged from 0% to 82%.

Understandably, studies that used a stricter threshold for social media addiction reported a lower prevalence rate than studies with a more relaxed threshold. For example:

- studies that used the monothetic approach (≥3 on all items) reported 5% prevalence rate, on average
- studies that used the polythetic approach (≥3 on at least two-thirds of the questions) reported a 24% prevalence rate.

Cultural variables also appeared to moderate prevalence rates of social media addiction. Prevalence rates reported for individualist societies (North America, Western Europe, Northern Europe) were lower than prevalence rates in collectivistic societies (Africa, Asia, Middle East): 14% versus 31% respectively, on average.

Prevalence rate estimates may also be affected by biases within studies. Low-quality studies that do not strictly control sampling procedures and do not use properly validated measurement tools result in systematically different estimates

Discussion

Is 25% a large number for the prevalence of social media addiction? This essentially means that one out of four people, on average, is addicted to social media.

Do you think one could argue that the criterion is too mild and we need to revise it so that the rate becomes lower?

ATL Thinking, Communication

of prevalence, compared to high-quality studies. The average estimate of prevalence rates in studies that used a properly validated version of BFAS/BSMAS was 19%, while for newly developed translations that were not yet properly validated this figure rose to 28%.

The nature of the study sample also proves to be very important in prevalence research. In particular, some studies employed probabilistic sampling (for example, random or stratified) and others sampled non-probabilistically (for example, opportunity samples consisting of students attending the researchers' university). Probabilistic samples in this meta-analysis yielded an average prevalence rate of 32%, whereas non-probabilistic samples resulted in a 22% estimate.

Activity

Make a poster or an infographic outlining all major problems in measuring the prevalence of a health problem.

ATL Self-management

Diagnosing obesity

The most widely used approach to classifying **obesity** in adults is on the basis of **body mass index (BMI)**. This is calculated as body weight in kilograms divided by height in metres squared. The criterion of obesity that is commonly used worldwide is BMI ≥30 kg/m^2, and the criterion for overweight is BMI ≥25 kg/m^2.

Obesity classification for children is more difficult, however, because body composition fluctuates a lot as a child grows. For this reason, the WHO publishes BMI-for-age references for individuals under 19 years.

Change of obesity over time

The world is becoming increasingly obese and this increasing prevalence of obesity is a global concern. Obesity is associated with a higher risk of several chronic conditions. A number of interventions have been proposed over time, but none of them have yet managed to stop or even slow down the negative trend.

When we analyse the development of prevalence rates over time, there are three sources of change that need to be separated from each other:

1. **Age effects** are changes across the course of one's life. They occur in all cohorts and are independent of time.
2. **Period effects** are changes over time that equally affect all cohorts and all age groups.
3. **Cohort effects** are changes that are associated with being born at a particular point of time (in a particular year).

Discussion

Why is it important to separate the three effects (age, period, cohort)?

ATL Thinking

Imagine you have measured obesity in a group of 20-year-old participants and a year later, measured obesity in a group of 40-year-old participants. You observe some difference which could be due to one of three explanations:

1. It does not matter when they were born and when you tested them, but these samples are different in terms of age (age effects).
2. It does not matter what age they are or when you tested them, but these samples were born at different times and belong to different generations (cohort effects).
3. It does not matter when they were born and what age they are, but you tested them a year apart, and this time difference in testing could have affected the result (period effects).

In reality, all three effects probably take place at the same time.

Disentangling these three sources of variance is important because the forecast and the practical strategy depends on it. For example, do obesity rates in people who belong to different birth cohorts develop similarly over time? If they do, we can use experience gained with the previous generation to design prevention strategies for the new generation. If they do not, this may be a useless exercise.

Discussion

Check if Shramm et al.'s (2023) prediction was right regarding the rate of obesity in Denmark in 2025.

What do you think about predictive validity of psychological research?

ATL Thinking, Communication

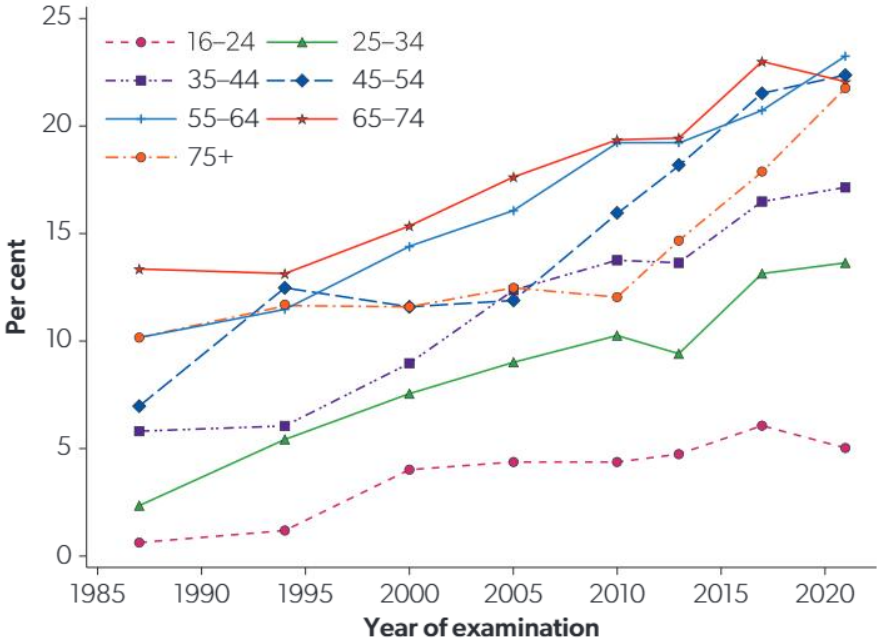
Prevalence rates of obesity

Schramm et al. (2023) investigated the prevalence of obesity among adults in Denmark in 1987–2021. These researchers used data from representative surveys of the Danish adult population conducted in 1987, 1994, 2000, 2005, 2010, 2013, 2017, and 2021. The total number of participants across these studies was almost 92,000.

In Schramm et al.'s (2023) study, the measure of interest was the BMI, calculated as weight (kg) divided by height squared (m^2). The cut-off point for obesity was 30 kg/m^2.

Overall, results of the study revealed that the prevalence of obesity between 1987 and 2021 increased from 6.1% to 18.4%. Continuing this trend and extrapolating into the future, the authors predicted that the prevalence rate of obesity in Denmark would reach 21% in 2025.

Overall, Schramm et al.'s study found significant period effects: rates of obesity increased over time irrespective of age group and birth cohort. In Figure 4.15, you can see that rates of obesity were increasing year after year in all age groups. No evidence for cohort effects was found in this study. In terms of age effect, they confirmed a previously known association: rates of obesity increase from childhood to mid-adulthood, then start decreasing again. Additionally, no differences were seen in the study between men and women.



▲ Figure 4.15 Change of obesity over time in Denmark in different age groups (Schramm et al., 2023)

Conceptual analysis

Perspective

Why is the prevalence rate of social media addiction higher in collectivistic societies than individualistic societies? Why do rates of obesity in Denmark (and globally) increase over time? Why do age differences exist? These and other questions may require a certain perspective, or a certain way of looking at things.

The most prominent perspective that is used for explaining such differences is sociocultural. For example, the changes happening over time are too fast to attribute to biological factors. Differences across nations are often correlated with differences along Hofstede's cultural dimensions. This suggests that it is cultural differences and rapidly changing socio-economic conditions that may be the primary factor affecting prevalence rates.

Causality

Think about the subtle differences between the following three questions:

1. What causes a health problem such as social media addiction?
2. What causes prevalence rates of a health problem to be different across populations?
3. What causes prevalence rates of a health problem in a particular population to change over time?

The first question is about the aetiology of health problems. The focus is on individual persons. How and why does a person "acquire" obesity or social media addiction? There may be a variety of considerations here including biological, cognitive, and sociocultural factors, as well as several theories of such aetiology.

The other two questions are about populations of people rather than individuals. Differences between populations are most probably explained by sociocultural variables such as cultural values, socioeconomic characteristics, globalization, accessibility of food and technology, and so on.

Measurement

Measurement is highly relevant to prevalence rates of health problems because the prevalence of a problem depends on how we measure it.

We have seen in this section that social media addiction, for example, shows a large variation of prevalence rate estimates from one study to another. One of the factors influencing this variation is different approaches to determining the cut-off point between "addiction" and "no addiction" on a standardized scale such as BFAS.

As we have also seen here, the validity of the measurement tool itself may be a huge factor impacting prevalence estimates. Cheng et al.'s (2021) study showed that studies that used an insufficiently validated version of the questionnaire systematically overestimated the prevalence of social media addiction.

Bias

We need to remember that estimates of prevalence rates of a health problem are just that—estimates. It is not a precise measurement. The accuracy of this estimate depends on many factors, such as:

- representativeness of the sample (probabilistic sampling is more representative than non-probabilistic, but the latter is easier and more frequently used)
- validity and reliability of the measurement tools
- diagnostic criteria being used (such as the selected cut-off score).

The problem here is that, when we obtain two different prevalence rates for two different populations, we want to be sure that it is a reflection of actual existing differences between people rather than a consequence of random or biased estimates.

Change

Prevalence rates change from population to population, but they also change over time.

When we look at change over periods of time, it is important to separate three sources of variance: age effects, period effects, cohort effects. We have seen from Schramm et al.'s (2023) study that separating these effects is not always an easy task.

It appears that the prevalence of many health problems has been increasing steadily over the last few decades. This could be explained by sociocultural factors: the spread of social media, availability of fast food and foods containing added sugar, sedentary lifestyle and entertainment, and so on.

Responsibility

There are no major ethical considerations in conducting a study of prevalence rates. Some sensitive questions may be asked in the survey so we need to observe the rules of confidentiality and data protection.

On the other hand, interpreting the results of such studies and applying their findings to construct public health policies has a lot of ethical elements.

As we have seen, bias can easily creep into the estimates of prevalence rates, and this means that statistics can be easily tweaked. Way too often, we see statements like "25% of young people suffer from social media addiction" with no indication of where such estimates came from. Even citing a research study does not make a difference because studies vary so widely in their estimates: a biased policy maker could always choose a study to better align with the message they are trying to deliver.

4.8 Stress and health

Inquiry questions

- Can stress make us addicted to social media?
- Is stress always a bad thing?
- What makes it difficult to research the relationship between stress and health problems?

What you will learn in this section

Key learning:

- Stress is a response of the organism to something demanding. However, stress itself is also demanding.
- In times of stress, the body reacts with activation of the sympathetic nervous system and release of adrenaline and cortisol. Other body systems are temporarily suppressed. If stressors are constantly present, it can lead to health problems.
- Not all stress has negative consequences for health. Some amount of stress is necessary for the organism. It is negative stress (distress) which is experienced chronically that presents the biggest danger.
- According to the transactional model of stress (Lazarus and Folkman, 1984), cognitive appraisal plays an important role in mediating the bodily response to a stressor: in order for a situation to be stressful, it must be perceived as stressful.
- A study has demonstrated that COVID-19-related stress could lead to increased social media addiction, especially when social media was used actively and resulted in feelings of flow. However, the study was correlational.
- Some research suggests that time distortion can be used as a more objective indicator of social media addiction that can supplement traditional self-report measures.

Key terms: stress, homeostasis, stress hormones, fight-or-flight response, eustress, distress, acute stress, chronic stress, transactional model of stress and coping, primary appraisal, secondary appraisal, coping strategy, social media addiction, active use of social media, flow, time distortion

In a wider context

Stress is widely recognized as a major cause of health problems. We have already discussed stress implicitly in other contexts. For example, stressful life events and other environmental factors that contribute to the aetiology of mental disorders (such as depression). Stress even features in the name of the diathesis–stress model.

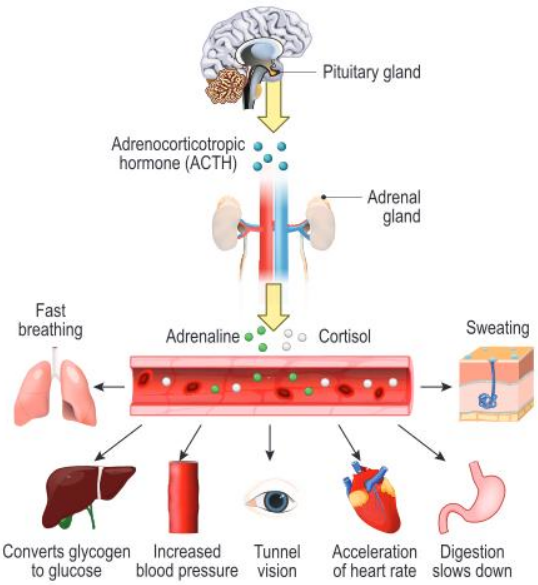
In this section, we delve deeper into the concept of stress and how it is related to health and well-being. The IB curriculum places this topic under the more

generic heading of "Health and well-being", but this is for example purposes only. Stress could be equally considered as a precursor of mental disorders. Keeping the focus on examples we covered earlier, we will consider the role of stress in the development of social media addiction.

What is stress?

Stress is a physiological or psychological response of an organism to an adverse environmental condition (also referred to as a "stressor"). Stress is a tricky phenomenon because, being the organism's response to something demanding, it itself can be something demanding. Stress itself is an adverse condition, and in this sense stress is stressful.

Hans Selye first used the term "stress" in the biological context, defining it as "the non-specific response of the body to any demand placed upon it" (Selye, 1976). In biology, an organism's optimal condition for living is known as **homeostasis**. Anything that causes the organism's condition to move too far away from the state of homeostasis can be experienced as stress. For example, life-threatening conditions will be experienced as extremely stressful and the organism will try to restore the balance, consuming resources if necessary. The fact that resources are being consumed is in itself stressful also, so it is important that homeostasis is restored quickly. If it does not happen, stress may become a chronic condition. Chronic stress has profoundly negative and long-lasting effects on the organism.



▲ Figure 4.16 Stress response system

Biology of stress

When the organism responds to a stressor, the sympathetic nervous system becomes very active. Activity of the sympathetic nervous system is associated with the so-called "**fight or flight**" **response**. It involves dilation of pupils, increased heart rate, narrowing of blood vessels, increasing air flow in the lungs, increased metabolism, sweating, and secretion of the **stress hormones** adrenaline and cortisol.

Cortisol increases glucose (sugar) in the bloodstream and enhances the brain's use of glucose, which gives a short-term boost to energy. Cortisol can also suppress non-essential body functions such as digestion, reproduction, and growth. Instead, energy is diverted to skeletal muscles and the brain—things that can be useful when rapid action is needed.

Think about a time when you were scared or thought you were in danger. You might remember your heart racing and your attention becoming focused on the source of danger. This was a useful response in our animal ancestors. All of their resources were mobilized to enable them to run as quickly as they could or fight as fiercely as possible.

After the perceived stress has passed, the levels of stress hormones gradually return back to normal. Other systems that were temporarily suppressed gradually resume their activities. However, if stressors are constantly present, then the "fight-or-flight" reaction in the body remains activated, which may have detrimental consequences.

In the modern world, we are no longer under the threat of being attacked by a predator, but there are other things that cause stress: homework, exams, fast pace of life, relationship problems in the family. The human body treats these environmental factors in the same way as it treated direct physical threats in our primitive ancestors.

Chat with AI

What is presented here is a basic understanding of the biology of stress response. To dig a little deeper, consider having a conversation with your favourite generative AI. Make sure to be curious and ask questions! Here are some sample prompts you could use:

- Can you provide a step-by-step explanation of the full biological process involved in a stressful (fight-or-flight) response? Make it suitable for a high school student.
- How exactly does cortisol suppress other body functions such as digestion?
- Why are people afraid of a lion in a zoo cage, but not afraid of a fast-moving car on the road?
- Is stress experienced by the modern person the same as stress experienced by human ancestors, in biological terms? What are the similarities and differences?

TOK

It looks like human culture has evolved very rapidly and human biology has not been able to catch up. Our minds are very well suited for dealing with the complexities of cultural artefacts (such as art or calculus), but these minds sit in bodies that are still optimized for life in nature. We no longer have to climb a tree when we are scared, but our hands still get sweaty in order to increase the grip between our palms and the tree bark.

Do you think the fact that our minds are hosted in a biological body has had any impact on knowledge? What kind of impact? Would your conclusions be different in, for example, natural sciences and the arts?

Types of stress

In 1975, Selye divided all stress into **eustress** and **distress**.

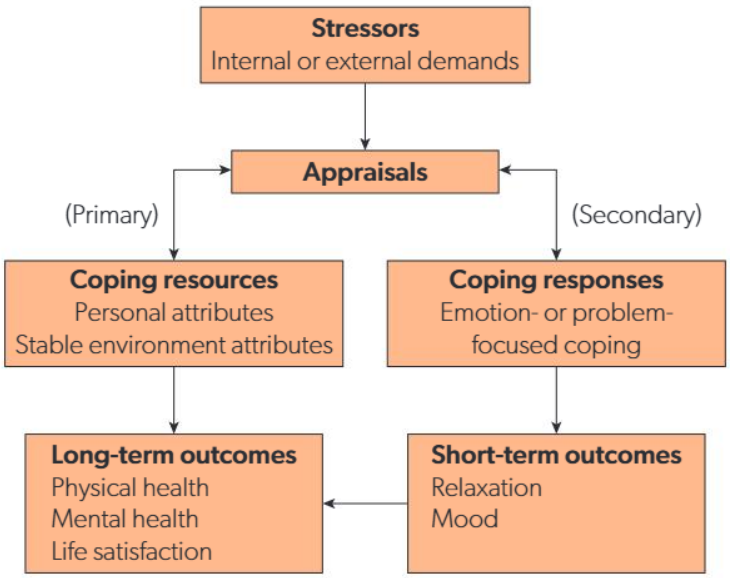
Eustress literally stands for "noble stress". This is when stress performs a positive function such as challenging a sportsperson to push the boundaries during training. Every new experience represents some form of challenge, so every experience involves some degree of alarm or arousal. Cognitive appraisal plays a large role. Depending on how an individual perceives it, an event may be interpreted as an interesting challenge or a threat (distress).

Distress is the negative type of stress: it is not completely resolved by coping or adaptation. It may be persistent and it may result in negative consequences for the organism.

We should also distinguish between **acute stress** and **chronic stress**. Acute stress includes any stress that is suffered for a short period of time, for example, a traffic jam, an argument with a friend, criticism from a teacher. Our body is designed to be good at handling episodes of acute stress: all body systems rapidly change their functioning for a short while. When the danger is gone, all body systems quickly return to normal. Chronic stress is prolonged, long-lasting exposure to stressors. An example would include living in an abusive relationship, having a demanding job, raising children alone, living in poverty. Acute stressors may turn into chronic stressors: think about a bus driver who has to deal with terrible traffic every day, or a student whose teacher is unreasonably critical all the time. Our body is not designed very well to cope with chronic stress. Over time, chronic stress may gradually increase your resting (normal) levels of muscle tension, heart rate, and blood pressure, thus creating a permanent "fight-or-flight" response that mobilizes some systems of the organism at the expense of others.

Appraisal

The **transactional model of stress and coping** proposed by Lazarus and Folkman in 1984 explained the cognitive processes that mediate the influence of stressors on the bodily response. They argued that in order for a situation to be stressful it must be perceived as stressful.



▲ Figure 4.17 Transactional model of stress and coping (Lazarus and Folkman, 1984)

According to the theory, appraisal occurs in two stages: primary and secondary appraisal.

1. At the stage of **primary appraisal,** the situation is assessed in terms of the impact the situation will have on the individual's well-being (not important, positive, or negative).

2. At the stage of **secondary appraisal,** individuals evaluate the extent to which they can cope with the situation. Generally speaking, two outcomes are possible: positive (I can do it; even if I fail at first, I will try again; I will try something else if this does not work) and negative (I cannot do it, I will not even try because my chances are too low).

Appraisal results in adopting a specific **coping strategy**. According to Lazarus and Folkman, there are two major coping strategies:

1. Problem-based coping: this is when you are focused on finding the practical solution to the problem.
2. Emotion-based coping: this is when you are focused on managing your emotions in time of stress.

Stress as a cause of social media addiction: Zhao and Zhou (2021)

Zhao and Zhou (2021) had a correlational hypothesis: that people who experienced greater stress related to COVID-19 were at a greater risk of addiction to social media.

Participants in the study were 705 college students from China. It was a correlational study using questionnaires. The measures used in the study included:

- Experience of stressful events related to COVID-19. For example: got the infection, experienced loved ones dying from the infection, had no access to medical care, stayed home for a long time, and so on. These were "Yes" or "No" closed questions, with a total possible score of 10.
- Addictive social media use, as measured with the brief version of BFAS.
- **Active use of social media** was measured with another short questionnaire—participants had to report how frequently they had been engaging in various activities on social media platforms, such as updating status, giving "likes", or writing comments and so on.
- Flow experience related to the use of social media was measured using an 11-item questionnaire. A sample item is "While using social media, I'm deeply engrossed". **Flow** is a psychological phenomenon when a person is so focused on something that they become deeply absorbed in that activity and stop noticing time.
- Participants were also asked to estimate the amount of time per day they spent on social media platforms, during the period of a severe pandemic.

Results of the study confirmed that the level of stress experienced by the participants during COVID-19 positively correlated with all other measures: active use, social media flow, and addiction to social media. More sophisticated correlational analyses also showed that we can think of active use and flow experience as variables that mediate the relationship between COVID-19-related stress and **social media addiction**. In other words, experiencing high levels of stress related to the pandemic was positively correlated with social media addiction. This was even more so if the individual engaged in active use of social media, and even greater if they experienced flow while using it.

Activity

Team up with a partner and have a "stress interview" with each other. Make this a semi-structured interview: all the questions need to be covered, but the sequence can be changed. You are free to allow the conversation to deviate somewhat to get additional detail. Some example questions:

- Give me an example of something that you commonly experience as eustress and something that you commonly experience as distress.
- In these two examples, what is the role of cognitive appraisal in the way you experience stress?
- What are your preferred strategies to cope with exam stress? What exactly do you do?

ATL Communication, Self-management, Social

Discussion

What alternative explanations could you suggest to fit the data observed in the study?

ATL Thinking

Exact correlations are given in Table 4.3.

	COVID-19 stress	Active use	Social media flow	Social media addiction
COVID-19 stress	1	0.15*	0.13*	0.16**
Active use		1	0.28**	0.25**
Social media flow			1	0.46**
Social media addiction				1

$p < 0.01$
**$p < 0.001$*

▲ Table 4.3 Correlations found in the study of Zhao and Zhou (2021)

The way researchers interpreted their findings was by suggesting that users who experience COVID-19-related stress first adopt an active use strategy. In an attempt to cope with their emotions, they start sharing more experiences online and reacting more to other people's experiences. This may lead them to experience social media flow, that is, the continued pleasant state of being reinforced and supported. In turn, this may lead to the development of addiction. However, while this explanation seems feasible, other trajectories could also fit the data obtained in this study: "the correlational nature of the study allows only hypothetical conclusions about the causality of the described associations" (Zhao and Zhou, 2021).

Social media addiction and time distortion

Zhao and Zhou (2021) outline the potentially important role of feelings of flow in developing social media addiction. When people are under stress, they may be seeking this feeling because it relieves them from stress. However, the feeling of flow itself may become addictive to people.

The problem is that the phenomenon of "flow" is not well understood because it is a highly subjective mental state and self-report may not be a reliable measure. For example, research participants are often asked to estimate the amount of time they spend on social media. However, what if their time perception is distorted when they are on social media sites, and what if they genuinely underestimate the duration?

Discussion

What would you name as the most essential limitation of self-report methods in psychology?

What do you consider the most important limitation of clinical interviews?

ATL Research

Turel, Brevers, and Bechara (2018) investigated the phenomenon of **time distortion** in people who may be considered at-risk for social media addiction. They argued that it is essential to come up with other, more objective measures to supplement self-reports.

A sample of 274 university students was recruited. The study was conducted in a computer lab from which all clocks had been removed. Students were asked to complete an online survey about Facebook use. At the same time, participants had to use a computer but could not check their Facebook. This was expected to induce a craving to check their Facebook page in those who were addicted to it. The task was designed to take approximately 20–25 minutes. On the last screen of the survey, participants were asked to provide their best guess of the time it took them to complete the survey.

Facebook addiction was measured using the BFAS. The score on this scale was used to split participants into two groups: low-risk and high-risk for Facebook addiction. Time distortion was calculated as the estimated completion time divided by the actual completion time.

Results demonstrated that there was a difference in time perception: the low-risk group slightly underestimated the time of survey completion, whereas the high-risk group quite significantly overestimated it. On average, if a participant spent 20 minutes completing the survey, they would report 18 minutes and 28 minutes in the low-risk group and the high-risk group, respectively.

Overall, it was concluded that time distortion on a task that makes participants think about Facebook, yet prevents them from checking their Facebook account, may be a reliable objective indicator of social media addiction. It may even be used to categorize participants into groups of low risk and high risk. They do not suggest replacing other measures with this indicator, but rather supplementing them.

Conceptual analysis

Perspective

Stress leads to health problems. This is not debated. Where different perspectives exist is our understanding of the exact role of stress among all other factors and variables. For example, one suggestion would be that stress translates into health problems "automatically", whereas another perspective (Lazarus and Folkman) is to give special consideration to cognitive appraisal as a variable that mediates the influence of stressors.

Stress itself may be viewed as a complex response that has all the components: biological (fight-or-flight response, release of cortisol and adrenaline, etc.), cognitive (primary and secondary appraisal), sociocultural (social support networks, risk factors, and vulnerability factors).

Causality

Does stress cause health problems or do health problems cause stress? As usual, the answer is probably both, and that there exists a dynamically developing interaction between these things. For example, encountering large amounts of daily stress at school and in the family may cause adolescents to turn to social media to seek peer support. However, turning to social media may further remove them from real-life relationships, thus further contributing to sources of stress.

It is sometimes a bit of a chicken-and-egg problem, and in a practical sense it does not really matter which came first. What matters with any form of addiction is that, when one enters the vicious cycle, it may be hard to get out. Therefore, we need to investigate various factors that influence this dynamic to figure out how to cure or better prevent health problems.

Measurement and bias

Measurement is a large problem in investigating the influence of stress on social media addiction. First, measuring social media addiction relies on measures such as BFAS. Most researchers use some kind of a cut-off point to determine what score on the survey would be taken as an indicator of "addiction".

Another problem with measurement is the self-report nature of data collection tools used in most studies. It is hard to imagine how else data can be collected. One can probably define signs of addictive behaviour in animals, but that would only work with addiction to substances. Social media addiction is a genuinely human phenomenon. Additionally, researchers are limited to using pre-existing groups. However, with pre-existing groups we never know: is it that people experiencing stress turn to social media as a coping strategy, or is it that people who have a predisposition to an excessive use of social networks also tend to get stressed more easily?

Approaches that could partially solve the problem are longitudinal studies and attempts to find more objective indicators of social media addiction, like the one that we saw in the Turel, Brevers, and Bechara (2018) study.

Responsibility and change

Researchers and ethics committees have to make difficult decisions about each study, carefully weighing potential harm against potential benefits. Whenever we investigate the relationship between stress, health problems (such as social media addiction), and related variables, we have a chance of discovering something that explains how health problems arise and how they are sustained. By extension, we have a chance to discover a way to help those suffering from health problems: either by preventing the problem or by giving people effective tools for coping.

At the same time, every such study has a potential to harm both the study participants and society in general. Questions on the surveys may make people uncomfortable or touch upon sensitive topics. Participation in a study may serve as a trigger for some individuals to deepen their addiction or to relapse. Think about a drug addict who has managed to stay "clean" for a while but then they fill out a survey about drug addiction and it serves as a cue for them, causing them to refocus on the problem and start using again. It may not be as obvious with social media addiction, but the mechanism is the same. Remember how participation in a survey in Turel, Brevers, and Bechara's (2018) study apparently distorted participants' time perception and made them feel an urge to check their social media (only if they were already addicted).

4.9 Social learning and health problems

Inquiry questions

- Do we learn unhealthy behaviours by observing others?
- Can we make someone's behaviour healthier by being a role model for them?
- If an intervention programme based on social learning theory is effective, does it mean that the theory is correct?

What you will learn in this section

Key learning:

- Health problems arise from health behaviour, and behaviour can be changed through intervention. To make it happen, we need a good theory that explains how health behaviour occurs, and we need an intervention strategy based on that theory. Social learning theory has been widely used for both purposes.
- The components of the theory that are especially prominent in the context of health behaviours are observational learning (especially in explaining health problems) and self-efficacy (especially in designing intervention programmes).
- Health interventions based on SLT break a distant goal into a series of more attainable goals and make sure that the individual acquires self-efficacy in achieving these goals. This is done through verbal persuasion, vicarious experience, and one's own performance accomplishments.
- It is difficult to obtain experimental evidence of the role of social learning in the aetiology of health problems for ethical reasons. However, one way is through the study of effectiveness of treatment or intervention programmes that were designed based on the principles of SLT.
- One of the problems of this area of study is that for a prevention programme to be "based on the principles of SLT" may mean many different things. It is often unclear which parameters of SLT were responsible for the effects observed in the study.
- It is also important to consider both statistical significance and effect size. It is possible for results of a study to have theoretical but not practical value.

Key terms: health behaviour, observational learning, self-efficacy, performance accomplishments, vicarious experience, verbal persuasion, intervention programme, prevention programme, social cognitive theory, social media-based interventions, statistical significance, effect size

Exam tip

Social learning theory is included in the content required for Papers 1A and 1B. Please see the "SAQ" feature on page 144 for an outline of social learning theory that would be sufficient for your preparation.

Social learning theory is also a topic in its own right: Unit 2.6. This means that this content can also be used in the formulation of exam questions in Paper 1 Section C.

Material presented in this section is focused on the application of social learning theory to health behaviour. We assume that you have studied Unit 2.6 prior to studying this section.

In a wider context

A term that is widely used in health psychology is **health behaviour**. Although it sounds simple, it actually has a very profound meaning. Some people are used to thinking about their health as something that happens to their bodies. They see it as something unrelated to how they think, what kind of values they have, what kind of personality, and so on. Indeed, you catch the flu because a virus enters your system, not because of the kind of person you are.

However, it would be an oversimplification to think so. As the term suggests, your health is related to your behaviour. Even with the flu, whether or not you catch it depends on how strong your immune system is. Strength of the immune system can be influenced by such factors as stress, sleeping patterns, engaging in sports, and exercising. It also depends on personal hygiene. All these behaviours stem from your values, attitudes, and personality, and they all influence the probability of the germ entering your system in the first place.

The influence of health behaviours is even greater in some other health problems, such as obesity or addiction. Such diseases and health problems can be said to be "related to lifestyle". However, lifestyles can be altered by positive changes in behaviour. Various promotion strategies and **prevention programmes** have been designed for this purpose. Any such intervention should be based on a theory of behaviour change. This is a theory that, in an evidence-based way, explains how people change their behaviour and what factors contribute to such a change. Albert Bandura's social learning theory has been applied to health behaviour and is viewed as one of the most prominent theories of behaviour change in health-related areas.

Social learning theory in the context of health behaviour

Social learning theory can explain the origin of health problems. The idea that is frequently used in this regard is **observational learning**.

Using the example of social media addiction, the theory suggests that individuals learn social media behaviours through observing and interacting with other people. We are often invited to a social network, either by a friend or others in the community, or we hear others talk about it and we want to follow their example. Perhaps the secret of their popularity lies in the ability of social networks to constantly grow by attracting new members.

Social learning theory has also been used to figure out how to change addictive behaviours.

Social learning theory places a strong emphasis on the idea of **self-efficacy**: one's belief that they are competent to perform the behaviour needed to influence the expected outcome (the "I can do it" belief). According to the theory, self-efficacy beliefs are influenced by the following information:

1. **Performance accomplishments**—the person's history of previous successful attempts to perform the behaviour and achieve the desired outcome.
2. **Vicarious experience**—experience obtained through the observation of others and successful and unsuccessful outcomes of their efforts. In

order for self-efficacy to be learned vicariously, the observed model should successfully achieve desired outcomes through investing effort and overcoming difficulties. If the result is achieved with ease or by a lucky chance, this will not contribute to the observer's self-efficacy.

3. **Verbal persuasion** is a technique used in health education to convince people that they have a great degree of control over their own health behaviours.

Bandura believed that self-efficacy plays a central role in the regulation of behaviour: "Among the mechanisms of personal agency, none is more central or pervasive than people's beliefs about their capabilities to exercise control over events that affect their lives" (Bandura, 1989, p. 1175).



▲ Figure 4.18 Social learning

Health interventions based on social learning theory break the distant goal (lifestyle change) into a series of smaller, more attainable goals. This is important because self-efficacy is believed to be the most important factor of change, and personal accomplishments are believed to be the most important resource for self-efficacy. People need to feel that they can be successful, even if it is something small. Attainable subgoals provide immediate positive rewards and encourage self-motivation by nudging people towards "believing in themselves".

People very often misrepresent poor health practices as "diseases": we suffer from health problems not because we make poor choices, but because a disease "happens to us". With such a mindset, people may seek the help of medical practitioners and fail to accept (at least partial) responsibility for their own health problems. The idea of health behaviours shifts this responsibility back to the client: have you done anything that allowed this "disease" to "happen" to you? (Blair, 1993).

Chat with AI

To deepen your understanding of the role of social learning in health behaviour, consider having a conversation with your favourite generative AI. You could use the material presented here as a starting point for designing some of your prompts. Here are some examples:

- How does social learning theory explain why and how people become addicted to social media? Would that be similar to how it explains other health problems such as obesity?
- What does it mean when an obesity prevention programme says it is "designed on the principles of social learning theory"? What specific features, activities, or methods of intervention does this imply?
- Could you give me several examples of well-known health interventions (for obesity or social media addiction) that have been designed based on social learning theory?

Exam tip

There are many ways one might investigate the role of social learning in health problems. One approach is to look at the effectiveness of **intervention programmes** based on social learning theory, as we are doing in this section. However, this is merely an example. Other studies have taken different approaches. For example, a large number of correlational studies have studied the role of self-efficacy (high versus low) in the probability of relapse after a treatment of drug addiction.

Activity

Design an eight-week intervention programme to reduce the rates of social media addiction in your school.

What theoretical principles will your programme be based on? Which behaviours will be your target variables? How exactly will you use the ideas of observational learning and self-efficacy?

What specific week-by-week activities will it include?

It may be difficult to come up with a detailed programme from scratch. Consider first writing a paragraph that describes the problem, in the specific context of your school. Then create an outline of your programme. Once you are finished with the draft, run your draft together with the problem description through a generative AI. As part of your prompt, make sure to stress that you are building a programme on the basis of social learning theory. Ask AI to suggest improvements or things to consider, and take it from there.

ATL Research, Communication, Self-management

Experimental evidence of the role of social learning in health problems

If we focus on the study of aetiology of health problems, then we are mostly limited to non-experimental evidence. This is because it would be unethical to divide people randomly into groups and manipulate a variable that is likely to create a health problem.

One way to use the experimental design with people and therefore be able to make cause–effect inferences is through the study of effectiveness of treatment and prevention programmes. The logic is as follows:

- We design a prevention programme based on the principles of a specific theory.
- The prevention programme works.
- Therefore, the theory is correct about the causes of the health problem and the variables targeted in this theory are indeed the ones that have caused the problem.

This logic is not perfect. Here are just a few reasons why:

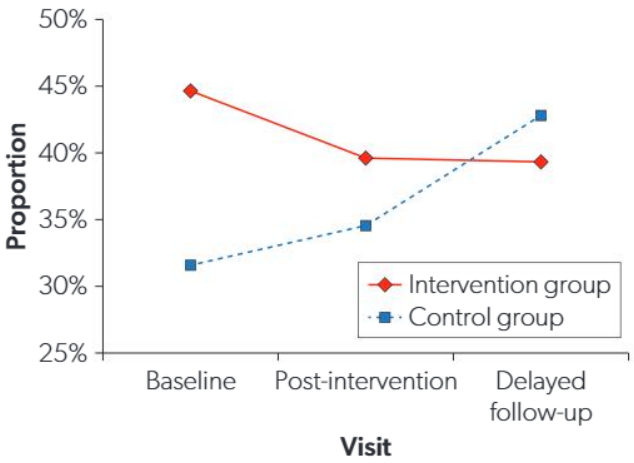
- A theory that explains what causes a health problem is usually a sophisticated theoretical construct that has multiple parameters. So if a treatment that was designed "based on the principles" of this theory "works", it is not clear what exactly made a difference.
- It is not always clear how to measure whether a treatment has "worked". For example, in obesity would it be a reduction of BMI that we consider an indicator of effectiveness? By how much should the BMI decrease for us to believe that the effect has been significant? How long term should the change be?
- Related to the two points above, reviews have shown that there may be a large variation in the findings of research studies that investigate the effectiveness of treatment and prevention programmes for health-related behaviours. When findings are varied, it most likely means that effectiveness of such programmes may depend on many factors and a large number of studies would be needed to make a reliable conclusion.

Social learning and obesity prevention programmes

Black et al. (2010) conducted a study to investigate the effectiveness of Challenge! This is a health promotion and obesity prevention mentorship programme among urban Black adolescents. A total of 235 Black participants (aged 11–16) were recruited from low-income urban communities. Participants were randomly split into those who took part in the intervention programme (the intervention group) and those who did not (the control group). The intervention programme was "anchored in social cognitive theory" and was delivered by college-aged Black mentors. The intervention included 12 sessions. Measurements were taken at the start of the study, 11 months after the start (at the end of the intervention period) and 24 months after the start (delayed follow-up).

For the sessions, mentors came to the adolescent's home and also went on field trips to community sites such as parks. The idea behind the intervention in the Challenge! programme was that the mentor would be a relatable role model and through interacting with the mentor, adolescents could learn a healthier lifestyle and grow more confident and adopt healthy behaviours. The intervention was designed to help adolescents identify with the model (which in social learning theory is one of the factors of vicarious learning). The mentors were of similar age and social status and the activities were selected to be relatable. For example, one activity included a rap music video promoting healthy eating and exercise. Role modelling and support were central to the intervention. Each session included a challenge, such as convincing someone to drink water instead of soda. Mentors helped the participants plan a realistic challenge, come up with an implementation plan, reflect on successes and challenges, and so on.

Results showed a moderate effect of the intervention. The percentage of obesity declined 5% in the intervention group and increased 11% in the control group.



▲ Figure 4.19 Proportion of overweight and obese adolescents in the intervention group and the control group across time (Black et al., 2010)

The intervention increased physical activity after 11 months but not after 24 months (it was a short-lived effect). However, snack and dessert consumption in adolescents in the intervention group declined significantly at both 11 and 24 months. Participants with the highest BMI benefitted from the programme the most and showed the greatest improvement.

Social-media-based interventions

Social media can be a source of role models. Observing these models vicariously and watching them being reinforced for their health behaviours contributes to observational learning, especially if the participant can identify with the model. **Social-media-based interventions** are programmes where participants are asked to follow carefully curated social media content in order to learn healthy behaviours.

In a meta-analysis of 28 RCTs investigating the effectiveness of social media-based obesity prevention, Loh, Yaw, and Lau (2023) revealed a statistically significant weight reduction in social media-based interventions as compared to control

Exam tip

In some parts of this chapter, we are referring to **social cognitive theory** (SCT) rather than social learning theory (SLT). Remember: SCT was a further development of SLT, so broadly speaking it is the same theory.

TOK

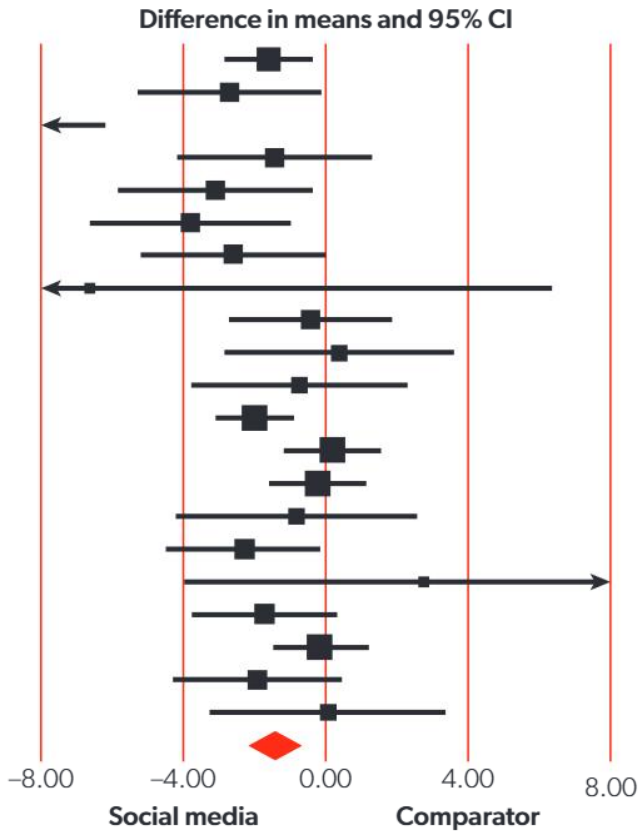
Does a significant scientific result have to be practically useful?

There are many examples in psychology where the result obtained in a study was statistically significant, but not large. Such a result has theoretical, but perhaps not practical value.

Think about how this relates to fundamental and applied science. What other areas of knowledge can be seen as having a similar distinction? Is it worth pursuing knowledge that we know will not be practically applicable?

groups (such as standard care, waitlist, or placebo): mean difference –1.45 kg (95% confidence interval (CI) between –2.15 and –0.75).

There was also substantial heterogeneity between studies. For example, the largest mean difference reported in a study (Chee et al., 2014) was 11 kg less in the intervention group compared to the control. In contrast, Stephens et al. (2017) reported the intervention group to be 2.8 kg heavier on average than the control group (as cited in Loh, Yaw, and Lau, 2023). Figure 4.20 summarizes the findings from this review.



▲ Figure 4.20 The difference between social-media-based obesity prevention and control groups across studies used in the meta-analysis of Loh, Yaw, and Lau (2023). Zero is the point of no difference. Lines around the central point denote a 95% confidence interval. The red shape denotes the overall mean.

When researchers further investigated potential reasons for such heterogeneity, it was revealed that the length of the intervention and publication status had a significant effect. Longer interventions resulted in more weight loss.

An additional difficulty was that most of the RCTs included in the analysis were conducted in Western countries and reported in English-medium journals, so we cannot be certain as to the generalizability of results to other societies and cultural contexts.

Another problem that the study highlights is the importance of taking both **statistical significance** and the actual **effect size** into consideration. For example, the mean weight difference between the intervention group and the control group across all studies was found to be statistically significant. At the same time, however, the actual value of this mean difference was 1.45 kg. Think about the practical significance of this value: is it worth the effort?

Conceptual analysis

Perspective

Health problems in this unit have been considered from the sociocultural perspective. People are social beings and being in society (observing others, learning from them, feeling a sense of belonging to them) has a great influence on their behaviour. That includes health behaviour. Social learning theory is a comprehensive theory that attempts to explain the exact mechanisms of learning new behaviours in a social context.

Bias

As always, research studies may be biased. In this particular area some of the most prominent potential biases include publication bias and sampling bias. Publication bias is prominent because there exists an incentive to publish "successful" studies and refrain from publishing "unsuccessful" ones. Sampling bias is prominent because various health problems may manifest very differently in different cultures (think about obesity in the USA versus India). However, most studies are conducted in Western societies.

Apart from separate research studies, bias may also be a feature of our theoretical conclusions derived from research. We have seen from results of meta-analyses that there may be a large heterogeneity of findings. Imagine a practitioner that designs a prevention programme for obesity. To ensure that this programme is evidence-based, they would have to look at published research studies and derive conclusions from them. But depending on what the selection of studies is, this conclusion may be quite different. So it is important to ensure that the selection of studies itself is as unbiased as possible.

Measurement

The most obvious challenge in this area of research is how to measure a health problem (such as obesity, drug abuse, or social media addiction). For example, in obesity is it:

- BMI
- waist-to-hip ratio
- waist-to-height ratio
- daily amount of added sugar consumed
- daily amount of exercise and physical activity?

When the focus of the study is on measuring the effectiveness of an intervention or a treatment programme, there is the added question of what to consider an indicator of effectiveness. In obesity, would it be a reduction in BMI? How large should it be for us to consider it significant? Would it be a change in behaviour such as diet and exercise? Would we consider an effect that is present 2 months after the intervention but disappears after 12 months "significant"?

Causality

With any health problem, we want to know what causes it. This knowledge could give us the ability to predict it and eventually learn to control it. One of the main obstacles to acquiring such knowledge is the multi-determined nature of health problems (they are caused not by a single factor but a complex interaction of factors). Another one is the absolute predominance of correlational evidence.

Activity

Consider making a large table with all the sections of this chapter in rows and the six key concepts in columns. Very briefly (just with key words) explain how the concept links to the content in each of the cells. You will soon see that there are a lot of repetitions and overlaps.

ATL Self-management

Normally, conducting a true experiment in this area is only possible in the context of treatment: we deliver a treatment to one group and a placebo to another. However, an important limitation in using such evidence is treatment-aetiology fallacy: if an intervention based on a certain theoretical principle "worked", it does not necessarily mean that the theoretical principle is a valid explanation of the health problem.

Change

Health behaviours are not static—they change. An overweight person can influence his or her weight greatly through a change in lifestyle. We can design intervention and treatment programmes to affect health-related behaviours and reduce health problems.

Therefore, apart from knowing what causes health problems, it is also important to understand the dynamics of their change. For example, how long does it take for changes in diet and exercise to be reflected in a person's BMI? How likely is it that people who switch to a healthier eating style will be able to maintain it over the long term? These and other questions determine what we know about the dynamics of development of a health problem—knowledge that is of paramount importance if we want to change anything.

Responsibility

Perhaps unlike some mental disorders, health problems cannot be treated medically. We cannot give people a drug that makes them less addicted to other drugs. There is no medicine against social media addiction.

This means that in targeting these health-related behaviours, we need to pull a complex network of other strings: social influences, cognitive appraisals, feelings of self-esteem. Social learning theory is a powerful explanatory tool that brings many of these factors together and explains how they interact. Because the interaction of variables is so complex, there exists a danger that unjustified simplification of the theory (or research findings) may lead to misconceptions about the health problem. For example, it would be way too simplistic to say that people "learn social media from others". Misrepresentation of the theory and research may lead to simplistic, ill-informed, ineffective, or even harmful prevention strategies. A blanket ban on mobile devices in a school, shaming as a means of fighting obesity, limiting access to information about illegal drugs—these could be examples of strategies that may do more harm than good. Whoever is designing and enforcing the strategy has the responsibility of justifying why this strategy is expected to be helpful in this particular context.

4.10 Prevention of one health problem

Inquiry questions

- How effectively can health problems be prevented?
- To prevent a health problem, is it better to change the person who is affected or the situation around that person?
- How can we know if a health prevention strategy is effective?

What you will learn in this section

Key learning:

- There are a variety of prevention strategies. Some of them target environmental conditions that influence a person's behaviour, others mainly target individuals themselves.
- One of the approaches for the prevention of social media addiction is to enhance an individual's self-control.
- According to the strength-model of self-control, people have a limited self-control resource: they may diminish this resource in one task and find it difficult to exercise self-control in a subsequent task.
- Researchers also divide self-control into proactive (preventing potentially health-risky situations from arising) and reactive (modifying one's cognitive response to the situation).
- Motivation is another variable that has been targeted. Enhancing motivation and making a commitment may lead to long-term improvements in social media addiction. Longitudinal studies are important to examine short-term and long-term effects.
- Mindfulness is another key variable that has been targeted in prevention programmes. It has been shown to contribute to social media addiction directly, but also indirectly through such mediating variables as attention control and FOMO (fear of missing out).
- Experimental, correlational, and qualitative evidence in this area each have their advantages and disadvantages. Triangulation of evidence is important.

Key terms: treatment and prevention, prevention strategy, self-control, strength-model of self-control, proactive and reactive self-control, motivation, motivational enhancement therapy, mindfulness, fear of missing out (FOMO), attention control

In a wider context

The difference between **treatment and prevention** is that treatment occurs in response to an already existing problem, whereas prevention is a set of measures designed to avoid the onset of the problem in the first place.

Exam tip

In the previous unit (Social learning and health problems), we assessed the validity of social learning theory (SLT) as an explanation for health problems by analysing the effectiveness of prevention and intervention strategies designed on the principles of SLT.

This means that the two sections are closely interrelated, and there is a lot of conceptual overlap between these two content points.

The design of prevention strategies depends on our beliefs about the aetiology of a health problem: if we believe that a particular health problem is caused by a particular factor, then our prevention strategy will target that factor.

When we develop a prevention strategy, we are guided by our theoretical understanding of how health problems appear. If the prevention strategy is actually effective, this also serves as evidence in favour of our theory. This is why assessing the effectiveness of intervention strategies is important both for theory and practice.

Overview of prevention strategies

Prevention strategies come in a variety of forms. To some extent the choice of strategy is determined by the health problem that is targeted. For example, prevention programmes for drug addiction will be somewhat different from prevention programmes for social media addiction, although there may be overlap because some generic protective factors are the same (e.g., **self-control**).

Some prevention strategies mainly target the environmental conditions that influence a person's behaviour. For example, we could use "nudges" to encourage healthy eating or automatic notifications to remind people about the time they have spent on social media. Other prevention strategies mainly target the individual himself or herself. For example, they may enhance self-efficacy and goal-setting, making individuals more likely to take conscious control of their health behaviours.

In this section, we will use prevention of social media addiction as an example, but the main principles of constructing prevention strategies and the main problems with research in this area are applicable to other health problems.



▲ Figure 4.21 Prevention

The role of self-control in social media addiction

Self-control is often mentioned as a protective factor against excessive use of social media. For example, individuals who exercise self-control can check their social media accounts one or two times daily and refrain from repetitive use impulsively throughout the day.

The question is: what can be done in order to develop and strengthen self-control?

Baumeister et al. (1998) suggested a **strength-model of self-control** where self-control relies on a limited resource that gets depleted if used continuously, much like a muscle runs out of energy if flexed for a long time. According to the strength-model, engaging in self-control quickly consumes this resource, leaving the individual in a state of "ego depletion" and diminishing their ability to control their behaviour afterwards.

Supporting this idea is an interesting study by Brevers et al. (2018). In this study participants were randomly split into two groups:

- The experimental group was asked to read five sentences describing an effortful action and, after each sentence, close their eyes and imagine themselves engaging in that activity for 30 seconds. Examples of such sentences included: "To study instead of going on the internet", "To study instead of going out with friends", "To wake up instead of pressing the snooze button".
- The control group went through a similar procedure, except that the sentences they were required to read did not contain any effortful action and in fact only referred to daily objects—for example: "A building that contains five levels", "An elevator that goes up to the second floor", "A room painted in white". Participants visualized these scenes for 30 seconds each.

After the visualization exercise, participants engaged in a gambling task where they had to make a series of financial decisions involving risk. Results of the experiment showed that participants who engaged in more effortful visualizations (such as "To study instead of going on the internet") put less effort into thinking about the decision-making problems and made riskier decisions on the subsequent task. This seems to support the theory that self-control relies on a limited resource: participants depleted this resource when thinking about the effortful actions, so they did not have sufficient self-control left to help them with the gambling tasks.

This theoretical model of self-control later saw some development. Researchers then started paying more attention to proactive behaviours that allow a person to avoid challenging situations. As a result, self-control strategies were divided into proactive and reactive.

- **Proactive self-control** strategies include situation selection and situation modification. These strategies prevent people from finding themselves in situations demanding the kind of self-control that they do not have.
- **Reactive self-control** strategies are targeted at modifying one's cognitive response to the situation rather than the situation itself.

Brevers and Turel (2019) conducted a naturalistic investigation of self-control strategies in the use of social media. The sample consisted of 751 undergraduate students in an American university, aged 18–49. Participants completed an online survey. The survey was open-ended. Participants had to name strategies that they use to regulate their use of social media, and for each strategy they were asked to provide a one-sentence clarification.

After responses were collected, three trained coders read them and extracted recurring themes using inductive content analysis. The result of this analysis was identifying eight themes each pertaining to a specific self-control strategy of social media use. See Table 4.4 on the following page for the list of self-control themes (strategies) identified in this study, as well as sample participant responses in each category.

TOK

What is the value of a theory in the acquisition of knowledge? Theories may be wrong. Why can't we get rid of theories entirely and simply base our knowledge on observed data?

A related concept relevant to natural sciences is a theory-laden observation. It has been claimed that every observation is theory-laden—that is, it is impossible to not be influenced by a theory when conducting an observation (scientific or otherwise).

To what extent do you think this applies to psychology? How about history?

Activity

Interview your partner about their social media use. Some of the questions you could consider asking:

- How do you control your social media use?
- Do you feel like using social media is affecting your well-being? In what ways?
- What specific strategies do you employ in order to protect your well-being?
- Do you often compare yourself to others? In what ways do you think it is affecting you?

ATL Communication, Self-management, Social

Theme (strategies)	Sample response
1. No need to control	"Never had social media"
2. No strategy	"I don't want to control my social media use"
3. Full prevention of access	"Charge my phone in a different room", "Leave my phone in the locker room while at work"
4. Partial prevention of access	"I put my phone to charge four metres away from me", "I put my phone in my purse instead of my pants pocket to make it harder for me to grab"
5. Modifying a feature on the device	"Leave my phone on mute at all times", "Turn off notification sounds"
6. Delimiting use time	"Stop using social media after 11 pm", "Check it only when I have a rest period between classes"
7. Self-talk	"I remind myself that work is more important than social media", "Tell myself that there is an important test coming up"
8. Straightforward self-control	"Keep working", "Ignore my phone", "Finish important tasks before checking social media"

▲ Table 4.4 Themes and sample responses in Brevers and Turel (2019)

Activity

Why did Brevers and Turel (2019) decide to use a qualitative research method? Was it the right choice, given the research question they wanted to investigate?

Split into groups, with each group taking one of the following alternative research methods:

- experiment (e.g., field experiment)
- correlational study
- interview (e.g., focus group).

Within each group, create an outline of the study using this alternative method. What would you have to change? How would you organize the procedure? What would be the advantages and disadvantages of your study compared to the original study?

Present to other groups and discuss.

ATL Research, Communication, Social

From Table 4.4, strategies (themes) 7 and 8 may be considered reactive self-control and strategies 3–6 may be considered proactive situational strategies. Overall, the study suggests that the most effective approach for therapists and other health professionals would be a combination of teaching social media users to enact reactive strategies (such as self-talk) while simultaneously trying to use proactive strategies to reduce tempting situations.

The role of motivation in social media addiction

We have considered the role of self-control and how it could be targeted in prevention programmes for social media addiction. Another variable that is often targeted is **motivation**.

Manwong et al. (2018) investigated effects of a group activity-based **motivational enhancement therapy** (GA-MET) on social media addiction among high school students in Thailand. The study involved 244 students aged 12–15, randomly sampled from eight classrooms. They were randomly split into the intervention group and the control group.

The control group participated in regular school activities. The intervention group participated in the GA-MET programme. It included a number of activities performed in each classroom for 45–50 minutes per week with a facilitator (a psychologist). The eight activities were:

1. ice breaking
2. group discussion about the experience of using social media
3. watching short movies about social media addiction followed by a group discussion
4. doing a self-assessment to determine their individual level of social media addiction
5. a discussion about experiences of refraining from social media use

6. a group activity where participants shared their predictions about the future if the current rate of social media use is maintained
7. a role-play exercise where participants acted out scenarios of regulating social media use
8. a goal-setting exercise where participants wrote their individual goals of behaviour change on cards—the cards were then shown in front of their classroom.

Chat with AI

Have an AI-augmented discussion in a small group about how you could improve this eight-week programme. "AI-augmented" means that a generative AI chatbot will be part of this conversation but will not be dominating it. Everybody should get a turn and the chatbot should occasionally get a turn. When it is the AI's turn to contribute, consider (as a group) what would be the most suitable question to ask it in order to move the discussion productively forward.

Make sure to provide AI with a context of your discussion. Once the context is provided, examples of questions at various stages of the conversation could be:

- Here is an alternative activity that we are discussing for week 4. We think that it is better for enhancing motivation because [insert information]. Do you agree?
- Is there anything in existing scientific research that you are aware of that suggests that 12 to 15-year-old children do not benefit from group discussions in class?
- Here is our suggestion for the goal-setting exercise. Could you suggest any modifications to make it more suitable for this age?

Activities 1–4 were designated as a "contemplation phase" that was meant to increase participants' motivation to change. Activities 5–8 were designated as the "commitment and strengthening" phase.

Three measurements were conducted in the study: at the start of the programme (baseline), a follow-up after 8 weeks, and a follow-up after 12 weeks.

Interestingly, the difference between the intervention group and the control group was not large at 8 weeks (when the programme just ended), but they became larger at 12 weeks. We could conclude that the programme triggered some changes in the participants' behaviour that manifested in the outcome variables over time. It could be the case that more effects would be observed if the follow-up observation was continued even further.

At 12 weeks, the intervention group showed a reduced duration of social media use during weekdays, higher self-esteem scores, and lower depression scores. Interestingly though, the scores on a social media addiction scale did not differ significantly. Several explanations of this are possible. For example:

- More time is needed for the effects to settle in: first the amount of time on social media gets reduced, then addictive behaviours catch up.
- The scale is not completely valid or reliable.

Activity

Which of these alternative explanations seems to be more likely to you?

Can you design a follow-up study to test which of the potential alternative explanations is true?

Swap your study suggestion with a partner and give each other feedback.

ATL Research, Thinking, Communication

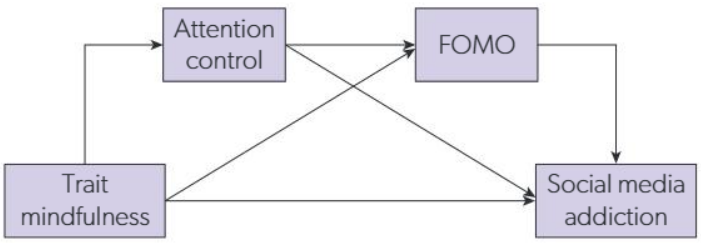
- The GA-MET programme actually only has an effect on the frequency with which social media is used, but not on participants' vulnerability to developing an addiction.

The role of mindfulness in preventing social media addiction

Chang et al. (2023) investigated the effectiveness of **mindfulness** programmes as a prevention strategy against social media addiction. They especially targeted **attention control** and **FOMO (fear of missing out)**.

- Mindfulness is a state of being aware of where we are and what we are doing, as well as our surroundings at the present moment. Mindfulness is an antidote to being overwhelmed, being driven by impulse, and exhibiting obsessive behaviour. The ability to sustain mindfulness is known as "trait mindfulness", and it is believed that this ability can be trained.
- Fear of missing out (FOMO) is the anxiety associated with not knowing what others are doing. This anxiety is typically expressed as an obsessive desire to know what is happening with others. It distracts people from their own experiences in the moment: you cannot be properly focused on what is happening to you if you keep thinking about what is happening to others. Therefore, it is believed that mindfulness training, with its focus on the here and now, may be a prevention strategy that mitigates the effects of FOMO.
- Mindfulness teaches people how to control their attention. For example, mindfulness meditation typically involves focusing your entire attention on a single selected object for a long period of time and resisting any distractions.

Chang et al. (2023) hypothesized a serial mediation model (see Figure 4.22) where trait mindfulness is associated with social media addiction both directly and indirectly through attention control and FOMO. It was a correlational study with a sample of 446 college students from China. Data were collected via questionnaires. Social media addiction was measured with a Chinese translation of the BFAS. In the instructions, the names of social networks such as Facebook, Twitter, and Instagram were replaced by their Chinese counterparts: QQ, Weibo, WeChat, and so on.



▲ Figure 4.22 The hypothesized model of relationships between variables in Chang et al. (2023)

Results of the study showed that all variables were significantly correlated with each other (see Table 4.5). Note how all correlations are in the expected direction—for example, FOMO correlates negatively with mindfulness and attention control, but positively with social media addiction. Further correlational analysis of mediating variables also confirmed that all pathways shown in the

image are significant. FOMO was a more significant moderator of the effect of mindfulness on social media addiction than attention control.

	Mindfulness	Attention control	FOMO	Social media addiction
Mindfulness				
Attention control	0.40 *			
FOMO	–0.44 *	–0.35 *		
Social media addiction	–0.52 *	–0.37 *	0.67 *	

* $p < 0.01$

▲ Table 4.5 Correlations between all variables in Chang et al.'s (2023) study

For college students, it was concluded that mindfulness may be a protective factor against developing social media addiction. Apparently, mindfulness is associated with better control of attention and less FOMO, both of which are factors contributing to unhealthy social media use.

Conceptual analysis

Perspective

In terms of the triad biological–cognitive–sociocultural, most of the intervention strategies considered in this unit target cognitive variables. They try to change the way a person interprets the situation and reacts to common challenges, through enhancing such characteristics as self-control, motivation, or mindfulness. Of course there are other prevention strategies, predominantly targeting environmental (sociocultural) variables, but these are mostly in the hands of policymakers, not psychologists.

In terms of different perspectives on the same issue, there exist multiple theories about what the major risk and protective factors are in developing health problems (e.g., social media addiction). Depending on the theoretical perspective, different prevention strategies have been proposed. Most of them have shown some promising results.

Causality

There are numerous interacting variables in health problems. It is probably unreasonable to expect that one single characteristic or personality trait will turn out to be the ultimate cause of one or more health problems. It is always a combination of things.

From the practical point of view perhaps it does not matter that much which specific variable has triggered a change, as long as we know that an intervention programme in general "works". However, it can also be argued that without such knowledge practical interventions may be biased or misled.

Measurement

Measuring a health problem (such as social media addiction) may be problematic because this process is not entirely standardized and not entirely objective. In other units, we already considered problems surrounding the use of various

questionnaires to measure social media addiction (and other health problems), as well as the controversy related to choosing a cut-off point on a scale. Additionally, measuring the rate of improvement also presents challenges. How do we know that an intervention programme has resulted in an improvement? One of the aspects that is important to consider here is the possible difference between short-term and long-term effects. A significant short-term effect may turn out to be transient and short-lived.

Bias

It is theoretically valuable and significant for us to try and isolate variables and test their importance in the development of a health problem. Such was the study of Brevers et al. (2018) who showed the significance of self-control. However, such studies will inevitably be artificial and not immediately applicable to real-life scenarios. Therefore, the investigation of real-life intervention programmes (e.g., Manwong et al., 2018) is equally important. The trade-off here is that such studies will inevitably include more confounding variables.

Triangulating evidence from both these types of research, as well as qualitative research studies such as Brevers and Turel (2019)—in order to delve more deeply into participants' experiences and interpretations—is the way forward.

Change

As we have seen here, a person may change the situation that is conducive to the development of a health problem, or a person can change his or her cognitive response to such situations. Both strategies have been investigated and it appears that they are most effective in combination.

An important point here is that people are not helpless victims of health problems such as social media addiction. It is in our hands to control such problems and make a positive difference.

Responsibility

One aspect of responsibility in this area is conducting research studies while carefully observing ethical guidelines. Although they are designed to bring a positive difference, prevention programmes may still be viewed as somewhat "invasive", because we attempt to modify participants' characteristics such as self-control and mindfulness. Some of these interventions may be seen as "manipulative" by some participants. Another aspect here is, if we know that an intervention programme is going to prevent participants from developing a health problem, are we deliberately allowing participants in the control group to develop the problem? Is it ethical?

Another aspect is obviously applying results of research in practice. Once we have established the effectiveness of a prevention programme in research, there is now the question of rolling it out. For example, mandating that all students in a given school participate in the programme.

Exam-style practice questions

Paper 1 Section A (4 marks)

Explain genetic inheritance with reference to one example.

Paper 1 Section B (6 marks)

Two identical twins share 100% of their genotype. However, they are not equally likely to develop a mental disorder. For example, if we take a sample of identical twins and conduct clinical interviews to diagnose depression, we may find that there are many pairs where one of the twins has depression but the other twin does not.

Explain this pattern of results using the diathesis–stress model.

Paper 1 Section C (15 marks)

No matter what perspective we take on mental disorders, they can never be fully understood without considering cultural differences. Evaluate this claim with reference to one or more perspectives and one or more mental disorders.

(Concept: Perspective. Content: Cultural differences in mental disorders. Context: Health and well-being)



5 Human relationships

Introduction

It is argued that we cannot understand the behaviour or experiences of a human being in isolation from others. We influence and are influenced by the behaviour of other people. In this chapter we are going to consider two groups of topics: group behaviour and interpersonal relationships.

The first of these (group behaviour) deals with how people are affected by the group.

A sense of belonging to a group is especially important to us when we are growing up and in fact can have a huge influence on our development, as you already know from Chapter 3: Human development (see such topics as enculturation of social norms, sociocultural factors in development, attachment, theory of mind, peer influence). Social learning theory (SLT) is a useful framework to think about how exactly the process of socialization occurs.

Another powerful theory in social psychology—social identity theory (SIT)—explains why we behave in certain ways towards certain groups of people. It attempts to explain prejudice and discrimination as well as the well-known phenomenon of in-group bias (when someone treats members of their own group more favourably). Neither SIT nor SLT can be considered purely "social" theories. They both explain what is happening in a person's mind when they are processing social information such as group membership. In this sense, there is a blurred line between cognitive and sociocultural perspectives that blend in these powerful theories.

Culture is also a huge influence. Many psychologists believe that culture should be given top priority in the list of factors affecting human behaviour. Their argument is that biological evolution has long been replaced by cultural evolution, and that anything that is human is also cultural. We will look more closely at cross-cultural variation in behaviour, trying to capture it through Hofstede's concepts of cultural dimensions.

Finally, we will look at two specific types of social influence (how one person can influence the behaviour of another)—conformity and compliance techniques. The key distinction between them is that in conformity, there is no direct request. You behave in a certain way because other people in the group behave so. In contrast, compliance always includes a direct request from one person to another. However, this request is framed in a way to increase the likelihood that you will comply. Various techniques can be used that act upon various psychological principles—for example, the principle of reciprocity.

The second group of topics is interpersonal relationships. Interpersonal relationships include such phenomena as romantic relationships and friendships, but also interpersonal conflict. The dynamics of relationships between two people can be quite different from that between two groups, or between a person and a group. Arguably, in a relationship the dynamics of causality are more complex than social influence. One person affects the behaviour of another, who in turn affects the first person's behaviour. "Behaviour" here is used very broadly: it could be values, attitudes, beliefs, cognitive appraisals, personality traits, overt actions. Additionally, relationships involve multiple interactions, so they may evolve considerably over time.

We will focus on romantic relationships. Firstly, we will consider how they are started and maintained, and we will look at it from two perspectives: biological and cognitive. Then we will focus on how people use communication to maintain their relationships and what communication strategies they may use to resolve relationship conflict.

The following table provides an overview of all topics in this chapter. Remember: the headings in the IB Guide (such as "Group behaviour") are used for convenience only and will not be used in the formulation of exam questions.

Group behaviour	Interpersonal relationships
• Acculturation • Conformity • Compliance techniques • Cultural dimensions • Social identity theory • Social learning	• Chemical messengers • Cognitive explanations • Communication/language • Strategies for improving relationships

5.1 Social identity theory

Inquiry questions

- Why do people discriminate against others?
- Why do people feel a need to belong to their group?
- How does group membership affect our behaviour?

What you will learn in this section

Key learning:

- Social identity theory (SIT) is better understood when placed in the context of previously existing theories of intergroup conflict: the aggressive personality theory and the realistic group conflict theory.
- Social identity theory is based on six key claims (examined in this section) and supported by studies using the minimal group paradigm. The main idea of the theory is that the mere perception of belonging to a group is sufficient to trigger out-group discrimination and in-group favouritism.
- Tajfel et al.'s classic experiments used the minimal group paradigm to demonstrate that people tend to maximize their in-group's profit at the expense of the out-group. They even tend to sacrifice their in-group's profit for the sake of positive group distinctiveness.
- Tajfel's theory has a large explanatory power. For example, it explains why conflict and discrimination can occur even in the absence of competition over resources between the two groups. It has also contributed to our understanding of stereotypes. It can be useful for designing strategies for overcoming inter-group conflict.
- On the other hand, Tajfel's theory has limitations: the self-esteem hypothesis is debated, the positive–negative asymmetry continuum has been discovered, and it has been argued that the theory's explanatory power is better than its predictive power.

Key terms: realistic group conflict theory, social identity theory, social categorization, out-group discrimination, in-group favouritizm, self-esteem, distinctiveness, social comparison, minimal group paradigm

In a wider context

Social identity theory (SIT) was developed as a theory of intergroup conflict: its aim was to explain why conflict and discrimination occur.

Social identity theory (SIT) is better understood when placed in the context of previously existing explanations of intergroup conflict, such as:

- Theories that stressed intrapersonal reasons (i.e., factors within the personality). According to these theories, group conflict emerges because groups are composed of people with "aggressive", "violent", or "evil" personalities. Such theories ignored social variables.

- **Realistic group conflict theory** (originated in the studies of Sherif, 1956). This theory suggested that intergroup conflict is caused by real conflict between group interests (competition over limited resources). In other words, groups conflict with each other because they want to possess resources (food, territory, energy, romantic partners) that are not abundant enough for everyone to possess them.

What is social identity theory?

SAQ Social identity theory

Social identity theory (Tajfel and Turner, 1979) is based on six key claims:

1. Competition over resources is not always necessary for the development of conflict between groups.
2. The mere perception of belonging to two groups (**social categorization**) is sufficient to trigger **out-group discrimination** and **in-group favouritism**.
3. Social categorization provides individuals with a means of building their social identity. People define themselves in social terms, in terms of being similar to or different from, "better" or "worse" than members of the other groups.
4. Individuals strive to achieve a positive social identity because it increases their **self-esteem**.
5. Positive social identity is based on **distinctiveness**: the in-group must be perceived as positively different from ("better than") certain out-groups. This involves the process of **social comparison**.
6. When social identity is not positive, individuals will try to either leave the group or make the existing group more positively distinct.

Some of these claims have been supported by research studies using the so-called **minimal group paradigm**. In this experimental paradigm, participants are randomly classified as members of two groups on the basis of a very trivial criterion (like the letter their name starts with or even based on a coin flip). The groups are purely cognitive as there is no objective reason for the group members to compete.

It is a good idea to thoroughly analyse one of Tajfel et al.'s (1971) studies presented in this section to understand what the theory is based on, as well as its inherent limitations. This will help you apply your knowledge of SIT to any unseen situation that can be presented in Paper 1B.

For Paper 1A purposes, however, you will be limited in time, so it is a good idea to know about a simplified real-life example to illustrate the six claims. Alternatively, practise explaining one of Tajfel's studies in just a few sentences.

Henry Tajfel (1919–82) was a social psychologist from Poland who was best known for his pioneering work on social identity, prejudice, and discrimination. He played the key role in creating social identity theory.

There is a lot to explore in this theory because it is a construct that has been shaped over decades in multiple research studies. However, let us start with the classic experiments in which the theory was born. These experiments do not support all six claims outlined above, but they certainly shed light on three claims: 1, 2, and 5.

Tajfel et al. (1971): minimal group experiment 1

The main theoretical claim that was tested in this experiment was claim 2. This is that the mere perception of belonging to two different groups is sufficient to trigger in-group favouritism and out-group discrimination, no matter how trivial the basis of group membership is.

Participants in this study were 64 male students from a school in Bristol (England), aged 14–15. They were tested in eight groups of eight participants.

Each group of eight was further split in half according to the principles of the minimal group paradigm. This means that the difference between the groups was not significant and only important in participants' minds but not in reality.

Here is how it was achieved in practice. At stage 1 (categorization), participants were rapidly shown 40 slides with clusters of dots on a screen and requested to estimate the number of dots. Participants thought that estimating the number of dots on the screen was the main point in that experiment. However, after the slideshow was finished, the experimenter told participants that there was another study going on (as in, "since you are already here, why don't you participate in the other study too?"). Participants were told that for convenience, they would be divided into groups based on the previous task—a group of four "overestimators" and a group of four "underestimators".

At stage 2 (distribution of rewards), participants were taken to another room one by one where they worked on their own in separate cubicles. They had to distribute rewards and penalties in real money to others. For this, they were given a booklet with 18 matrices and required to choose one of the options for the distribution of rewards and penalties. Table 5.1 shows an example of such a matrix.

You are in the group of: overestimators														
Rewards and penalties for:	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Member # 2 of your group	–25	–21	–17	–13	–9	–5	–1	0	2	4	6	8	10	12
Member # 4 of the other group	12	10	8	6	4	2	0	–1	–5	–9	–13	–17	–21	–25

▲ Table 5.1 Example matrix. The numbers on top (1–14) represent the choice that the participant is required to make. Based on: Tajfel et al. (1971)

In this matrix, as you move from left to right, the reward of your own group member increases, but the reward of the member of the other group decreases. Some matrices in the booklet involved two members of the different groups (as in the example given in Table 5.1). In some matrices, the choice was between two anonymous members of the in-group, and in some it was two members of the out-group. In the example above, the choice that gives maximum profit to an in-group member is 14, and the point of maximum fairness is 7.5 (that is, between 7 and 8).

Activity

Complete the following steps individually, then get together in a small group and compare your answers.

- How were the constructs of "in-group discrimination" and "out-group favouritism" operationalized in this study?
- How do we know that this study employed the minimal group paradigm?
- Why was the matrix designed this way?

ATL Research, Communication

Participants always allotted money to others and never to themselves. All participants were anonymous because names were replaced by codes. They were told that at the end of the task they would be given whatever amount of money others had allocated to them.

Results of this study revealed that the mean choice in the different-group matrices was 9.2 (on the scale from 1 to 14 as shown in the example in Table 5.1). The choices in the same-group matrices (two members of the in-group or two members of the out-group) were more closely clustered around the point of fairness (7.5). These results indicate that participants demonstrate in-group favouritism and out-group discrimination when categorized into groups based on trivial criteria.

Chat with AI

Social identity theory may be complex to explain, especially if you are writing an answer to a question in Paper 1A. Therefore, it is important to explain things concisely: briefly but without losing depth and clarity.

Practise writing a response to the following question: "Explain social identity theory using one example". Follow your comfortable pace and style of writing, but time yourself. Count the number of words in your response.

Your next task is to condense your answer to a maximum of 300 words (or however much you can write comfortably in 10 minutes). This is where AI can help. Try making your answer shorter yourself and at the same time ask your favourite generative AI to make it shorter for you. Compare your approach to "shortening" your response to that of AI.

What can you learn from this comparison?

Tajfel et al. (1971): minimal group experiment 2

To deepen the findings and to further test the claims of social identity theory, Tajfel et al. (1971) conducted another experiment published in the same paper. This time the key claim that was being investigated was claim 5 (suggesting that positive social identity is based on distinctiveness).



▲ Figure 5.1 Kandinsky's painting

Participants in this experiment were 48 boys of the same age from the same school. The procedure was largely the same as in the first study, but the way boys were categorized into groups and the matrices used in the booklets were different.

At stage 1 (categorization), participants were shown slides with pairs of abstract paintings—one by Kandinsky and one by Klee—and asked to choose the one they liked better. They were told that they would be split into groups based on whether they preferred Klee or Kandinsky, but in reality, the allocation was random.

The matrices at stage 2 (distribution of rewards) were manipulated in such a way that three different contradictory strategies could be used:

1. maximum joint payoff (MJP): choosing the option that provides the highest total number of points no matter the group membership
2. maximum in-group payoff (MIP): the best outcome to your own group, regardless of what the other group's outcome would be
3. maximum difference in favour of the in-group (MD): choosing the option that maximizes the difference between the payoffs for the members of the two groups.

Table 5.2 is an example of a matrix that was used in the study.

Rewards and penalties for:	1	2	3	4	5	6	7	8	9	10	11	12	13
Member # XYZ of your own group (Klee)	7	8	9	10	11	12	13	14	15	16	17	18	19
Member # XYZ of the other group (Kandinsky)	1	3	5	7	9	11	13	15	17	19	21	23	25

▲ Table 5.2 Example matrix. The numbers on top (1–13) represent the choice that the participant has to make. Based on: Tajfel et al. (1971)

In this matrix, as you move from left to right, MJP and MIP are maximized but MD is minimized. If you care about the biggest total reward and do not care about group membership, you will choose option 13. The same is true for MIP. However, if you mostly care about your group being better than the other group (more positively distinct from it) and you are even ready to sacrifice your group's profit for the sake of this positive distinctiveness, then you will choose options closer to the left side of the matrix. In option 1, your group performs seven times better than the other group.

Results of this experiment showed that participants demonstrated a clear preference for the MD strategy—that they preferred their group to get less money if this meant that their group would compare more favourably to the other group. Therefore, the study supports the idea that subjects sacrifice group and personal gain to achieve favourable intergroup differences—the idea of positive distinctiveness (claim 5 in the series of claims that we used to outline SIT).

Discussion

Why was it necessary for Tajfel et al. to conduct this additional experiment?

ATL Communication, Thinking

Evaluation of Tajfel's experiments

A major strength of procedure in Tajfel et al.'s (1971) studies was the high level of control over confounding variables. The experiments were designed very carefully to ensure that the mere perception of belonging to two different groups is the only thing that is actually different. Also note the sophistication and the

precision with which the variables were operationalized: the matrices were carefully designed and used in a very quantifiable way to measure in-group favouritism and out-group discrimination.

It is commonly noted that Tajfel's studies were conducted with only a very limited sample of participants, so generalizability is questionable. However, it is worth noting that Tajfel's original studies triggered a very large number of follow-ups and replications from other authors. The same hypotheses were tested in multiple variations of the minimal group paradigm, with multiple populations.

Another argument is that conditions in the study were quite artificial and did not closely resemble participants' normal everyday activities—this may limit ecological validity of the experiments. For example, teenage boys may be naturally competitive, and the matrices used in the study involved forced choice that was suggestive of competition. This may have triggered demand characteristics (i.e., the boys may have felt that it was expected for them to behave competitively in the experiment). However, this concern was addressed in follow-up studies as well. Some of these studies demonstrated that social categorization effects remain even when the groups are made to be explicitly random and forced choice is removed from the matrices.

Evaluation of social identity theory

Tajfel's theory has a large explanatory power. For example:

- It explains why conflict and discrimination occur even in situations where there is no competition over limited resources.
- It provides an explanation for how and why people respond to status inequality. For example, under what circumstances groups start engaging in protests.
- It has contributed to explaining stereotyping. From the point of view of an in-group, stereotypes may be reliable guides as to how to feel about and act towards members of another group. Since positive group distinctiveness is, according to the theory, important for our self-esteem, stereotypes too may be said to be "important" to our self-esteem.
- Social identity theory has played a role in developing evidence-based strategies of resolving intergroup conflict. For example, researchers of SIT have argued that if you want to reduce conflict between two groups, you need to find a way to make their group membership less noticeable or important. You can introduce another categorization that cuts across the existing groups. Suppose you want to reduce conflict between boys and girls in a middle school class, so you introduce the idea of "houses" and divide them into two mixed-gender teams. Another SIT-based strategy of conflict resolution that has been researched is redrawing the boundaries between groups in such a way that expands the in-group to include what used to be the out-group. Such is the phenomenon of former enemies teaming up against a third party.

Researchers have also outlined limitations of SIT:

- SIT attempts to explain social identity and behaviour through the concept of self-esteem. Two predictions can be made based on the theory: (1) your self-esteem will increase if your group is judged to be better than other groups; and (2) people from low-status groups whose self-esteem is initially low should seek more opportunities to present their group in a positive light in

Activity

Imagine you are a psychologist who has been invited to a school that is experiencing a problem with strong segregation among students (and teachers) belonging to two different cultural communities. You are tasked with designing a strategy to help overcome this segregation.

What suggestions do you have based on the principles and predictions of SIT?

ATL Thinking, Self-management, Communication

order to restore self-esteem to a normal level (Brown, 2000). However, these predictions have received little empirical support. Debate around the self-esteem hypothesis remains unresolved to this day.

- Another controversy that was uncovered is the so-called positive–negative asymmetry phenomenon. It was observed that if you conduct a study that is designed similarly to Tajfel et al.'s classic experiments but replace the matrices allocating rewards with matrices allocating punishment (such as exposure to aversive noise), then the effects of in-group favouritism and out-group discrimination are not nearly as strong. Therefore, it seems that discrimination and favouritism are evident in the positive domain, but not so much in the negative domain. Why would this be the case?
- Finally, it has been argued that the explanatory power of SIT (i.e., its ability to explain why a certain social behaviour occurs) is better than its predictive power (i.e., the ability to predict how exactly groups will behave in given social circumstances).

TOK

Why is the ability to predict important for a theory? Does it depend on the area of knowledge? For example, are there any theories in history that are able to make predictions?

Conceptual analysis

Perspective

The starting point of SIT is of course a social factor—the fact of belonging to a particular group. The end result is also social—it is our attitudes towards other people, our behaviour towards other people. However, the theory is also cognitive because it explains how we cognitively process this social fact. It is not group membership per se that matters, but perceived group membership. Perhaps it is this interplay between the social and the cognitive that gives SIT its great explanatory power.

Bias

The most obvious form of bias described (and explained) by SIT is the so-called "in-group bias", also referred to as in-group favouritism and out-group discrimination. It is a very general bias which accounts for other more specific forms of bias in human behaviour and judgement. For example, evaluating the performance of your team in a sports competition as more successful than it actually is.

As always, there is a possibility that the theory itself could be biased. For example, it seems like we might have overestimated originally how widely the theory applies. As you have seen in this section, doubts have been raised that SIT applies equally well to maximizing positive outcomes (distributing rewards) and avoiding negative outcomes.

Measurement

Theoretical constructs of SIT such as "in-group favouritism" are quite subtle and not easy to capture. A lot depends on how we operationalize them. From its onset, research of social identity has been characterized by rigorous operational definitions. For example, in Tajfel et al.'s (1971) studies "in-group favouritism" and "out-group discrimination" were operationalized as the participant's choice on a matrix that outlined various possible distributions of rewards between members of the two groups. Such an approach allows for precise measurement and easy replication, which contributed immensely to the advancement of research in this field over the years. The existence of a strictly defined "minimal group paradigm" also contributed to better replicability.

Causality

From the start, research supporting SIT has been strictly experimental. Conducting true experiments allows us to make cause–effect inferences with more confidence. Carefully controlling the experimental situation by following the principles of the minimal group paradigm ensures that we eliminate potential confounding variables and alternative explanations. This contributes to the theory's strong stance in science.

However, we have also considered a few challenges that SIT research is facing. One of these challenges is the mismatch between the theory's explanatory power and its predictive power.

Change

Social identity theory can be used purposefully to change human behaviour in desirable ways. Think about school education. We constantly promote the inclusion of children in various social groups: the school choir, extracurricular clubs, the house system within some schools, the "group" of scientists and academic workers, and so on. Perhaps by manipulating perceived group membership we can prevent such negative social phenomena as prejudice and discrimination and promote positive ones such as helping and pro-social behaviour? Racial segregation in classrooms is another problem that could be addressed.

Responsibility

Manipulating social identity is quite easy. We have seen from the research studies in this section that people actually cling to an opportunity to embrace another group membership that sets them apart from others. Think about boys in Tajfel's study who readily became "overestimators" or "Kandinsky lovers" and acted in a way that was discriminatory to the opposite group. If people so readily embrace an identity that is clearly made up and has nothing to do with their personality, how important would it be to them to cling to an identity that they may perceive as an expression of who they are, such as supporting a football team or being part of a musical subculture?

If introducing new identities is so easy and its effect on behaviour is so powerful, we must tread lightly. Any policy should be carefully analysed based on evidence, with potential costs weighed against potential benefits.

5.2 Cultural dimensions

Inquiry questions

- How can we describe the ways in which cultures differ from each other?
- Is it true that all cultural differences may be explained by a small set of parameters?
- Does social identity theory apply equally well to people from different cultures?

What you will learn in this section

Key learning:

- Cultural dimensions are a necessary component of the etic approach because they enable cross-cultural comparisons along a common set of constructs.
- Cultural dimensions were determined on a basis of a large cross-cultural survey.
- To determine the "dimensions" that underlie multiple cultural differences, researchers analyse the pattern of pairwise correlations between each pair of statements on the survey and identify groups of statements that are closely correlated with each other, but not with statements from other groups.
- Hofstede's model of cultural dimensions has been incredibly influential in research. It has a variety of practical applications. For example, in managing a multinational company.
- Limitations have also been pointed out. For example, the lack of representativeness of samples, the difference between individual and national level of analysis, within-nation heterogeneity, and problems with reliability and construct validity of the VSM (Values Survey Module).
- Social identity theory predicts that people will trust members of their in-group more than members of the out-group. Research shows that this depends on the cultural dimension of individualism–collectivism. For members of collectivistic societies, information about group membership may be far less important than knowledge about networks of personal relationships.
- Therefore, SIT may be more applicable to individualistic than collectivistic societies.

Key terms: culture, cultural dimensions, individualism versus collectivism, power distance, pairwise correlation, correlation matrix, factor analysis, reliability, construct validity, individual and national level of analysis, within-nation heterogeneity, VSM (Values Survey Module), trust and cooperation

In a wider context

One of the key aspects of human relationships is how an individual is influenced by the group. We study this aspect through several topics such as SIT, conformity, compliance techniques, and **culture**. When you grow up, it is impossible not to be influenced by culture.

As you know already (see Unit 2.7: Cultural factors in cognitive processes), the study of cultural influences on behaviour can be emic and etic. The emic approach to study a culture is to immerse yourself in it and to get an insight into its unique characteristics. This is done without imposing any concepts that could be alien to the culture. This approach allows deep insights, but it does not allow cross-cultural comparisons. To compare cultures to each other, we need to have a set of measures against which the comparison is performed, assumed to be universal to all cultures. That is exactly what **cultural dimensions** are.

What are cultural dimensions?

SAQ Cultural dimensions

The idea of cultural dimensions was proposed by Geert Hofstede who studied cultural differences and observed that they appeared to follow certain common patterns. Hofstede suggested defining culture as "the collective programming of the mind which distinguishes the members of one human group from another" (Hofstede, 1980, p. 25). In 1967–73, Hofstede conducted a worldwide survey of values among IBM employees. At that time IBM was a multinational corporation, with the total number of survey participants reaching 117,000 (Hofstede, Hofstede, and Minkov, 2010).

Results of the survey were analysed to identify hidden dimensions underlying observed patterns of responses (a statistical technique known as **factor analysis**). In order to confirm original findings and generalize them to other populations, a series of follow-up studies were conducted between 1990 and 2002. The samples included commercial airline pilots, students, and civil service managers. Around this time countries were also profiled against the value dimensions (value scores were established for each country).

Hofstede originally identified four dimensions but later he refined the theory and added two more (5 and 6 in the following list):

1. **Individualism versus collectivism.** In individualistic cultures, people define their personality in terms of their own personal characteristics, their success, and their unique features. Among the strongest values in society are personal autonomy, competitiveness, and self-sufficiency. In collectivist cultures, identity is linked to the social group, and values associated with belonging to a group take priority over personal values.
2. **Power distance** index (PDI). A higher PDI score indicates that hierarchy between the less powerful and the more powerful is firmly established in the society and rarely questioned. A lower PDI score indicates that people question authority and believe that authority figures are not particularly superior to themselves.

Exam tip

Note: you do not have to list all six cultural dimensions in your responses. It is equally acceptable to choose one and elaborate on it a bit more.

3. Masculinity versus femininity. Masculine values include achievement, autonomy, and competitiveness; feminine values include caring, cooperation, modesty, and compassion.
4. Uncertainty avoidance index. This refers to the extent to which people can tolerate ambiguity. Cultures that rank high on this index tend to avoid risk-taking and are intolerant of events that are unexpected or unknown. Cultures that rank low on this index are more tolerant of uncertainty and embrace new situations more easily.
5. Long-term versus short-term orientation. Cultures that score low on this dimension (short-term orientation) are conservative, they rely on the past for guidance, and they value traditions highly. A high score (long-term orientation) indicates a culture that is pragmatic and oriented towards the future and future challenges.
6. Indulgence versus restraint. Societies that promote indulgence allow a free gratification of basic and natural human desires, such as having a holiday or just having fun. Societies that promote restraint regulate gratification of needs with strict social norms.

Nation scores

After conducting the survey in representative samples in most countries of the world, Hofstede and colleagues arrived at national scores for each of the six dimensions. The national score on each dimension can take values between 1 and 100. For example, Guatemala scores very high (95) on power distance whereas Austria scores very low (11). Individualism is high in the USA (91) and low in Malaysia (26).

How can we measure dimensions?

What exactly is a "dimension" of cultural difference and how do you operationalize it? Every existing measure of cultural dimensions is a questionnaire with a set of statements that participants are required to self-rate. The most commonly used rating is the Likert scale (e.g., from "absolutely disagree" to "absolutely agree" with several gradations in between). The statements initially get formulated based on a theoretical construct.

Let's use a hypothetical example. Suppose you are building a questionnaire that will measure a range of (hypothetical) cultural dimensions. Here are the examples of statements you could include (given in the parentheses are names of cultural dimensions to which we believe each of these statements belongs):

1. Family well-being is more important to me than personal achievement (collectivism).
2. I cannot imagine myself behaving impolitely towards an older member of my community (collectivism).
3. Every holiday I try to be reconnected with my extended family (collectivism).
4. I think it is not acceptable to be friendly with my boss (power distance).
5. As a student I thought it was impolite to speak to the teacher unless asked (power distance).
6. I do not question decisions of my managers (power distance).

Discussion

The country-wide scores on cultural dimensions obtained in Hofstede's research can be found (and interacted with) on the website called Hofstede Insights.

Locate your country of origin and review its dimension scores. Do you agree? Does this go in line with your personal experiences? Partner up with someone in your class (preferably from a different cultural background) and exchange your impressions. Note: you can compare two cultures graphically on the website.

ATL Communication, Social, Self-management

After data is collected from a representative sample of participants, researchers calculate **pairwise correlations** between each pair of items. For example, the correlation between items (1) and (2) is:

- positive if there is a tendency for people who agree with the first statement to also agree with the second statement
- negative if there is a tendency for people who agree with the first statement to disagree with the second one (or vice versa)
- not present if there is no tendency: people who agree with the first statement are equally likely to agree or disagree with the second one.

Activity

This is a good time for you to refresh your knowledge about correlations. See Chapter 1 and re-read relevant material if you find it difficult. Here are the questions you should be comfortable answering:

- How can you graphically represent a correlation between two variables?
- What will a perfect correlation ($r = 1$) look like on this graph?
- What will a non-existent correlation ($r = 0$) look like on the same graph?
- What is considered a "large" correlation?
- What is considered a "small" correlation?

Your favourite generative AI is also a good option to help you decide the answers.

ATL Research, Self-management

The researcher then builds a **correlation matrix**—a table of pairwise correlations—and examines it. This is what a correlation matrix may look like in our simplified example (Table 5.3).

	1	2	3	4	5	6
1. Family well-being is more important to me than personal achievement.	1	0.65	0.78	0.21	0.15	0.16
2. I cannot imagine myself behaving impolitely towards an elderly member of my community.	0.65	1	0.80	0.23	0.11	0.09
3. Every holiday I try to be reconnected with my extended family.	0.78	0.80	1	0.13	0.16	0.22
4. I think it is not acceptable to be friendly with my boss.	0.21	0.23	0.13	1	0.87	0.79
5. As a student I thought it was impolite to speak to the teacher unless asked.	0.15	0.11	0.16	0.87	1	0.92
6. I do not question the decisions of my managers.	0.16	0.09	0.22	0.79	0.92	1

▲ Table 5.3 Hypothetical correlation matrix

If "dimensions" really exist, and if they are exactly what we predicted them to be, then the following conditions must be true:

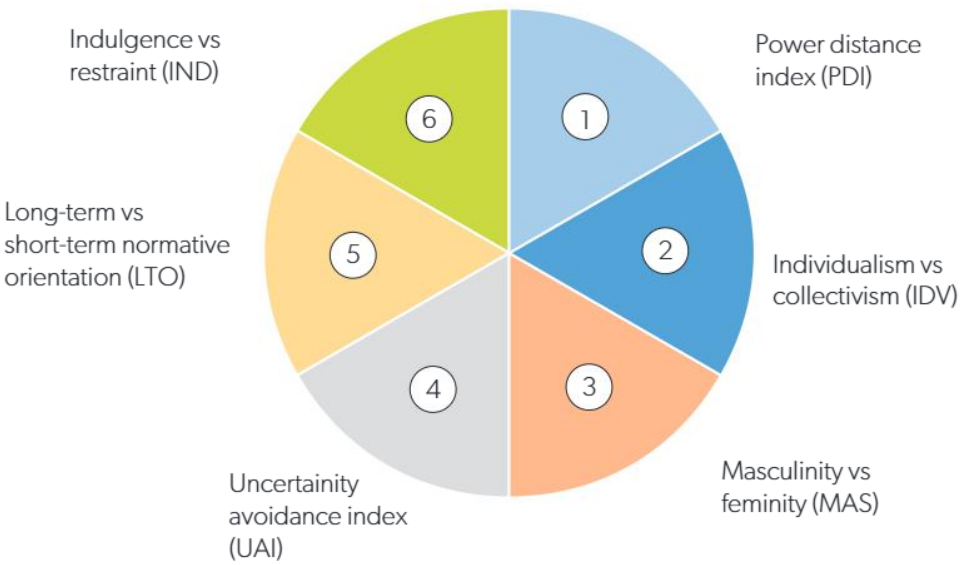
- Statements that belong to one hypothetical dimension should all be correlated with each other. For example, statement (1) should have a significant correlation with statements (2) and (3), and statements (2) and (3) should also be correlated.
- Statements that belong to different hypothetical dimensions should not be correlated with each other. For example, there should not be any significant correlation between statements (1) and (4), or between (2) and (6).

From the point of view of measurement, cultural dimensions are groups of statements on a questionnaire that are strongly interrelated while being relatively independent from all other statements.

In reality, researchers also use a more complicated statistical procedure called "factor analysis" that allows them to uncover the groups of interrelated variables in the correlation matrix more objectively.

TOK

Our discussion here links to the more generic problem of knowledge and measurement. Is what we know about things always determined by the instrument that we use to observe (or measure) them? Is it possible to know something directly, independently of the observation tool? Would the answer to these questions differ depending on the area of knowledge e.g., natural sciences, history, mathematics?



▲ Figure 5.2 Hofstede's six dimensions of culture

Evaluation of the model of cultural dimensions

Hofstede's model of cultural dimensions was originally developed in the context of business and organizational psychology. Its most obvious immediate applications were in multinational companies who work with partners and clients all over the world. For example, how much social chat should there be in a meeting before moving on to business discussions? What is the expectation about dining and gift-giving? Should you be reserved and polite in your demeanour or relaxed and friendly?

This model is also important in managing an organization. If your company employs people with various cultural backgrounds, it is crucial to create an environment in which people are comfortable with each other and simultaneously productive at their work. For example, there is evidence that the cultural dimension of power distance plays a role in how empowered people like to be in the workplace.

However, critics have noted that Hofstede's model of cultural dimensions has a number of limitations.

First, the representativeness of Hofstede's samples has been questioned. His original research, although very large in scope (over 100,000 participants) was conducted in a business setting, mostly among engineers and sales personnel working at one of the most elite organizations at that time. Due to biased employment opportunities at the time, most of Hofstede's participants were male.

Second, we need to clearly separate the **individual and national level of analysis**. Hofstede's theory and measurement tools were designed to derive national scores, not to assess individuals. For example, if a country such as Japan gets a high score on collectivism, it does not mean that every Japanese person shares collectivistic values. In fact, it is very possible that a randomly selected Japanese person will value independence and success, be tolerant to ambiguity, and will treat authority figures as equals.

Third, Hofstede's research ignores **within-nation heterogeneity**: it assumes that one nation has one culture. The theory of cultural dimensions is probably so famous and attractive because it provides an explanation for a complex variety of cultural variables by suggesting that they can all be understood in terms of only a small number of parameters (i.e., the theory is parsimonious). However, within-nation heterogeneity is lost in an attempt to achieve this.

Evaluation of Hofstede's surveys

Geert Hofstede's survey for measuring cultural dimensions is called the **Values Survey Module (VSM)**. The first version of the VSM was published in 1980, which was also the year when the cultural dimensions model was published. It is that version that was used to establish national scores on cultural dimensions. Other versions of the VSM were numbered 82, 94, 08, and 2013. At the time of writing, VSM-2013 is the most recent version.

Gerlach and Eriksson (2021) conducted an exploration of validity and **reliability** of VSM-2013. Results of their study suggested low reliability of scales and some problems with **construct validity**. They also noted that results of surveys conducted in various years in one and the same country could be quite different. It is complicated because the versions of the survey that were used in those years were also different. There may be two explanations for this:

(1) The survey is reliable, but cultural values themselves change in each society with the course of time.

(2) Cultural values are stable, but the surveys used to measure them are not reliable, so results fluctuate.

Unfortunately, there is not enough data for us to be certain which of the two explanations is correct.

Discussion

The full version of the VSM and the manual describing how to use it are freely available on the website of Hofstede's research group—just search for "Hofstede Values Survey Model". Review the survey questions. What are your impressions of the survey?

ATL Research, Communication

Activity

Construct validity is the extent to which the scales of the questionnaire measure what they are believed to measure (i.e., the underlying construct).

Consider the following four items from the VSM-2013 that comprise the power distance scale:

1. In choosing an ideal job, how important would it be to you to have a boss (direct superior) you can respect? (If you say that it is important, you score high on power distance).
2. In choosing an ideal job, how important would it be to you to be consulted by your boss on decisions involving your work (you score high on power distance if you say "not important").
3. How often, in your experience, are subordinates afraid to contradict their boss (or students their teacher)? (If you say "often" or "always", you get a higher score on power distance).
4. An organization structure in which certain subordinates have two bosses should be avoided at all cost (if you answer "yes", you get a higher score on power distance).

Now if you look at the theoretical description of the construct "power distance" and compare it to the four items that are used in VSM-2013 to measure this construct, do you think it is a good match? Do the items capture the essence of the construct as it is defined theoretically? Are the items sufficient to capture this essence? Is their focus too narrow? Should there be more of them? Do they actually measure what they are supposed to measure? All these and similar questions are questions of construct validity.

ATL Research, Thinking

Cultural dimensions and social identity theory

Knowing SIT, it seems reasonable to predict that people will trust someone they share the same group membership with more than a stranger. An interesting question in this regard is what is the role of culture in this? For example, when a group of Australians learn that they all come from the same university, will this increase **trust and cooperation** behaviour between them? How about the same scenario in a group of Japanese people?

Buchan et al. (2002) had participants play a game of investment. They found that, for participants from individualistic societies, the amount of monetary investment based on trust in unknown others increased when participants believed that all players shared the same social group identity. However, the same was not true for participants from collectivistic societies: their investment strategy did not change once they learned this new information. This suggested that knowing that others share the same social group identity did not make them trust these people more.

Exam tip

Material presented here assumes that you are familiar with SIT—see Unit 5.1 for details.

Other studies have confirmed this counterintuitive finding. For example, Yamagishi et al. (2003) compared Australian and Japanese participants and found that for Australians, learning that they share the same group membership (e.g., same country, same university) was sufficient to generate higher levels of trust in members of the in-group. However, this was not the case for Japanese participants (Brewer and Yuki, 2007).

Another finding was that participants from Japan (but not from the USA) demonstrate more trust and cooperation towards a member of the out-group if they learn that they share a common personal acquaintance with this person.

These findings suggest that trust is depersonalized for American participants: it is based on the mere fact of belonging to a social category. If they know that the person shares the same group membership as themselves, they will be more likely to trust that person. However, trust is personalized for Japanese participants. It is not so much affected by the knowledge of what social category the person belongs to, but it is affected by whether or not that person is connected to them in the network of personal connections. If they share a common personal acquaintance, there will be more trust and cooperation between them even though "formally" they may belong to different social groups (Brewer and Yuki, 2007).

Therefore, we might conclude that SIT is more suitable for individualistic than collectivistic societies, at least when it comes to explaining trust and cooperation between members of the in-group.

Activity

Design a research proposal for the investigation of cultural differences in shyness. It should use a specially designed questionnaire to measure shyness and it should establish nationwide scores of shyness across a variety of cultures.

What would be the requirements for your sample? What will you need to do to create the questionnaire?

ATL Research, Self-management

Conceptual analysis

Perspective

Cultural dimensions were proposed to explain how culture influences the behaviour of an individual. Therefore, they obviously "belong" to the sociocultural perspective on behaviour.

The concept of perspective also implies different theories or different ways of looking at the same phenomenon. The idea of cultural dimensions belongs with the etic perspective as opposed to emic. It implies that cultures can actually be compared against a set of dimensions that are universal to each of them. From the emic perspective, this approach bears the danger of "imposed etic" when we try to understand a culture through the lens of a concept that is alien to it.

Causality

Cross-cultural research is almost exclusively correlational. This is because we cannot manipulate culture as an independent variable (and even if we could, that would not be ethical). The value of the model of cultural dimensions is that it provides a framework through which to understand and describe existing cultural differences in behaviour. No claim is made that cultural characteristics influence behaviour. However, it is plausible to assume that they do. Why else would there be nation-level differences in the behaviour of large masses of people?

Bias

The danger of bias always exists in cross-cultural research. The idea of cultural dimensions itself may be viewed as a bias from the emic perspective (the imposed etic bias). Conclusions derived from surveys of cultural dimensions may be biased if surveys themselves are biased. The assumption that one nation represents one culture may be far-fetched. Within-nation heterogeneity is frequently ignored. When we apply results of research into nations on the whole to separate individuals, this can lead to biases. An individual who comes from a collectivistic nation does not necessarily have to share collectivistic values.

Measurement

In this section, we have seen that any conclusions about cultural dimensions are inseparable from our confidence that the surveys actually measure what we think they measure. Surveys are not static: they get updated, and therefore more recent data is obtained with the use of more recent surveys. When results of these recent surveys and the earlier ones do not match, it creates a dilemma: is it because culture has changed or because the survey itself lacks reliability? We have also seen here that the problem of construct validity is always pertinent in survey construction.

Change

The concept of change may be considered from the point of view of variation from culture to culture: cultural values change from one society to another. In this respect, the model of cultural dimensions is extremely helpful because it allows us to capture this change, categorize it, and make sense of it.

Another aspect of change is alterations over time. Cultural values can change gradually, perhaps now more so than ever with the spread of globalization processes and cultural interpenetration. This presents a problem for measurement.

Responsibility

The idea of cultural variation is a sensitive one because its misinterpretation or misrepresentation may lead to enhanced stereotyping and other negative consequences. It may be tempting to use the theory to oversimplify how we view others and make assumptions about individuals that are not justified. One needs to be careful in clearly delineating what the idea of cultural dimensions can and cannot be applied to, how it can and cannot be used.

5.3 Acculturation

Inquiry questions

- How do your cultural values change when you move into another culture?
- What is the relationship between acquiring values of the new culture and preserving the values of your culture of origin? Can you have both?
- What is the best strategy to adapt to a new culture?

What you will learn in this section

Key learning:

- Unidimensional models of acculturation assumed that adopting new cultural values and maintaining values of the heritage culture were two extremes of the same continuum: as you acquire one, you lose the other.
- Berry's two-dimensional model of acculturation assumes two independent dimensions that create four possible acculturation strategies: assimilation, separation, integration, marginalization.
- Two main approaches used to measure acculturation are questionnaires and proxy variables.
- Limitations of Berry's model of acculturation include: arbitrary cut-off scores for the dimensions, the assumption that the dimensions and strategies are universal for all types of migrants and contexts, and lack of support for the marginalization strategy.
- Schwartz's extension of Berry's model adds two extra sets of dimensions: the affected components (practices, values, and identifications) and the context of acculturation.
- The immigrant paradox is the finding that a greater degree of acculturation is associated with worse health outcomes. One way to explain this is that most research relied on proxy variables of acculturation and proxy measures implicitly assume the unidimensional model (e.g., more years of residence = greater acculturation = less attachment to the culture of origin).
- Remote acculturation is a modern form of acculturation where people adopt values, practices, and even identities of foreign cultures in which they have never lived.

Key terms: acculturation, enculturation, unidimensional model of acculturation, two-dimensional model, acculturation strategies, assimilation, separation, integration, marginalization, cultural practices, cultural values, cultural identification, acculturation context, immigrant paradox, remote acculturation, proxy measure

In a wider context

Acculturation is a process of adjusting to a new cultural environment. Most frequently it occurs when an individual moves from one culture into another—for example, as a result of migration.

This is to be distinguished from **enculturation**, which is the process of learning your own culture. Enculturation occurs when you grow up surrounded by cultural values and you acquire those values through direct instruction, social learning, and other mechanisms.

Models of acculturation

SAQ Models of acculturation

Prior to the 1980s, most studies of acculturation assumed that it is a unidimensional process: as a person acquires the norms and values of a new culture, they automatically lose connection with their country of origin. The more you are influenced by the new culture, the less of an influence you feel from your old culture.

This one-dimensional model of acculturation was criticized. It did not seem to adequately explain the richness of the process.

Berry (1992) proposed an influential theory of acculturation that classifies individual **acculturation strategies** along two dimensions:

1. The extent to which the person is receptive to the new culture.
2. The extent to which the person maintains links to their culture of origin.

A major development in this model is that these two dimensions are seen as independent of each other, so a person's individual score may be high or low on both dimensions or high on one dimension and low on the other.

The intersection of these two acculturation dimensions allowed Berry to distinguish four acculturation strategies:

1. **assimilation**—when an individual adopts the values and norms of the host culture but loses connection with and attachment to the culture of origin
2. **separation**—when an individual rejects the host culture and instead reinforces links with their culture of origin (this is often accompanied by immigration to cultural enclaves)
3. **integration** (also called biculturalism)—when an individual manages to both adopt the values and norms of the new culture and maintain close ties to their culture of origin
4. **marginalization**—when an individual rejects both cultures.

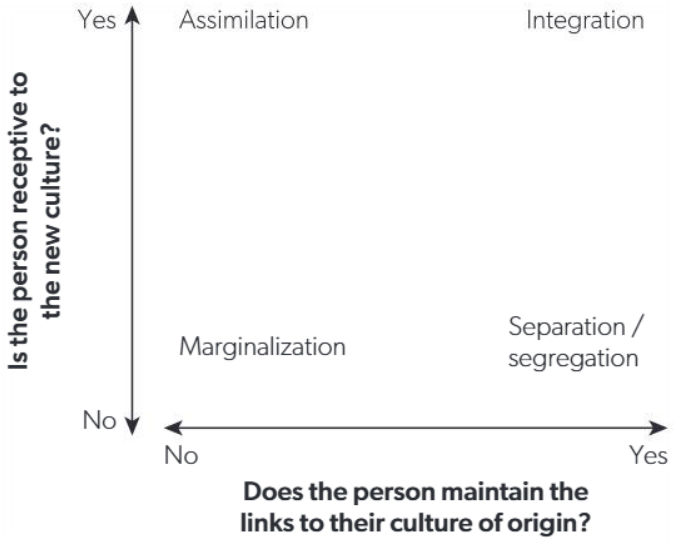
In a way, the two dimensions can be thought of as enculturation (maintaining the heritage culture) versus acculturation (seeking relationships with the dominant culture). This idea has far-reaching implications: it means that enculturation and acculturation are not necessarily opposite processes and can act together to pursue a common goal if cultural diversity is accepted in the society.

➜

Exam tip

Exam questions will not ask you to use more than one model of acculturation in your response (unless you want to). It is recommended that you focus on Berry's two-dimensional model and explain it. You will find several suitable examples in this section or you can come up with your own.

Later on in this section, we will also discuss Schwartz's model of acculturation which is a further development of Berry's model. You can choose to use Schwartz's model in your response instead.



▲ Figure 5.3 Berry's acculturation model

Plenty of research studies based on this model demonstrated that the strategy of integration is associated with most positive behavioural outcomes—for example, better health.

Measurement of acculturation

There are two main approaches used in research to measure acculturation: questionnaires and proxy variables.

Questionnaires can include items formulated around such things as participating in **cultural practices**, family beliefs, adherence to traditional values, and other "domains" of acculturation. The idea is that on the questionnaire, we ask participants a range of questions each tapping into a specific behaviour (e.g., "Were most of the movies you watched recently Western?"). The more questions they answer positively (or in a specific way), the higher their acculturation score will be.

Questionnaires that have been based on the **unidimensional model of acculturation** may be different from the ones based on the **two-dimensional model**, and results of the measurement will also reflect this. We observe what we expect to observe. For example, the Acculturation Rating Scale for Mexican Americans (ARSMA) asks participants to indicate their preferences between American and Mexican cultural items (food, music, language, TV) in a forced-choice kind of way. It is impossible to get high scores for both cultures.

In contrast, a different version of the same questionnaire called ARSMA-II assumes the bidimensional model of acculturation and includes parallel questions that mirror each other, such as: "I have difficulty accepting some behaviours exhibited by Anglos" and "I have difficulty accepting some behaviours exhibited by Mexicans". Participants need to answer both these questions (Cabassa, 2003).

A **proxy measure** is an indicator that is used when something cannot be measured directly. Proxy measures of acculturation include: generational status (e.g., first-generation migrant, second-generation migrant), proportion of life

Discussion

It is common for large-scale studies of acculturation that have large representative samples to use proxy measures. This is because it is practically difficult to administer questionnaires to large groups of people. In contrast, smaller studies with smaller samples more frequently use questionnaires.

In what ways do you think this can potentially bias our understanding of acculturation?

ATL Research, Communication

spent in the host culture, years of residence in the new country. These measures are not perfect because the amount of time you have been exposed to a culture does not directly translate into how deeply you internalize its values. However, it is useful for the research process and it is a fair assumption that a group of people who have lived in a country for 30 years would be more acculturated, on average, than a group of migrants who have just arrived.

Limitations of Berry's acculturation model

Schwartz et al. (2010) outlined several limitations of Berry's two-dimensional model of acculturation.



▲ Figure 5.4 Indian women celebrating Holi festival with traditional dresses and ornaments

First, the model relies on our ability to differentiate between "low" and "high" levels on each of the dimensions (maintaining the culture of origin and acquiring the host culture). However, how do we know what to consider a "low" level of maintenance as opposed to a "high" level? The way this is usually done is by taking the mean point in a certain population and using this point as the cut-off (e.g., anything below average is "low"). However, a "low" level of host culture acceptance in one population and a "low" level of host culture acceptance in another population are not necessarily similar. With this approach, conclusions such as "the integration strategy is more frequently used by Asian immigrants than Caribbean migrants in the USA" cannot be made because the comparison is invalid.

Second, the existence of marginalization as one of the acculturation strategies in Berry's model did not gain strong empirical support in subsequent studies. It turned out that the marginalization group is usually either very small or non-existent. Schwartz and Zamboanga (2008) explain that the likelihood that a person will develop a cultural sense of self with no inputs from either the culture of origin or the host culture is minimal.

Third, Berry's acculturation theory has been criticized for its one-size-fits-all approach (Schwartz et al., 2010). It assumes the existence of the same two dimensions and the same four acculturation strategies for all types of migrants, all host countries, and all ethnic groups.

TOK

Acculturation is a complex phenomenon that depends on many factors. Listing all these factors within one theory is impossible: it will make the theory clunky and unusable. Such a theory will not be able to make good predictions because its response to any question will be "it depends".

A good theory should be parsimonious: explaining a lot with a few variables. However, this raises the question: how many variables are enough, and how many are too many?

How can we approach this question?

Schwartz's extension of Berry's model

Schwartz et al. (2010) suggested a further expansion of Berry's two-dimensional model of acculturation. They claim that alongside the two dimensions in Berry's model, we should add two extra dimensions:

1. The exact components that are affected by the process of acculturation.
2. The context in which acculturation is taking place.

Schwartz et al. (2010) suggested the following three components that get affected by the process of acculturation: practices, identifications, values.

1. Cultural practices include the use of language, media preferences, participation in cultural traditions, and so on.
2. **Cultural identification** refers to attachments to cultural groups. A variety of research studies have shown a stronger ethnic identity (i.e., attachment to one's culture of origin) is typically associated with more positive health outcomes (i.e., better self-esteem and lower levels of obesity and substance addiction). However, the role of identification with the host culture is not yet fully understood. Does it provide a double benefit to identify yourself as a member of both cultures?
3. **Cultural values** refer to beliefs associated with various cultural dimensions (individualism–collectivism, masculinity–femininity, etc.). Someone who used to share collectivistic values of their culture of origin may start internalizing individualistic values of the host culture.

Given that acculturation is so multidimensional in nature, Schwartz claims that "acculturation is likely not a singular process that occurs at a single pace" (Schwartz et al., 2010, p. 248). Therefore it is an oversimplification to say that a person is "acculturated". We have to define which components (practices, values, or identifications) have or have not changed, and we also have to define how exactly they have changed along the two dimensions of accepting the new culture and maintaining the links to the culture of origin.

The other additional dimension (the context in which acculturation is taking place) refers to the situational variables that define the affordances and constraints, in a given location at a given time. The term "acculturation strategy" used in Berry's model implies that the migrant has a choice and that the outcome of acculturation is the result of this choice. However, this assumption is limited because depending on circumstances, migrants may experience certain constraints so they would not be entirely free to choose any strategy they want. For example, suppose there exists hostility in the host culture towards migrants of this ethnicity. This would make it difficult for people of this ethnicity to "choose" assimilation and integration as their acculturation strategies.

Activity

Consider the following two hypothetical individuals:

1. A second-generation Polish migrant in the USA. When her parents arrived in the country, they found it difficult to adapt. They hardly spoke English, had a strong accent, and clearly stood out. However, she was born and raised in the USA (this is what "second generation" means). Her accent is indistinguishable from the rest of the population and people assume she is American without question.

2. A second-generation Pakistani in the USA. She has fully acquired the host culture and has no accent at all, but it would still be common for her to be asked "Where are you from?" when making small talk, or to be complimented on her "good English".

How are **acculturation contexts** different for these two people? Analyse their situations using the concepts of Berry's and Schwartz's models. Which acculturation strategies would they be most likely to choose? Which of the three components (practices, identifications, or values) are likely to be affected differently?

ATL Thinking, Research, Self-management

The immigrant paradox

Many studies of acculturation have focused on health outcomes, such as stress, depression, or overeating. A well-known finding that emerged from such studies is the "**immigrant paradox**": a greater degree of acculturation is associated with worse health outcomes. For example, migrant workers with many years of residence are more likely to be obese than recent arrivals (Shah et al., 2015). Hispanics born in the USA are more likely to be diagnosed with mental disorders than those born abroad (Alegría et al., 2008), Hispanic adolescents who engage in Hispanic cultural practices are less likely to have addictive behaviours (Allen et al., 2008). There have been many findings like this from various ethnic and acculturation contexts.



▲ Figure 5.5 A protest against anti-immigration policies

Schwartz has noted that a key problem with the research that identified the immigrant paradox is that such research has relied—often implicitly—on the understanding of acculturation as a one-dimensional process. For example, suppose a study uses "years of residence" as a proxy measure of acculturation

and the finding is that there is a negative correlation between years of residence and a health outcome. Now that we accept a two-dimensional view of acculturation, how can we interpret this finding? It could be that the poorer health outcome is the result of acquiring the values of the new culture. Or it could also be the result of losing connection with the culture of origin.

Or it could be both.

The problem is that the majority of studies that led to the identification of the immigrant paradox have used proxy measures of acculturation, so interpretation of these results is ambiguous. We know that the effect exists, but we do not know exactly what is responsible for it. By extension, we do not know what the practical outcome of such studies should be: should we discourage migrants from adopting the new cultural practices? Should we encourage them to preserve social ties to their culture of origin? Should we do both? Should we use different tactics with different groups of migrants?

Remote acculturation

You do not have to physically move to another country to experience the effects of acculturation. In today's interconnected world, other cultures will reach you where you are. Ferguson, Ferguson, and Ferguson (2015) investigated this in their study of "**remote acculturation**" in a sample of urban Zambian adolescents.

Remote acculturation is a modern form of acculturation that is the result of the globalization process. People can adopt values, practices, and even identities of foreign cultures in which they have never lived. A sample of adolescents in Lusaka (the capital of Zambia) is unique for this purpose. Zambia saw a very rapid economic development with profound international influences from several key countries. Urban Zambian adolescents become exposed (through media and otherwise) to multiple foreign cultures at the same time. The most notable impact comes from the USA, the UK, and South Africa. The authors refer to this multidimensional acculturation jokingly as "I am AmeriBritSouthAfrican–Zambian".



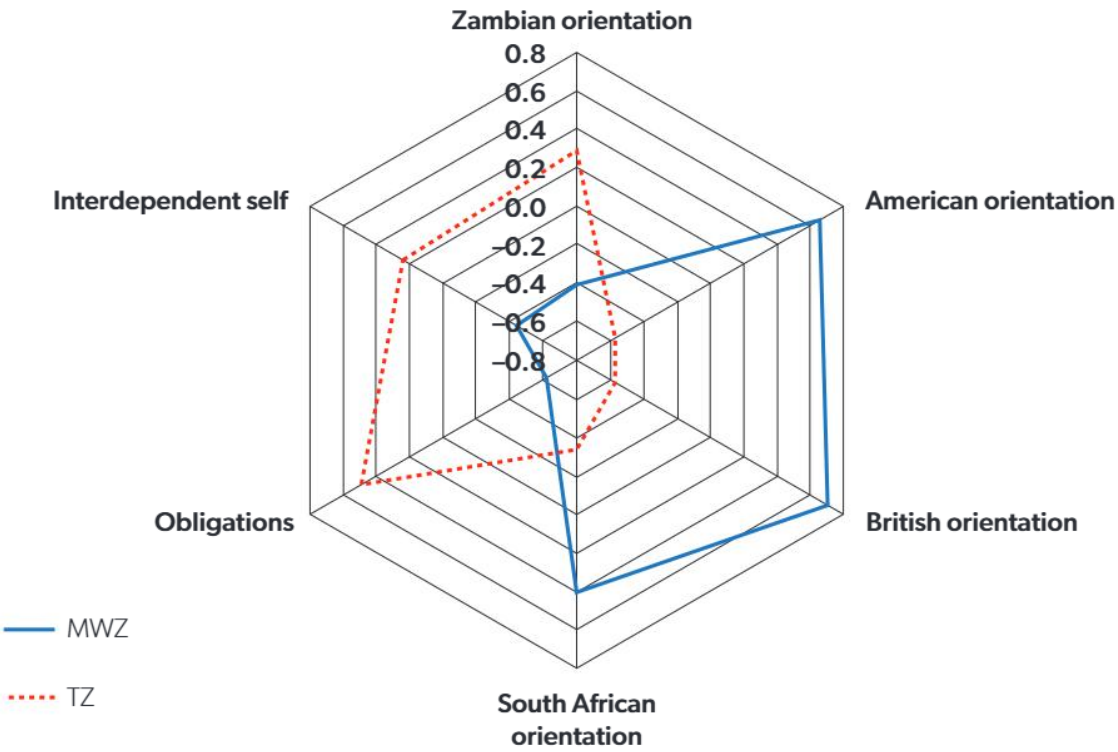
▲ Figure 5.6 Zambian adolescents

Researchers in this study measured three components of acculturation as recommended by Schwartz et al. (2010): behaviours, identity, and values. The sample consisted of 83 adolescents from a private school in Lusaka. All participants were born and raised in Zambia.

Analysis of data revealed several interesting findings.

First, Zambian adolescents differentiated very well among the three foreign cultures: South African, UK, and USA cultural influences were perceived as independent from each other. For example, participants who reported enjoying UK media did not necessarily enjoy American media as well, and vice versa. Therefore, adolescents clearly felt three distinct cultural influences rather than one single influence of "foreign culture".

Second, analysis of groups revealed that participants could be divided into two distinct groups: "Westernized Multicultural Zambians" (WMZ, 45% of the sample) and "Traditional Zambians" (TZ, 55% of the sample). The way they differed from each other is presented in Figure 5.7. TZ participants had a stronger orientation towards the Zambian culture and a weaker orientation towards the other three cultures (UK, USA, South Africa). Simultaneously, WMZ participants felt less obligated to help their families and viewed themselves as more independent from society.



▲ Figure 5.7 Results of the study: Ferguson, Ferguson, and Ferguson (2015)

Conceptual analysis

Perspective

Clearly the process of acculturation belongs with the sociocultural approach to behaviour. The existence of acculturation as a phenomenon in the first place is proof that sociocultural factors have an effect on people.

However, remember that the concept of perspective is related to various theories or explanations of the same phenomenon. There are perspectives on acculturation itself. Acculturation was initially viewed as a one-dimensional phenomenon in which the acceptance of a new culture implies weakening ties with the culture of origin. It is now increasingly recognized that acculturation is a much more complex process which unfolds along multiple dimensions, exists in multiple contexts, and includes multiple components.

Causality

Acculturation itself is a process that is not "caused" by anything: when you move into another culture, you go through the process of acculturation by definition. However, different people go through the process differently, and it makes sense to ask what causes these individual differences in acculturation strategies and outcomes.

Research has uncovered multiple factors that are associated with the choice of acculturation strategy. Some factors such as perception of immigrants in the society do not depend on the individual—they create the context in which acculturation takes place. Other factors seem to depend on the individual's personal choice.

Another question is, what outcomes are caused by acculturation? A lot of research has focused on health outcomes: mental disorders, addiction, obesity, and so on. Findings suggest that health outcomes are indeed influenced by the choice of acculturation strategy.

Bias

As we have seen in this section, there are so many complexities surrounding the process of acculturation that it is easy for a research study to overlook the influence of something important and therefore create biased results.

For example, some studies have assumed the unidirectional model of acculturation and have not considered the possibility that a person can both adopt new cultural values and maintain the connection to the old ones. Some studies have focused on one component of acculturation (i.e., changing practices) but overlooked other components (i.e., changing values and identifications). The samples have also been limited, with most acculturation research conducted in the context of cultural minorities adjusting to their life in the USA or Europe.

Measurement

Measurement of acculturation is performed in one of two ways: either by a questionnaire or by using a proxy variable.

Proxy variables are problematic because using such measures implicitly assumes the unidirectional model of acculturation (the longer you reside in a country, the more you accept the new cultural values and the more you weaken the ties with your culture of origin).

However, questionnaires also have their challenges. Every questionnaire is built on the basis of a theory. You can only measure the reality of things as seen through the lens of that specific theory upon which a questionnaire was built. Additionally, questionnaires are self-report measures. This makes them susceptible to a variety of response biases.

Change

Acculturation changes the individual in many ways: their behaviours, values, and identities. Acculturation also changes society on the whole: not only the minority, but also the majority culture of the host country.

We usually speak about "change" to describe either the naturally occurring modifications in behaviour or the purposeful actions that we can take in order to influence such changes. Acculturation is not something we can change in order to influence people's behaviour in desirable ways. Acculturation just happens. However, once we understand more about the flow of acculturation and how various contexts are associated with various outcomes, we can make changes in the context to ensure that acculturation flows in desirable directions and leads to positive outcomes.

Responsibility

There is a crucial transition between research findings and implications for practice. Suppose research studies have found that individuals who have weakened ties with their culture of origin demonstrate more negative health outcomes than those who have maintained the connection to their cultural roots. What does this finding imply for practice? What should we do now that we have this knowledge? Should we encourage the creation of cultural enclaves? Should we increase the status of minority languages?

Researchers are bearers of important information about the process of acculturation that policymakers may not have. It is the responsibility of researchers to bring this knowledge to the public in a clear and unbiased way, and to advocate for the social changes they believe their findings suggest.

5.4 Conformity

Inquiry questions

- What factors make people more likely to conform to the group?
- Is there conformity in animals?
- Can people conform to robots?

What you will learn in this section

Key learning:

- Conformity is a type of social influence that occurs in the absence of a direct request. There could be two reasons behind conformity: informational influence and normative influence.
- Conformity links to the principles of SIT. Groups we identify with are the most powerful influencers.
- Many experimental studies of conformity have followed a standard research protocol called the Asch paradigm.
- Follow-up interviews in Asch's classic 1951 study revealed that there were three explanations for conformity: distortion of perception, distortion of judgement, and distortion of action.
- Multiple variations of the classic Asch paradigm revealed a large number of additional findings, such as the relationship between conformity and number of confederates, and the effect of the presence of another participant giving correct responses.
- It has been argued that Asch's research has come to be misrepresented over time. Introductory psychology textbooks emphasize the negative conclusion (that people conform) and downplay the other side of it (that some people never conform). This tendency is stronger in more recent textbooks.
- Conformity is not a uniquely human phenomenon. It may have had a survival value and may be a product of evolution.
- People can conform to robots as long as robots are perceived as competent in solving the task.

Key terms: conformity, social influence, informational influence, normative influence, Asch paradigm, distortion of perception, distortion of judgement, distortion of action, biological basis of conformity

In a wider context

Conformity is one of the ways in which the group can influence an individual. Some other ways are direct persuasion, manipulation, compliance techniques, and others—these forms of influence assume that one person purposely tries to change the behaviour of another. Of course, people are also influenced by others

indirectly, by simply being among them. For example, belonging to a group forms our social identity, and social identity, as we have seen, is a powerful driver of behaviours such as in-group favouritism and out-group discrimination.

Definition and types of conformity

SAQ Conformity

Conformity is when people change their beliefs, attitudes, or behaviour to match those held by the group. Conformity is a type of **social influence**. A defining characteristic that sets conformity apart from other forms of social influence is that it occurs in the absence of a direct request or even an intention to influence the individual. People adjust their beliefs and behaviour by their own accord, simply because they perceive these new beliefs and behaviours to be more in line with the group's norms.

When analysing reasons behind conformity (i.e., what makes people conform?) a distinction is made between informational influence and normative influence.

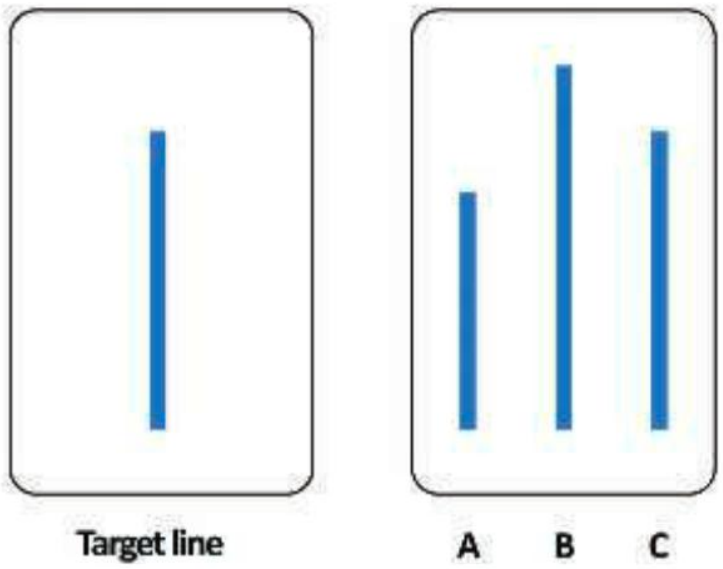
- **Informational influence** occurs when people conform to the group because they are uncertain about their own beliefs and so they operate on the assumption that the group's accepted belief is more accurate than their own. When you are uncertain about your own judgements, you are more likely to be influenced by the group's "default".
- **Normative influence** occurs when people conform to the group because they feel the need to be accepted by this group. The desire to be accepted by other people can be a powerful motivator of behaviour. The reason normative influence works is because people tend to believe that others will accept or like them more if they share the same beliefs, attitudes, or behaviours.

Conformity links to the principles of SIT. This is because we do not equally conform to all groups. Groups that we consider to be important and we consider ourselves part of (i.e., in-groups) are the most powerful influencers. Disagreement with such groups will negatively affect our self-esteem because of the way we identify with them.

In this section, we will consider a series of research studies that followed the so-called Asch paradigm. The Asch paradigm could be a good example of conformity to use in Paper 1A. It does not have to be any specific study.

Classic study of conformity: Asch (1951)

The most well-known classic study of conformity is the Asch conformity experiment. If you were a participant in this experiment, you would be invited to take a "vision test". After arriving at the laboratory, you would find yourself in the company of seven other participants and would be required to engage in a line judgement task. The task would include two cards—one card with a reference line and the other card with three comparison lines of varying lengths (see Figure 5.8). You would have to tell which of the three comparison lines was the same length as the reference line.



▲ Figure 5.8 Line judgement task

Discussion

In the previous IB Psychology syllabus, students were required to carry out a simplified replication of an existing research study as their coursework. It was an IB rule that studies of conformity were not allowed for ethical reasons.

Why do you think this was the case? Which ethical principles does the Asch paradigm clash with?

ATL Communication, Self-management

What you would not know, however, is that the other seven participants were all research confederates (actors) who knew in advance what responses they would be giving. You would always be the last one to provide an answer, and the confederates would sometimes confidently provide answers that are obviously incorrect. You would not know that the true aim of the experiment is to see whether or not you would also give an incorrect answer under the influence of the group.

This classic experimental procedure became so widely used in subsequent replications and follow-up studies that it is now referred to as the **Asch paradigm**.

Participants in Asch's (1951) study were male students at a college in the USA. All confederates in Asch's first study provided the same response, but on some trials they responded incorrectly. Participants completed a total of 18 trials. Confederates gave the incorrect response on 12 out of these 18 trials. After the line judgement task, all true participants had a post-test interview where the real nature of the experiment was revealed to them and they were asked a series of questions on what they were thinking about while engaging with the task.

There were 50 "genuine" participants in this study. An additional 37 participants were allocated into the control condition where they went through the same procedure, but without confederates and with only themselves and the experimenter present in the room.



▲ Figure 5.9 Participants in a conformity study (only the person sitting on the furthest right is a real participant)

As expected, participants in the control condition were almost never wrong on the critical trials (error rate <1%). In the experimental condition, 35.7% of responses on critical trials were incorrect, and 74% of participants overall gave at least one incorrect answer on one of the 12 critical trials. If the number does not seem too impressive, remember that the line judgement task was very simple and obvious: it is as if participants chose not to trust their own eyes just because a group of people they had never met before behaved in a particular way.

Results also revealed considerable individual differences:

- 12% of participants conformed to the group on all or almost all trials.
- 26% always stuck to their judgement and never conformed.

Analysis of the follow-up interviews led Asch to conclude that one of three things was happening to those participants who conformed to the group on all or some of the trials:

1. **Distortion of perception**: this is when participants appeared to be unaware that the judgements of the group were incorrect; they believed that confederates were giving correct responses all this time.
2. **Distortion of judgement**: this is when the participant believes that the group is giving an incorrect response, but starts doubting their own judgement due to the pressure. They hesitate and think, "What if I'm wrong and these other participants know better?"
3. **Distortion of action**: this is when participants believe what they see and their judgement is not swayed by the behaviour of the group. Nevertheless, they still change their action (response) in order to avoid the embarrassment of having to publicly oppose the opinion of others.

Activity

If you had to express the results of Asch's (1951) study in only one sentence, what would this be?

Compare your sentence with those of your classmates. Are there differences? What does it tell you about the transition from data to the conclusion? Is this transition vulnerable to bias?

ATL Communication

Chat with AI

Asch's paradigm has been used in a variety of modifications to test a large number of related hypotheses.

1. Use a generative AI with a function of providing references to investigate this topic. Here is a possible prompt: "Please give me a list of eight different modifications to the classical Asch paradigm of conformity (the line judgement task with seven confederates) that were used in research studies. For each of the modifications, formulate the main research finding in one sentence and provide a link to the source where I can find the relevant study."
2. As you keep studying this section, compare your findings to the material presented here.

Remember: generative AI is prone to errors in some cases, so do not blindly trust what it says. It is always useful to check the primary sources that the AI will provide.

Variations of the classic study

In follow-up studies, multiple variations of the Asch paradigm were tested. Here are some of the findings from these follow-up variations:

- The presence of another real participant in the group substantially decreases conformity. The same effect is achieved when one of the actors is instructed to give correct responses on all trials. In follow-up interviews, participants mentioned a feeling of warmth or closeness to this other person, yet they failed to recognize that the presence of this person may have somehow influenced their behaviour.
- The size of the majority group had an influence on conformity levels. However, this relationship between group size and conformity gradually slowed down and adding more people beyond seven people already made no difference.
- Conformity decreased when the participant had to provide the response in writing rather than saying it out loud (although all confederates spoke) (Asch, 1955).

There is a debate regarding the most probable explanation of Asch's findings: are they best explained by normative influence or informational influence?

Do we just pretend to be like the others in order to blend in and be accepted, or does the influence of the group actually alter our perception of things?

Interpretations of Asch's results

Another interesting angle in considering results of the classic experiments using the Asch paradigm is how you choose to interpret them. Do you emphasize the fact that 74% of participants conformed at least once during the 12 trials, suggesting that people are very conforming? Or do you emphasize the fact that 26% of participants never conformed, not even on a single trial, despite having to go repeatedly and publicly against the opinion of seven other people in the room? This latter interpretation would suggest that we are actually quite resistant to group influence and there are individuals among us who may not be swayed at all.

Friend, Rafferty, and Bramel (1990) analysed 99 introductory psychology and social psychology textbooks published between 1953 and 1984 and found that these textbooks repeatedly emphasized the conformity aspect of Asch's findings (that some people conformed) and repeatedly downplayed the independence aspect (that some people remained independent in their judgement). Moreover, this tendency appeared to become stronger over time. Griggs (2015) continued this line of research and added 20 introductory psychology textbooks that were current at the time of the study (i.e., published from 2012–14). His results largely coincided with the prior studies and showed that the trend was continuing. For example, only 15% of the textbooks mentioned the existence of totally independent participants. Only one of the 20 textbooks discussed the post-experimental interviews in Asch's classic studies.

TOK

Can textbooks be gatekeepers of knowledge? What other gatekeepers are there? Think about the process of writing a textbook (e.g., IB Psychology), especially selecting the material for it. Inevitably there is the problem of material selection: how do you decide what to cover, which studies to include and which to exclude, how to write overall conclusions? How objective is this process? Even if it cannot be entirely objective, are there any ways for us to reduce bias and subjectivity in this process?

Biological basis of conformity

If conformity is a common feature of human behaviour, it must have somehow evolved in our species. Therefore, it must have served some sort of survival function (Morgan and Laland, 2012). This survival function would be the **biological basis of conformity**.

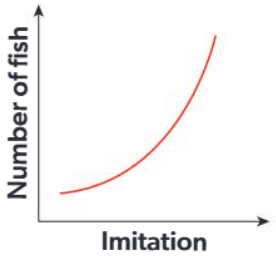


▲ Figure 5.10 ***Pungitius pungitius*** (nine-spined stickleback)

Pike and Laland (2010) investigated the copy-the-majority learning strategy in the nine-spined stickleback—a species of freshwater fish. Their study used a fish tank divided into three sections, with a feeder placed at each end of the tank. One of the feeders dispensed food frequently and the other one rarely did. The study followed three stages:

1. Training: fish could swim freely and work out which of the feeders was more abundant in food.
2. Conflicting public information: the fish were confined in the middle section and they observed two shoals of same-species fish feeding at the two feeders, but the rich and poor feeder were swapped compared to training.
3. Test: fish were allowed to swim freely again and it was recorded which of the feeders they prefer—the one consistent with their own experience or the one where other fish were feeding.

Results of this study showed that nine-spined sticklebacks indeed prefer social information to their own experience. Importantly, the number of fish in the shoal greatly affected imitation: the likelihood that the participant fish would swim to the same feeder increased exponentially with the number of fish. This demonstrates conformity in freshwater fish.



▲ Figure 5.11 Exponential dependence of probability of imitation on the number of fish in the shoal

Activity

How can you operationalize informational conformity and normative conformity so that they could be measured separately?

Think about a typical Asch paradigm experiment where participants take turns to say their answer out loud versus a similar experiment where participants do not speak but write their answer down.

In the case of speaking (but not writing), participants will know the judgement of others and this will create informational conformity. However, when participants have to say their answer out loud, they may experience a pressure to conform because speaking against the group may be uncomfortable.

Therefore, the same experimental manipulation affects both informational and normative conformity. Can you think of a way to manipulate one without manipulating the other?

ATL Research, Thinking

Discussion

One might wonder what place a study with freshwater fish has in a psychology textbook. What is the value of animal studies like this for understanding human behaviour?

ATL Research, Thinking

Conformity to robots

Salomons et al. (2021) discovered that people conform not only to other people, but to robots as well. Moreover, robots cause both informational conformity (believing the group to be correct) and normative conformity (feeling peer pressure). In their experiment, participants sat around a table with three small yellow My Keepon robots (see Figure 5.12). Participants played a game with questions that did not have a clear correct answer. In each of the 20 rounds of the game, participants first gave their preliminary answer, then learned about the answer given by the robots, and then gave their final answer.



▲ Figure 5.12 My Keepon robots used in the study

The three robots were given different appearances (dressed in different clothes), names (Julia, Chuck, Paul), and personalities (one made jokes, another one was shy). They made utterances such as "Hmmm" or "I am not sure about this one", but on critical trials they were programmed to remain silent to avoid any confounding variables. There were six critical trials in total where all robots gave the same response and this response was different from that of the human participant. In some trials if the human participant gave an "incorrect" answer, the robots turned their heads and stared at the participant for several seconds to add social pressure.

In line with what you would expect in a regular "human" conformity study, results showed that participants were more likely to change their initial answer if they knew that all the robots answered differently. The addition of staring somewhat increased this likelihood, but the increase was not statistically significant.

Conceptual analysis

Perspective

Conformity is a type of group influence on individual behaviour. Therefore, formally speaking, the study of conformity belongs with the sociocultural approach (the "social" aspect of it). However, we have seen once again in this section that all perspectives are closely intertwined. Conformity is not a uniquely human behaviour. Similar patterns of it also show in other species (such as freshwater fish). Therefore, conformity may have evolved as an adaptation, and it may have biological roots. Analysis of neural correlates of conformity reveals that different aspects of this behaviour are more integrated than we might think based on social psychology alone.

Causality

When we ask, "What causes conformity?", this question may mean several things:

1. Are we asking about environmental conditions and survival mechanisms that have favoured and selected conformist behaviour?
2. Are we asking about generic universal factors that change the likelihood of conformity in a given situation (e.g., group size or ambiguity of the task)?
3. Are we asking about individual differences—that is, what causes some people to conform and others to be independent even when they are placed in the same situation?

Some aspects of causality lend themselves more easily to experimentation, some less so. For example, universal factors that affect the likelihood of conformity can be manipulated: we can run the same experiment with different group sizes, different types of the task, and so on. Other aspects such as pre-existing individual differences cannot be manipulated. All we can do is establish a correlation between them.

Measurement

In the Asch paradigm, conformity was clearly operationalized and measured. Asch used the line comparison task because it involved a simple perceptual judgement that was not likely to be influenced by confounding variables. One could argue that such a task is a very narrow operationalization of conformity and that in real life, conformity takes many different shapes and forms. Narrow operationalizations are a common problem for construct validity.

Bias

Bias can manifest in a variety of ways and at various stages of the research process. One such example that we have seen in this section is the transition from study findings to a theoretical interpretation of their significance and implications. Classic research studies of Asch have been interpreted to suggest two diametrically opposite conclusions:

1. People are very likely to conform to the judgement of others even in the most obvious perceptual tasks.
2. People remain independent even when they have to repeatedly voice an opinion that goes against the majority.

Research findings and conclusions are different things, and it is often possible to derive different conclusions from the same research findings—that is, to interpret them differently. Of course, one's theoretical predispositions, cultural expectations, and even political agenda may influence the way we interpret the findings.

Change

Conformity behaviour is observed in animals as well as humans, and therefore it has been undergoing changes in the process of evolution. It is hard to establish what these changes are. However, it is remarkable that patterns of conformity established in species as unrelated to us as fish are almost the same. Could it be that conformity as a basic learning behaviour or adaptation mechanism has remained unchanged over millions of years?

There has been some research focusing on how conformity develops over the course of a human life—that is, at what age children start to exhibit conformity and how this changes with age.

Like many other behaviours, rates of conformity are not the same across different human populations. It has been established that culture plays a certain role here, with some cultures more conforming and some less.

Responsibility

Participation in a study of conformity could cause a certain amount of psychological harm. Moreover, a study of conformity involves a fair amount of deception. The real participant is not told that all the other people in the room are actually researcher's confederates. The participant may also experience discomfort when they are put in a situation where they have to publicly go against the opinion of the majority. When results of the study are later revealed to them in a debriefing session, their self-esteem may be affected. In any case, the study of conformity emphasizes the fact that "protection from harm" is an important ethical imperative and that "harm" should not be understood as merely physical.

5.5 Compliance techniques

Inquiry questions

- What is the best way to convince someone to comply with your request?
- Why do compliance techniques work?
- Do compliance techniques easily generalize from one situation to another?

What you will learn in this section

Key learning:

- Compliance is a form of social influence. Unlike conformity, compliance includes a direct request from one individual to another.
- Cialdini formulated six key principles of compliance that underlie most compliance techniques and explain their effectiveness: reciprocity, commitment and consistency, social proof, authority, liking, scarcity.
- Foot-in-the-door is a compliance technique when you first formulate a small request and then, after the person has agreed, you formulate your actual bigger request.
- Studies show that familiarity alone does not explain the effectiveness of foot-in-the-door. The simple act of agreement to comply with a smaller request does play a role, but actually carrying out the action is also important.
- The low-ball technique is different from foot-in-the-door because at the first stage people have already agreed to the target request, but at the second stage this request becomes more costly.
- Some common problems in this area of research are: population validity (students), ecological validity (demand characteristics), and construct validity (narrow operationalizations).
- These problems can only be removed by combining evidence from multiple studies.

Key terms: compliance, conformity, social influence, reciprocity, commitment and consistency, social proof, authority, liking, scarcity, compliance techniques, foot-in-the-door, low-ball technique, cognitive dissonance

In a wider context

Compliance is a form of social influence. Compliance is change in behaviour that occurs in response to a request. It is the presence of a request that makes compliance different from conformity. Another form of social influence apart from conformity and compliance is persuasion.

What are compliance techniques?

SAQ Compliance techniques

Compliance techniques are one of the content points that can appear in questions in Papers 1A and 1B. You will need to be able to explain what compliance techniques are and provide one example (in Paper 1A) or apply your knowledge to a given example (in Paper 1B).

Compliance is a form of **social influence**. It is a change in behaviour that occurs in response to a request. Compliance should be differentiated from **conformity**: conformity is when a person changes behaviour because all other people behave in a particular way, but those other people do not purposefully try to influence the individual and do not ask them to do anything. With compliance, things are different: there is a direct request to change behaviour.

The request may be either explicit (e.g., a call for donation) or implicit (e.g., advertisement). The individual may or may not be aware that they are the target of a **compliance technique**.

A compliance technique is a specific strategy used to elicit the desired response from an individual. There are many compliance techniques, for example: foot-in-the-door, door-in-the-face, low-balling, "that's not all", and others.

Robert Cialdini (1984) formulated six key principles of compliance that underlie most compliance techniques: reciprocity, commitment and consistency, social proof, authority, liking, and scarcity.

Real-life examples of compliance are very diverse. Think about buying an item because a salesperson encourages you to do so, helping a friend because they ask you for a favour, following a link suggested by an advertisement that you saw on a website while browsing. What is common across all these examples is that there was an intent to influence you in a very particular way, to get you to do something specific.

In this section, we will discuss the following examples of compliance techniques: foot-in-the-door and low-balling. For Paper 1A, you will only need one example—choose your favourite. For other papers (such as Paper 1C), you are not limited as to how many examples of compliance techniques you can use in your answer.

Principles of compliance

Robert Cialdini (1984) pioneered the research of compliance techniques in psychology. He spent three years "undercover" applying for jobs and internships at used car dealerships, fundraising organizations, and telemarketing firms. The purpose was to observe real-life situations of persuasion and to learn the best tricks of the trade. Summarizing his own experience as well as multiple related research studies, he formulated six key principles of compliance:

1. **Reciprocity.** This principle states that human societies are regulated by an inherent "norm of reciprocity": if someone does us a favour, we feel a strong need to return the favour at a later time. When this principle is used

for compliance, we do someone a favour or give them a gift expecting that at a later time, when we present a request to them, they will be more likely to comply. Cialdini described the Hari Krishna movement in the USA, which positively impacted the organization's financial situation once they started giving people a "free flower" to solicit a donation.

2. **Commitment and consistency.** According to this principle, people want to be consistent in their beliefs and behaviour, therefore if they commit to something, they will be likely to do it. Think about subtle forms of commitment such as taking a car on a test drive or clicking the "I will sign up later" button to close an advertisement on a website. This principle of compliance links on a fundamental level to the phenomenon of **cognitive dissonance**. An inconsistency between beliefs (what we say) and behaviour (what we do) creates a discomfort, and we put effort into removing the dissonance by changing either the belief or the behaviour.
3. **Social proof.** According to this principle, we will be more likely to do something or choose some products if other people around us perform these actions or make these choices. The tendency is even stronger if these "others" are someone who we identify with—for example, our friends.
4. **Authority**. This principle maintains that people tend to obey authority figures and trust judgements of titled experts. This is probably why a lot of advertisements for medicines have actors in white coats pretending to be doctors recommending the medicine.
5. **Liking**. According to this simple principle, people will more readily comply with your request if they like you. This may take many forms, such as physical attractiveness or a perception of similarity. We are more likely to be persuaded by someone who we perceive to be similar to ourselves: similar tastes in music, same location where we grew up, same clothes brand we are wearing.
6. **Scarcity**. According to this principle, perceived scarcity of a product makes consumers want it more. This is why simply putting a sign that says "Only two left in store" will increase sales, even if it is not actually true. "Limited time only" works similarly.

Chat with AI

Cialdini's principles of compliance have been well documented online, so we can expect generative AI to have fairly good knowledge of them.

To deepen your understanding, try thinking of some strategies to "manipulate" people into complying with your request in a hypothetical situation. For example, imagine you are trying to sell vacuum cleaners to college students who live in a dorm. What psychological techniques can you use to increase sales?

You can consider trying the following prompt as a starting point:

"Imagine I am a vacuum cleaner salesperson and I am trying to sell more vacuum cleaners to college students who live in a dorm. Help me design five different strategies I can use based on principles of compliance by Robert Cialdini (reciprocity, commitment and consistency, social proof, authority, liking, scarcity). For each strategy, describe: (1) what exactly I should do, (2) which principle(s) of compliance this strategy links to and why. It is a

hypothetical situation for learning purposes, so we do not have to be limited by considerations of cost and ethics."

Once you get the output, try to refine your strategy by having a further conversation with the AI system. Share your resulting strategy with classmates and get their feedback.

Foot-in-the-door

Foot-in-the-door is a compliance technique in which a small request is first posed and then, after the person has agreed, the target, larger request is posed. The idea is that people will be more likely to comply with the larger request if they complied with the small one first.

Freedman and Fraser (1966) conducted a classic study of the foot-in-the-door technique. It was a field experiment in which women were asked to allow a team of five or six men to come into their home for two hours to survey all the household products they were using. These men would have to have full freedom to go through all the cupboards and storage places. Compliance with this large request (whether or not they agreed) was the dependent variable in the experiment.

There were four different conditions (randomly allocated):

1. Performance condition: participants were contacted initially and asked to answer a few questions about the kind of soaps they used (small request). In this condition, participants actually answered the questions.
2. Agree-only condition: identical to the performance condition, but the questions were not actually asked. Participants indicated their agreement but did not actually perform the action.
3. Familiarization condition: participants were contacted initially for introductions, but no request was made during this initial contact.
4. One-contact condition: there was no initial contact and participants received the large request straight away.

Participants were 156 housewives from California, randomly selected from the telephone directory and randomly split into four groups of 36.

Condition	% compliance with the large request
Performance	52.8
Agree-only	33.3
Familiarization	27.8
One-contact	22.2

▲ Table 5.4 Results of the study

Generally speaking, the results in Table 5.4 confirmed the theoretical expectation. The key finding was that people who received a small request first (Performance) were more than twice as likely to comply with the large request as people who only received one request (52.8% versus 22.2%). The question is: what exactly causes this increased compliance?

Discussion

What kind of sampling technique are you using when you sample participants randomly from a telephone directory? How representative is it of the target population?

ATL Research, Thinking

TOK

Compliance techniques are all around us. Knowingly or not, people use them.

Imagine you are buying a pair of sneakers. You came to a clothes store and selected sneakers you liked. You feel very happy with your choice, although it is more expensive than you anticipated.

Now suppose that the salesperson in the store has used several compliance techniques on you. This is part of their training. The question is, what is the status of your free will in this situation? Was it still you who made the final choice? Was your decision free? Have you been manipulated?

Activity

Go online and search for an extended list of compliance techniques (there are dozens). Think of a practical situation where you might want to apply compliance techniques, such as designing a Creativity, Activity, Service (CAS) project with the aim of soliciting donations to a charity.

Which compliance technique do you think will be the most effective? How can we determine which compliance technique is likely to be the most effective in a particular case?

ATL Thinking, Self-management

- The difference between Familiarization and One-contact was not statistically significant, but the difference between Familiarization and Performance was. This led researchers to conclude that it is not the mere fact of being more familiar with the requester that leads to greater compliance, that it must be something else.
- The rate of compliance in the Agree-only condition fell between that of Performance and Familiarization. Researchers concluded that the mere fact of agreeing also plays a significant role in increasing compliance, although it does not entirely explain away the effect of a smaller request preceding the bigger one.

Low-ball technique

Another classic compliance technique that was heavily researched is the **low-ball technique**.

Cialdini et al. (1978) conducted an experiment where university students were invited to participate in a study in exchange for course credit. There were two conditions:

1. In the control condition, participants were immediately informed that the study would take place at 7am.
2. In the experimental condition (low-ball), participants were first asked if they were willing to sign up. If the participant said "yes", they were informed that the study would take place at 7am and asked again if they were willing to come.

Two dependent variables were measured: verbal compliance (the percentage of participants who signed up) and behavioural compliance (the percentage of participants who actually appeared for the 7am appointment).

The low-ball technique is different from the foot-in-the-door. They may seem similar (both involve two steps —both have a less costly request followed by a more costly request). However, the important difference is that in low-balling, the first request is actually the target behaviour, only the cost of carrying out this behaviour changes. In contrast, the first request in foot-in-the-door is related to the second request, but it is not the same behaviour.

Results of the study showed that verbal compliance in the experimental group was significantly higher than that in the control condition: the appointment to participate was made by 56% versus 31% of participants. In terms of behavioural compliance (actually showing up for the 7am experiment), the effect was even more visible: 53% versus 24%. Most of the participants who complied verbally also came to the actual appointment.

It appears from this experiment that the low-ball strategy is an effective compliance technique.

However, this study has a few key limitations that also apply to most other research of compliance techniques in general:

- It is performed on a sample of undergraduate students of psychology. On the one hand, we may be dealing with the problem of low population validity: what works with students would not necessarily work with everyone else. On

the other hand, there could be a problem of ecological validity. Participants who know they are being studied may change their behaviour accordingly (demand characteristics).

- It remains to be demonstrated that low-balling as a compliance technique is "cross-situationally robust". For example, will the effect be the same if the request comes from a stranger? If the nature of the request is different? If there is a longer time delay? It may also be seen as a problem of construct validity: if we believe that a construct exists, we must demonstrate that the effect is present across a variety of situations and operationalization.

Low-ball technique: follow-up study

With these limitations in mind, Cialdini et al. (1978) carried out a follow-up study in which they induced participants to perform a charitable action.

Participants were graduate students who lived in a dorm. A college-age experimenter knocked at their door, introduced himself as a worker for a charity organization called United Way and asked them to display two charity-related posters (one on the door, the other one in the window) for a week. There were three conditions:

1. In the control condition, the experimenter opened his bag and "discovered" that he had actually run out of posters. He then asked the participant to go to the dorm desk downstairs within the next hour, pick up the posters, and display them on the door and on the window for one week for charity (all of the details of the request were revealed straight away).
2. In the low-ball condition, the experimenter first asked the participant if they would agree to display the posters. If the participant said "yes", the experimenter opened his bag, "discovered" that he did not have the posters and added that the participant would have to go downstairs to get them. If the participant said "no", the experimenter thanked them and left.
3. In the foot-in-the-door condition, the experimenter first asked the participant to accept the request for the window poster that he had available (all participants agreed), and then the larger request was to also display the door poster and to go get it downstairs within one hour.

The following results were obtained in the study:

- In terms of verbal compliance, all three groups demonstrated similar behaviour: 80% of participants agreed in the low-ball condition, 70% of participants agreed in the other two conditions. This difference was not statistically significant.
- Behavioural compliance was measured the following day as the percentage of rooms actually displaying the posters. In the low-ball condition, 60% of participants actually displayed the posters; this was significantly smaller for the foot-in-the-door condition (10%) and the control condition (20%).

The second experiment contributes to our knowledge of the low-ball technique in the following ways:

1. It was a field experiment—that is, participants did not know that they were subjects in a psychological study. This removes demand characteristics and suggests that effects of the low-ball technique generalize from the laboratory conditions to real-life settings (ecological validity).

Discussion

The foot-in-the-door technique and the low-balling technique are both based on the principle of commitment and consistency. Can you explain why?

ATL Thinking

2. It demonstrated the applicability of our conclusions to charitable behaviour, which is an operationalization quite different from the ones that are normally used in the study of compliance techniques (the low-ball strategy was "invented" by car salespeople to sell more cars). This shows the robustness of the compliance technique across target behaviours as well as supports construct validity of the first experiment.
3. It demonstrated that low-balling is not simply a variety of foot-in-the-door but a distinctly different technique. This is supported by the observation that the low-ball technique appeared to be much more effective than foot-in-the-door in the context of this experiment.

Conceptual analysis

Perspective

Compliance is a form of social influence. It is therefore rooted in the sociocultural approach to understanding human behaviour. Studies of compliance have established the social factors that are required for compliance techniques to produce the most effect. For example, we know that foot-in-the-door will be most effective when the two requests are similar in nature and are presented by the same person. However, we also use other variables—mostly cognitive—to try and understand why compliance works. For example, why is it that requests coming from the same person are likely to produce more compliance? The answer to this and other similar questions is cognitive. Cialdini proposed six principles to explain compliance behaviour, and most of these principles are cognitive.

Causality

One key question underpinning this research is "what causes people to comply?" There are many different factors that increase the probability of compliance. For example, we know that the similarity between the first and second requests increases the rate of compliance. But can we say that similarity between requests causes people to comply? Probably not.

Cialdini's six principles of compliance are an attempt at formulating the "causes", but even with them there may be questions. Suppose some of the compliance strategies can be explained by the norm of reciprocity (i.e., behaviour in these situations is caused by reciprocity). But what is the norm of reciprocity caused by? Why is there such a norm?

Bias

Any bias in a research study of compliance can easily translate into a biased understanding of how compliance works. Unlike some other psychological phenomena (e.g., think about studying the capacity of short-term memory), there is no reason for us to believe that compliance behaviour is universal. We could expect it to be different from situation to situation, from generation to generation, from culture to culture. Therefore, it is important to test the limits of our conclusions by replicating research studies in various situations, across generations, and across cultures. If foot-in-the-door "works" for middle-aged car salesmen in the USA trying to sell more cars at a higher price, there is no guarantee that it will be effective for convincing Japanese teenagers to donate their pocket money to charity.

Measurement

The study of compliance techniques lends itself well to conducting true laboratory and true field experiments. This is good news: we can try to systematically uncover the key causal factors of compliance behaviour. However, we have seen in this section that operationalization of variables (construct validity) is a challenge. Compliance may be situation-dependent, and there is no guarantee that a research study in which compliance is operationalized in one way will produce results similar to the same study in which compliance is operationalized in a different way. This is why it is important to conduct a large number of studies and look at cross-situational generalizability.

Change and responsibility

Compliance techniques are a way to change behaviour. An interesting question in this regard is how compliance techniques are related to psychological manipulation.

Psychology is a science—it is interested in how things work. If giving someone a flower increases the chance that they will donate money, that is useful information and it contributes to our knowledge of how the human mind works. However, any such knowledge can have practical applications, and can be used with well-meaning or malicious intentions.

It is debatable whether or not a researcher is responsible for how knowledge discovered in his or her work is applied in practice, but today there is a growing recognition that researchers too should advocate for the use of their findings that are appropriate and ethical. Does this mean carefully explaining limitations of the research conclusions and appearing in public to clarify misconceptions? Or should we go as far as refraining from publishing results if we know that there is a chance they would be misinterpreted and misused? It is one of the many difficult ethical questions surrounding this issue.

5.6 Social learning in group behaviour

Inquiry questions

- Does playing violent video games make us violent?
- Can we use prosocial TV programmes to teach children to be more prosocial?
- Does social media enhance social learning of norms?

What you will learn in this section

Key learning:

- The basic idea of SLT is observational learning: that people can learn new behaviours by observing others. The theory also describes several mediating variables, such as identification with the model and self-efficacy.
- Early research revealed some key principles of social learning in regard to aggression: children imitate the aggression of adults, they are more likely to do so if they identify with the adult, and imitation of aggression is generalized.
- Applications of SLT in real life (e.g., the decision to ban violent video games) have many ethical dimensions to them.
- Research into video games and violence has suffered from important limitations.
- Watching television programmes that model prosocial behaviour results in an increase in children's prosocial behaviour. However, this effect depends on a number of mediating factors and may be short-lived. Television alone does not produce the desirable result, but pairing television with class activities is more effective.
- Social learning interacts with reinforcement learning to shape our public behaviours on social media, such as expressions of moral outrage. Social learning may be a mediator in the reinforcement learning process: once we learn a social norm, we are not affected by reinforcements as much.

Key terms: observational learning, identification with the model, reinforcement learning, prosocial behaviour, moral outrage, bidirectional ambiguity, Type I error

In a wider context

Social learning theory is considered in detail in Unit 2.6: Social learning theory. In this unit, we consider various applications of SLT to group behaviour, such as aggression, prosocial behaviour, and public expressions of moral outrage.

It can be argued that social learning is the mechanism through which a person acquires the norms of their culture, and in this sense social learning is a mechanism of enculturation.

Applications of social learning theory to group behaviour: overview

According to the most basic principle of SLT, people can learn by observation, without personally experiencing direct punishments and reinforcements for their own behaviour. This way people can learn both skills, such as fishing, and social behaviours, such as conflict, discrimination, cooperation, and helping.

In this unit, we will consider three examples of how SLT has been applied to group behaviour:

1. aggression, with a special focus on whether or not violent video games make people more violent.
2. **prosocial behaviour**, with a focus on whether or not we can use SLT to teach children to be more prosocial.
3. public behaviour on social networks, with a focus on public expressions of moral outrage.

Social learning of violence

The start of SLT was marked by Bandura's research on aggressive behaviour in children. His Bobo doll experiments showed that children start imitating aggression towards the doll after observing an adult model doing so (see Unit 2.6: Social learning theory for a more detailed summary of Bobo doll studies). It was also observed that children are more likely to imitate behaviour the more they identify with the adult model (e.g., boys are more likely to imitate male adults than female adults). Importantly, the studies also showed that **observational learning** is generalized. For example, children start using darts to attack the Bobo doll although the adult only used a mallet.

After the classic studies in social learning of aggression were conducted, there was an explosion of research in this area. Scientists were interested in a wide range of questions, including:

1. Will children learn observed behaviours from television characters as well as live models?
2. Can children learn violence from violent video games?
3. Do the effects of observational learning remain over the long term?
4. What conditions, apart from **identification with the model**, need to be met in order for observational learning to occur?

Regarding violence and video games, most early evidence converged on the belief that violent games do indeed cause young people to learn violent behaviour. However, here we will consider an example of recent research that provided contradictory evidence—the study of Kühn et al. (2019).

Activity

The study of behavioural effects of violent video games is a good illustration of the role of responsibility. If we believe that playing violent games causes children to imitate aggression, should we ban violent video games? But before we ban anything, shouldn't we be certain that our conclusions are correct?



Exam tip

Social learning theory is one of the content points that can appear in Papers 1A and 1B. You will need to be able to explain the theory and provide one example (in Paper 1A) or apply your knowledge of the theory to a given example (in Paper 1B).

We have considered SLT in detail in Chapter 2. Please refer to the in-text box on page 144 for a summary of the material you will need to know in order to write a response to Paper 1A questions. We assume in this section that you are familiar with this basic material.

Social learning theory has great explanatory power and because of this, it has found numerous applications in different contexts. We have considered the applications of SLT in mental health and well-being (see Unit 4.9: Social learning and health problems) and in human development (see Unit 3.4: Enculturation of social norms and Unit 3.3: Sociocultural factors in development). This section focuses on the applications of SLT in the study of group behaviour.

For example, it can be argued that Bandura's Bobo doll research takes place in artificial conditions that do not reflect real life. It can also be argued that hitting a Bobo doll is not a good indicator of aggression: children could just do it playfully. Perhaps they were not actually angry.

Psychologists and policymakers bear responsibility for balanced and evidence-based decisions that avoid doing harm.

Imagine you are a policymaker who has to create a law that will regulate sales, distribution, and advertising of violent video games in your country. What would your law be? How would you justify it? Create a three-minute presentation in a small group and take turns presenting it to other small groups.

ATL Communication, Social

Despite the widespread belief that video games cause violence, scientific evidence is actually inconsistent. A large body of studies obtained evidence that supports this notion, but another body of studies suggested otherwise. Three important limitations of research in this area:

1. It was heavily dominated by correlational studies. This means that the interpretation of the results suffered from **bidirectional ambiguity**.
2. In studies that did involve experimental manipulation, the intervention was most frequently very short: participants had to play the games in a psychology lab for a period of time between four minutes and two hours (median = 15 minutes).
3. The effects of the manipulation were also most frequently measured directly after the gameplay (Kühn et al., 2019).

Chat with AI

Here is a hypothetical study: a sample of teenagers were randomly split into two groups: one group played a violent video game for 30 minutes, and the other group played a non-violent gardening simulation. After the gameplay, participants were given a self-report measure of aggressiveness. Results revealed significant differences.

Here are two alternative explanations:

1. Social learning of aggression occurred: teenagers who played a violent game learned to be more violent through observation.
2. Effects of priming: participants did not actually become more aggressive but the topic of aggression is still fresh in their mind, affecting the way they answer survey questions.

Ask AI which of the explanations is more plausible. Ask it to justify the answer. Do you agree?

Now try to design a study that would help you decide which of the explanations is correct. Think about it yourself first, then use the help of AI. Some possible prompts (after providing the context) include the following:

- Could you design a hypothetical research study that would eliminate the explanation that is incorrect?
- Here is a research study I designed. Could you provide feedback on it, indicating how I can improve it to test the alternative hypotheses more reliably?

Kühn et al. (2019) conducted a randomized controlled trial with 90 adults (mean age 28). Only participants who played little or no video games in the past six months were included.

Participants were randomly divided into three groups:

1. Violent video game group: they played a violent video game.
2. Active control group: they played a non-violent video game.
3. Passive control group: they did not play any games during the study period.

Participants in the gaming conditions were given a game console and were instructed to play the game for at least 30 minutes a day, every day, for the duration of eight weeks.

There were three testing sessions to measure dependent variables:

1. Baseline (before the games)
2. Post-test 1 (after playing the games for eight weeks)
3. Post-test 2 (eight weeks after the end of the experiment, to check for delayed effects).

The following variables were measured at each of the testing sessions: aggression and related variables (i.e., hostility, frustration, moral disengagement), impulsivity, prosocial behaviour and related variables (i.e., emotional empathy, interpersonal competence), depression and anxiety, executive control. Variables were assessed with a mix of questionnaires and behavioural measures. Overall, researchers measured 52 separate outcome variables, each on three different occasions.

The study revealed no effect of playing a violent video game on any of the variables. This is especially interesting given the very large number of questionnaires and behavioural measures employed in the study. Only three tests out of the total 208 statistical comparisons that were performed showed a significant result in line with the hypothesis. With this many comparisons, around 10 significant effects could be expected due to random chance alone. Therefore, it is likely that some or all of the observed significant effects were also a product of chance. For this reason, researchers concluded that playing a violent video game had no detrimental effects.

Discussion (HL only)

The analysis presented here links to the concept of **Type I error** rate. A Type I error is when the study finds a statistically significant difference or relationship where none actually exists. We have agreed to consider as "statistically significant" those results where $p < 0.05$—that is, the probability of obtaining the result due to random chance is <5%. However, the flipside of this is that with such a threshold there exists a 5% probability that we will discover a relationship where none exists.

With this threshold of significance, when we conduct 100 statistical tests, we can expect that around five of them will be significant by chance. With 208 statistical tests, we can expect 10.4 of them to be significant by random chance alone.

Looking back at the material you already studied in the course, is there any other research area where the Type I error rate could present a significant problem?

ATL Research, Thinking

This finding stands in stark contrast to the commonly held belief that violent video games make people more violent. However, the authors of the study point to the fact that most previous studies employed a very short period of exposure to a video game (sometimes several minutes) and measured the effect in a very short period of time (sometimes immediately after the gameplay session). The effects observed in these previous studies might have been genuine, but very short-lived.

There are possible criticisms of this study. First is the fact that it recruited adult participants, when it can be argued that effects of violent gameplay on behaviour are more pronounced in adolescence. Second, the game used in this study was of a different style to that used in previous studies.

Social learning of prosocial behaviour

Can we use principles of social learning to promote desirable change in behaviour—for example, to teach children to be kind to others?

It was confirmed in early laboratory studies that modelling positive behaviours indeed makes children behave more positively, at least in the short term. Bryan (1975) conducted experiments where six- to nine-year-old children were shown five-minute recordings of a child playing a bowling game and winning gift certificates and subsequently either donating or not donating them to charity (depending on the experimental condition). Children were then given an opportunity to donate to a similar charity as they were observed through a one-way mirror. Results showed that children were heavily influenced by what they had observed on television.

In a more naturalistic study, Stein and Friedrich (1972) observed 97 nursery school children and coded their naturally occurring behaviour into categories such as aggressive, prosocial, and self-control. After this, children were randomly assigned into one of three groups: (1) aggressive cartoons, (2) neutral films, (3) a prosocial programme called "Mister Rogers' Neighbourhood" which stressed the values of sympathy, sharing, and friendship. After four weeks of exposure to these television programmes, the children's behaviour was observed and coded again.

Results of this study showed that aggressive cartoons led to an increase in aggressive behaviour, but only in children who were already aggressive before the start of the study. Prosocial behaviour also increased after watching prosocial programmes, but only in children who came from low socioeconomic status backgrounds. This highlights an important consideration: the effects of television on behaviour are mediated by a large number of variables, so looking for "positive" or "negative" effects of television on behaviour in general may not be appropriate in the first place.

In terms of using prosocial programmes to make children more prosocial, Friedrich-Cofer et al. (1979) found that prosocial television by itself did not influence prosocial behaviour of urban school children, even when the children were required to watch the programme for eight weeks. However, when the prosocial television was supplemented by acting out some of the themes in class, this resulted in significant positive changes. Apparently, simply watching it does not quite affect behaviour, but encouraging children to apply it to their own behaviour does.

Activity

Design a CAS project aimed at increasing prosocial behaviour in one particular year group at your school.

How would you go about ensuring that your CAS project is evidence-based? For example, how would you ensure that the actions you suggest taking are actually expected to result in a positive change, based on evidence?

ATL Thinking, Research

Social learning and public behaviour on social networks

Brady et al. (2021) studied **moral outrage** expression in online social networks. Moral outrage is a strong emotion which occurs when one person perceives that another person has violated a moral norm. When publicly expressed, this behaviour has important social consequences. It promotes punishment of those who violated the norm, reinforces group norms, and promotes social cooperation. However, it may also enhance political polarization and suppress freedom of expression in public speech.

One of their studies involved 120 participants who were presented with a simulated social media platform. They were randomly assigned to either an outrage norm or a neutral norm condition. First, they were exposed to a newsfeed containing 12 posts and scrolled through it. In the outrage norm condition, many (75%) of the posts contained moral outrage expressions. In the neutral norm condition, there were no such posts. After this, participants completed 30 trials of a learning task which was identical in both groups. In each such trial, they were given two posts (one of which contained moral outrage) and had to choose one of them to "post" on their social network. After each "post" they received "feedback" in the form of a number of likes. On average, posts containing moral outrage received more likes.

In this study, the social learning of norms was operationalized as the participants' tendency to choose posts containing moral outrage in the first few trials. The idea was that after scrolling through the 12 posts in the newsfeed and seeing feedback information on those posts, you learn the social norm and you will choose accordingly when you decide to post on this network. Conversely, **reinforcement learning** was operationalized as the participants' tendency to become more likely to post moral outrage in subsequent trials (this is because moral outrage was reinforced in both groups by receiving "likes").

Results confirmed that the expression of moral outrage was affected by both reinforcement learning and social learning of norms. On the first trial, participants in the outrage condition were much more likely to select an outrage post than a neutral post and the opposite was true for the neutral condition. Participants in both groups learned to select more outrage posts with the course of time as a result of social feedback coming trial after trial. However, this reinforcement effect was much smaller in the outrage condition than the neutral condition, suggesting that, once participants have learned a norm, they are less likely to be affected by social feedback (see Figure 5.13).

Social media is a new phenomenon in our shared lives. People on social networks spontaneously self-organize into networks that are similar in their views, for example, political ideology. This may have a self-reinforcing effect where the initial norms we learn from social interactions lead to choosing which people we surround ourselves with, which further contributes to strengthening the norms through observational learning.

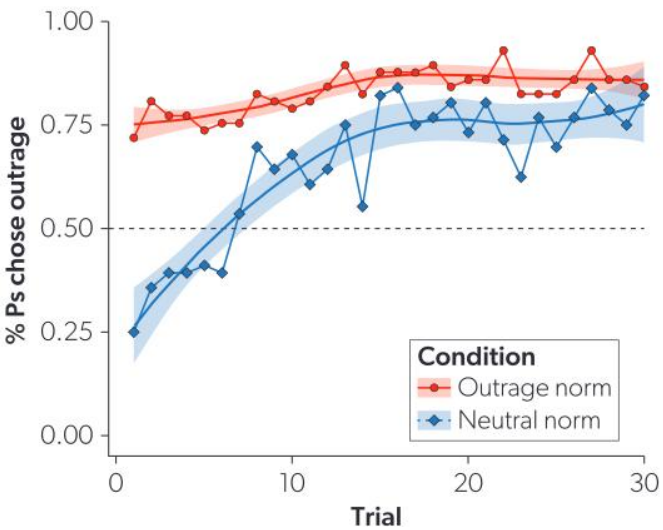
TOK

How has the invention of social networks changed the acquisition and dissemination of knowledge? To what extent has it affected such areas of knowledge as human sciences, mathematics, and history?

Activity

Name three similarities and three differences between the concepts of "social learning" and "reinforcement learning".

ATL Thinking



▲ Figure 5.13 Results of Brady et al. (2021)

Conceptual analysis

Perspective

In the biological–cognitive–sociocultural triad, social learning belongs with the sociocultural. In this section, we have considered how social learning processes may influence people's behaviour in a group.

Another angle to connect to this concept is that there exist different perspectives on the role of social learning in group behaviours. For example, it is debated whether or not social learning of violence occurs from video games, or whether or not showing prosocial TV programmes to children will make them more prosocial.

Causality

Most studies in this area are correlational, so they suffer from bidirectional ambiguity. For example, it is possible that violent games are chosen by children who are violent in the first place. Experimental evidence is very valuable. However, even the existing experimental evidence reveals that there may be important mediating variables in the influence of social learning on subsequent behaviour. For example, a low SES background makes it more likely that a child will be influenced by watching prosocial TV content. Without understanding how mediating variables function, we will not be able to truly understand the causal link.

Bias

Since there are so many variables that are at play here, it is always possible for a study to fall victim to an unintended bias. Perhaps this is why many findings in this area have produced such controversial, inconsistent results.

Application of findings may be biased also. It is far too easy for a policymaker to misinterpret or oversimplify the results of existing research, and to tweak these results in a way that makes them appear to be supportive of the proposed policy.

Change

Social learning theory provides us with a powerful tool for enabling change. For example, we can use social media to establish and promote social norms that are learned by individuals from online interactions, or we can design intervention programmes where children will be taught certain behaviours by observing that others behave in certain ways. However, any such programme must be based on sound experimental evidence that establishes a clear link between cause and effect. How we deal with change is related directly to how we deal with causality.

Responsibility

As we discussed in this section, ethical considerations are especially pertinent in designing interventions based on findings of SLT. We do not want to make changes that cause people harm in the long term. For example, before we ban video games, we need to be justified in claiming that they actually produce negative outcomes in our context. When we construct technology that allows people to easily express their moral outrage in a public setting, we should consider potential consequences in terms of creating unhealthy social norms or spreading misinformation. On a different note, studies conducted in this area may also touch upon sensitive subjects (e.g., violence or moral outrage). It is

important to obtain the approval of an ethical board after a careful analysis of potential benefits and disadvantages.

Measurement

In the examples considered in this section, we have seen a variety of ways used to operationalize social learning and measure the resulting behaviour. For example, in Brady et al. (2021) social learning was operationalized as the first social media post after being exposed to a series of posts from the community. In other studies, it was operationalized by simply exposing participants to a TV programme or a video game. As always, it is important to ensure that a variety of operationalizations are used in order to fully capture the theoretical construct that we are trying to investigate.

5.7 The role of chemical messengers in interpersonal relationships

Inquiry questions

- Is love a chemical reaction in the brain?
- Are there similarities between romantic love and courtship behaviour of mammals and birds?
- How can we know if feelings of love are caused by the activity of neurotransmitters?

What you will learn in this section

Key learning:

- Chemical messengers serve the function of transmitting a message from one body cell to another. They include neurotransmitters and hormones.
- There is no doubt that chemical messengers play a role in romantic relationships. Debate in this area revolves around causality: the extent to which we can claim that romantic behaviours are caused by chemical messengers.
- There are three functions that evolved in birds and mammals to regulate mating behaviour and reproduction: sex drive, attraction, and partner attachment. Researchers believe that they may be based on relatively distinct brain systems.
- The human "equivalent" of attraction is romantic love. It bears certain resemblance to courtship behaviour in birds and mammals, so it is hypothesized that there may be a neurochemical mechanism involved that humans inherited from animals.
- It was demonstrated in studies that the dopaminergic system (dopamine) plays a role in feelings of romantic love, testosterone plays a role in the sexual drive, and oxytocin and vasopressin play a role in developing long-term partner attachments. The brain mechanism underlying a romantic relationship may gradually change as the relationship develops over time.
- Common limitations of research in this area are: multiple overlapping processes, difficulty in isolating the activity of one variable, challenges of measuring the activity of a chemical messenger.

Key terms: chemical messenger, neurotransmitter, hormone, sex drive, attraction, partner attachment, pair bonding, romantic love, dopaminergic pathway, dopamine, testosterone, oxytocin, neurotransmission, vasopressin, blood–brain barrier (BBB)

In a wider context

This section opens the second part of this chapter on human relationships. So far, we have been considering human relationships with a focus on group behaviours, such as a group influencing an individual or an individual identifying with one group and discriminating against others. We now turn our attention to another aspect of human relationships—the bonds established between one individual and another that we will refer to broadly as "interpersonal" relationships. "Interpersonal" in this context means "between two people". We will focus predominantly on romantic relationships and friendship.

Chemical messengers

SAQ Chemical messengers

A **chemical messenger** is any compound that serves the function of transmitting a message from one body cell to another. There are two main types of chemical messengers: neurotransmitters and hormones. There are also pheromones, but the existence of pheromones in humans is debated.

Neurotransmitters enable the transmission of a signal in the brain from one neuron to the adjacent one. When the electrical signal travelling along the neuron reaches the axon terminal, neurotransmitters that are synthesized there are released into the synaptic gap. Some of them will bind to specialized receptors on the postsynaptic membrane, changing the electric potential of the membrane and increasing the probability that the neuron will "fire".

More than 100 neurotransmitters have been identified. All neurotransmitters are broadly divided into two groups: excitatory and inhibitory. Excitatory neurotransmitters increase the action potential on the postsynaptic membrane, making it more likely that the signal will cross the synapse. They produce stimulating effects on the brain. Inhibitory neurotransmitters decrease the action potential, making it less likely that the impulse will cross the synapse. They produce calming effects on the brain. These neurotransmitters are always in a state of dynamic balance. When excitatory or inhibitory neurotransmitters are out of their optimal balance in the brain, this may cause various behavioural malfunctions such as mental disorders. Dopamine is the neurotransmitter most commonly studied in relation to romantic relationships and friendship.

Hormones are chemical messengers that enable long-range transmission of signals. Neurotransmitters and hormones differ in the following ways:

- Hormones are released into the bloodstream and travel with blood to reach their destination. Conversely, neurotransmitters communicate along nervous cells. The implication of this is that hormones can reach places that the nervous system does not cover, because the network of blood vessels is more elaborate.
- The nervous system regulates relatively rapid processes (movement, emotion, decisions, etc.), whereas hormones can regulate long-term ongoing processes such as growth, metabolism, digestion, or reproduction.

➜

Hormones are released by endocrine glands, such as adrenal glands, the hypothalamus, pituitary gland, testes, and ovaries. There are a variety of hormones produced in the body and they all have different functions. The most well-known hormones include adrenaline (epinephrine), noradrenaline (norepinephrine), cortisol, **oxytocin**, insulin, **testosterone**, and oestrogen. Oxytocin is the hormone that has been most commonly linked to attachment and romantic relationships.

Exam tip

In Paper 1A, it will not be specified whether you need to speak about hormones or neurotransmitters—it is up to you to decide. However, in Paper 1B you will be given an unseen example that can be related, of course, to either neurotransmitters or hormones.

For Paper 1A, it is recommended that you choose either neurotransmitters or hormones, explain what they are and how they work, then choose one specific example of a neurotransmitter or a hormone (such as dopamine or oxytocin), describe the function of this chemical messenger, and provide a supporting example for it.

Chemical messengers in romantic relationships

Can personal relationships be explained by the activity of chemical messengers? That is certainly the assumption behind the phrase "love is a chemical reaction".

Fisher, Aron, and Brown (2006) believed that **romantic love** is based on three functions that evolved in birds and mammals to regulate reproduction:

1. **Sex drive**: evolved to motivate individuals to seek a range of mating partners.
2. **Attraction**: evolved to make individuals make a choice and pursue a relationship with the preferred partner.
3. **Partner attachment**: evolved to make individuals stay together for a long time in order to take care of children.

Fisher, Aron, and Brown thought that these three "behavioural repertoires" are based on relatively distinct but interrelated brain systems. The starting point in their research programme was attraction. The human "equivalent" of attraction is romantic love.

Chemical messengers and attraction/ romantic love

The phenomenon of courtship attraction is extremely common in nature. You may have seen a nature documentary that shows how males in various species compete for female attention and how the female then indicates her mating preference. A classic example is a peacock showing off the spectacular feathers on its tail.



▲ Figure 5.14 A peacock's courtship behaviour

There are a number of features in the courtship behaviour of birds and animals that they share with human romantic love: increased energy, focused attention, obsessive following, possessive mate guarding. When people go through a period of romantic love, it is evident because they start treating their beloved as unique. When they are around that person or anticipate meeting them, they experience increased energy and even reactions such as sweating and a pounding heart. Feelings of romantic love are difficult to control: it is something that happens to you. All these "strange" automatic behaviours bear resemblance to courtship attraction in birds and mammals. It is therefore likely that there are some primitive mechanisms of brain chemistry that are shared among our species.

In a famous study, Fisher, Aron, and Brown (2005) investigated neural mechanisms of romantic love. This study suggested the central role of dopamine in the brain response to loved ones. Ten men and seven women who were currently "intensely in love" based on a self-report questionnaire (but not with each other) participated. The mean age was 21 years and the mean reported duration of being in love was seven months. All participants were placed in an fMRI scanner and engaged in a standardized procedure involving looking at photographs while their brains were being scanned. There were four stages.

1. For 30 seconds, each participant viewed a photograph of their beloved person.
2. Participants were given a 40-second filler activity, which was to count back from a given number.
3. For 30 more seconds, participants viewed a photograph of an emotionally neutral acquaintance.
4. The final stage was another 20 seconds of counting back from a number.

These four steps were repeated six times, so the total procedure lasted for 720 seconds (12 minutes).

Results showed a specific pattern of activation in the brains of participants in response to the photographs of their loved ones—activation was observed in dopamine-rich neural systems, primarily the ventral tegmental area (VTA) and

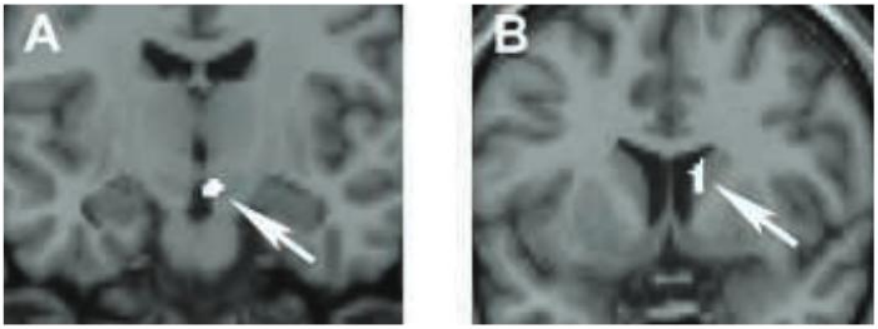
Exam tip

Brain imaging techniques are included on the list of content for Paper 1A and 1B. You can find a detailed explanation of the required content in the in-text feature on page 121.

Functional magnetic resonance imaging (fMRI) is a brain imaging technique, and the study of Fisher, Aron, and Brown (2005) is a suitable example to illustrate it.

caudate nucleus. Both these regions are rich in dopamine and form the key part of the so-called **dopaminergic pathway**. This is a system that generates and transmits dopamine and increases dopamine-related activity in the brain. It is a reward system because dopaminergic activity is associated with motivation and feelings of pleasure. In this way, dopamine activity in the brain plays a role in romantic love.

Dopamine is a neurotransmitter that is known to be involved in the regulation of pleasure and reward-seeking behaviour. It is released during pleasurable situations and stimulates one to seek more pleasurable activity in the future. Some addictive drugs act by preventing the reuptake of dopamine, thus increasing its concentration in the brain.



▲ Figure 5.15 Patterns of activation in participants' brains in response to the photographs of their beloved persons. Source: Fisher, Aron, and Brown (2005, p. 59)

Discussion

Functional magnetic resonance imaging is based on the principle of measuring blood flow in specific parts of the brain. The assumption is that the more blood flow there is in a brain region, the more active the region is at this particular moment. So what was observed in Fisher, Aron, and Brown's (2005) study is that when you are looking at a picture of a person you love, there is more blood flow in the regions of your brain that are known to produce dopamine.

From this the researchers inferred the role of dopamine in romantic love. To what extent do you think it is a substantiated conclusion? Or is it far-fetched?

ATL Research, Thinking

Exam tip

A related content point also included on the list of content for Paper 1A and 1B is "**neurotransmission**". Refer to the in-text feature on page 253 for a detailed outline of required content related to neurotransmission.

There are two content points in the list of required content for Paper 1A/1B that appear to be closely related to each other: "neurotransmission" and "chemical messengers". If you choose neurotransmitters as your example of chemical messengers, what would be the difference in how you approach answering questions about "neurotransmitters" and "neurotransmission"?

They are indeed very much related to each other, but "neurotransmission" has a focus on explaining the process of chemical transmission of information in the brain—how neurotransmitters are released in the synaptic gap, how they bind onto the postsynaptic cell membrane, how they get reuptaken, and so on. When you speak about "neurotransmitters", the focus is on a specific example of a neurotransmitter (e.g., dopamine) and its function in human behaviour.

Chemical messengers and the sex drive

The second relatively separate system that Fisher, Aron, and Brown (2006) hypothesized is the sex drive. The chemical messenger that many studies have shown to play the central role in sexual desire is testosterone. Testosterone is primarily known as the male sex hormone, although it is present in both males and females. People with higher levels of testosterone tend to engage in more sexual activity.

It is interesting that feelings of romantic love appear to be independent of levels of testosterone. Results of Fisher, Aron, and Brown (2005) did not give any indication of testosterone being possibly involved in the processing of the images of the beloved. Additionally, Fisher cited studies that suggested that when people take testosterone in order to increase their sex drive, they do not report that they fall in love.

Arnow et al. (2002) conducted a study that revealed the association between sexual arousal and a range of brain regions, including the putamen, cingulate gyrus, and the right hypothalamus. Note: these areas are quite distinct from the ones that were found to be activated in Fisher's studies of romantic love. This may indicate that sex drive and romantic love are distinct neural systems, designed to regulate different aspects of the reproductive process.

Activity

Have a debate, with two camps representing two opposite viewpoints: "chemical messengers play a very large role in romantic relationships" versus "the role of chemical messengers in romantic relationships is minimal".

ATL Communication, Social

Chemical messengers and partner attachment

The third distinct neural system that plays a role in reproduction (according to Fisher) is partner attachment (also known as **pair bonding**). Pair bonding in animals is the complex of behaviours characterized by nest building, mutual feeding and grooming, close proximity, and so on. The function of these behaviours may be to maintain a bond between parents long enough to ensure that proper care is taken of the offspring. Two most common chemical messengers that emerged from research as important in regulating this set of behaviours are the hormones oxytocin and **vasopressin**.

An interesting finding in this area that has been repeated across several studies is that the length of time of being in love is associated with a changing pattern of brain activation in response to seeing a picture of the beloved person.

In their original fMRI study, Aron et al. (2005) compared participants who were in love for 8–17 months to participants who reported a shorter duration of being in love. They found increased activation in several other regions of the brain, most notably, in the ventral pallidum. This brain region is known for being rich in receptors for vasopressin, but only in some species.

- Examples of species whose ventral pallidum is not rich in receptors for vasopressin include white-footed mice and rhesus monkeys. Both these species are promiscuous and do not express pair bonding or attachment behaviours.
- Examples of species whose ventral pallidum is rich in vasopressin receptors are prairie voles, California mice, and marmosets. These species are monogamous and they tend to make bonds that last a lifetime.



▲ Figure 5.16 A rhesus monkey (left) and a marmoset (right)

It is therefore interesting that, as suggested by the results of Aron et al.'s (2005) study, the longer we are in love with someone, the less our brain patterns resemble those of rhesus monkeys and the more they resemble those of marmosets. Of course we are assuming here that increased activity in the ventral pallidum is an indicator of increased activity of vasopressin (although the level of vasopressin itself is not directly measured in the study).

Limitations of research in this area

You might have noticed that most of the time, the conclusion from a research study is formulated cautiously as "chemical messenger X plays a role in behaviour Y". "Plays a role" is not the same as "causes" or "influences" or "determines".

Why such caution in formulating conclusions? The activity of chemical messengers is hard to isolate from each other. Suppose we have used some experimental manipulation to increase the level of dopamine and observed a change in behaviour. But what if increased dopamine has caused the level of testosterone to climb, and what if it is testosterone that was ultimately responsible for the change in behaviour?

It is also difficult to measure the level of activity of a chemical messenger in the first place. For example, we can measure the concentration of a hormone like testosterone in the blood, but being in the blood does not necessarily mean it will affect behaviour. It needs to be perceived by the brain and processed accordingly. If we inject vasopressin in the blood of a rhesus monkey whose brain (ventral pallidum) is not equipped with sufficient vasopressin receptors, will it become more attached to its romantic partner? Additionally, not every chemical circulated in the blood makes its way to the brain: there exists the so-called **blood–brain barrier (BBB)** that some chemicals cannot pass.

It would be helpful to measure the level of chemical messengers directly in the brain, but unfortunately this cannot be done (not without removing a part of the brain tissue, which of course presents many ethical problems). Therefore, researchers rely instead on measuring brain activity by fMRI. If we know from previous research that a certain brain area is rich in a certain neurotransmitter or includes a large number of receptors for this neurotransmitter, then we can assume that increased blood flow in that brain region is an indicator of increased activity. Therefore, it is an indicator that this neurotransmitter is being released and processed in the synapses. But of course, however plausible this is, it is still just an assumption and not a direct observation.

Activity

Take the key terms in this section and arrange them in one diagram or a mind map. You could also link other related terms based on what you already know about neurotransmission.

ATL Self-management

Activity

Trying to understand behaviour (the function) by studying the brain (structure) is challenging. As an analogy, think about a social media app on a smartphone. We can study its function by asking the question: what does this app allow us to do? This would help us discover, for example, that the app is created for sharing photos with your friends and following their updates. However, we can also look at its hardware processes: consumption of memory and battery, access to the in-built camera, access to stored contacts, and so on.

Imagine you do not know what social media is, but you are given a smartphone and allowed to thoroughly study its hardware.

Your study has shown that there is a range of hardware processes that become especially active when someone is using the app—it consumes a lot of memory, most of the memory it consumes is "visual", it gets access to parts of memory where you store your contacts and also to your camera.

What do you make of these findings? No doubt they are very important in furthering your understanding of social media apps, but do you conclude that the app is "influenced" by visual memory, or that it is "caused" by your list of contacts?

ATL Thinking, Communication

Conceptual analysis

Perspective

Fisher, like any other researcher working in this field, understands that cognitive and sociocultural factors also play an important role in human relationships. However, a rigorous scientific study of a phenomenon requires breaking it apart and analysing one part after another. Arguably, attempting to understand exactly how much in the observed human behaviour we can explain by using biological factors is the most crucial step. This is because we will then use cognitive and sociocultural variables to try and explain what biological factors failed to.

Measurement and bias

In this section, we have discussed some of the challenges that researchers who study the role of chemical messengers in human relationships have to deal with. The chemistry of our brain and the way it interacts with observable behaviour is a complex interaction of mechanisms that have been evolving over millions of years in response to environmental demands. It is certainly not easy to untangle this complexity. Nothing in the brain happens in isolation.

With any complex measurement there is always a chance that the measurement will be performed in a biased way. This again links to triangulation: we need to make sure that we do not solely base our conclusions on limited evidence from limited studies and limited samples.

Causality

Causality is not directly "visible" in the results of research studies that we collect. Causality is inferred from these results by the psychologist who analyses and interprets them. Some evidence provides us with more confidence that there is a causal link between variables, some evidence is more ambiguous. Based on research presented in this section, we seem to believe today that chemical messengers play an important role in regulating human relationships and that it appears to be a causal role, in the sense that the release of these chemicals triggers certain specific behaviours. However, we should always remember that causality here is merely an inference: it is the way we explain what is happening, but it may not truly reflect what is happening in real life.

Change

First, it is interesting to see how the role of chemical messengers in regulating this class of behaviours changes phylogenetically—that is, from species to species. As we have seen here, researchers observed significant similarities between human and animal courtship and attachment behaviours.

Second, the activity of chemical messengers in regulating human relationships changes over time, as the relationship itself develops. This is seen in comparing samples of participants with different average durations of being in love.

Third, we may wonder if the role of chemical messengers in human relationships is different from population to population. We have not considered any relevant studies, but we are assuming that these mechanisms are something fundamental that all humans share as a species.

Responsibility

An obvious dilemma in the study of the role of chemical messengers in human relationships is that we would know more, and that we would be in a better position to make cause–effect inferences if we could conduct invasive research. Invasive studies are ones that alter the physical composition of the body: brain lesions, genetic modifications, injections of high doses of hormones, and so on. However, such studies are obviously unacceptable. We use animal research to partly solve the problem, and the ethics of animal research is another complex topic in its own right (see Chapter 1: Research methods and data interpretation).

5.8 Cognitive explanations for interpersonal relationships

Inquiry questions

- Can we predict whether a relationship will be maintained?
- How do we decide whether or not to break up?
- Do we analyse costs and benefits to decide if we should stay in a relationship?

What you will learn in this section

Key learning:

- Cognitive explanations suggest that relationships are influenced by how we perceive our partners and interpret our interaction with them.
- A number of cognitive models of romantic relationships exist. For example, social exchange theory (SET) suggests that individuals calculate costs and benefits to decide if a relationship is worth continuing. Equity theory suggests that the longevity of a relationship depends on its perceived fairness.
- According to Rusbult's investment model, there are three components of commitment (satisfaction, perceived quality of alternatives, investment size). Longevity of a relationship depends on the sum of these three factors. The theory also described maintenance mechanisms used to salvage the relationship.
- Empirical support for cognitive models of relationships mainly comes in the form of correlational studies. Experimental studies are also possible, but commonly take the form of responding to one of several versions of a hypothetical scenario. Longitudinal studies provide valuable information because they give us an insight into how cognitive variables develop over time, indirectly suggesting causality. They also allow us to investigate the predictive validity of the models.
- One common problem in this research is self-report measures, due to the nature of the phenomenon that is investigated.
- Another common problem is the existence of different possible interpretations of the same correlational data. Conclusions from individual studies are somewhat speculative, and we heavily rely on triangulation to reach greater certainty.
- Additionally, conducting longitudinal studies involves using the same self-report measures repeatedly, which may lead to biased responding and self-fulfilling expectations.

Key terms: social exchange theory (SET), equity theory, investment model of commitment, satisfaction, perceived quality of alternatives, investment, commitment, maintenance mechanisms, entrapment, predictive validity

In a wider context

Just like any other behaviour, interpersonal relationships are influenced by a range of biological, cognitive, and sociocultural factors. Biological factors were considered in the previous unit using the example of chemical messengers and the role they may play on various stages of a developing romantic relationship. This unit will focus on some of the existing examples of cognitive explanations of interpersonal (romantic) relationships.



▲ Figure 5.17 An elderly couple

Cognitive theories of interpersonal relationships

Cognitive explanations of interpersonal relationships suggest that relationships are influenced by how we perceive our partners and interpret our interaction with them.

For example, **social exchange theory (SET)** suggests that individuals perform a rational calculation of costs and benefits in order to decide if a relationship is worth pursuing. It views a romantic relationship somewhat like a business transaction. Thibaut and Kelly (1959), the authors of the theory, look at a relationship as a series of exchanges where individuals are interested in gaining more profit at a lower cost.

Another example is **equity theory** of interpersonal relationships (Hatfield and Traupmann, 1980), which suggests that a partnership will be stable if both partners perceive their relationship as fair: what they get from the relationship is roughly equal to what they bring into it. According to this theory, you will be dissatisfied with the relationship if you feel like your partner benefits from the relationship more than you do. However, the opposite is also true: if you think you benefit more than your partner, then the relationship will not be a stable one either. This is because you may experience feelings of guilt and shame.

This kind of reasoning has frequently been applied to cheating. When one partner cheats on the other and the other partner suspects them, the cheating one may feel guilty and the one who is being cheated on may feel deceived. This may lead to a break-up. It has been shown that these feelings emerge even if cheating is never actually exposed.

Yet another example of a cognitive explanation is the investment model of commitment. We will focus on this model as the main example in this section to look at the topic more deeply.

Investment model of commitment

The **investment model of commitment** was originally proposed by Rusbult (1980) to explain why people remain in relationships or break up. The investment model proposes three components of **commitment**.

- **Satisfaction** (i.e., generally having positive feelings about the partner and the relationship). The level of satisfaction will be high when: (a) the relationship provides high rewards, and (b) the costs are low.
- **Perceived quality of alternatives**. An alternative could be a different partner, but not only. It could also be, for example, solitude (remaining single) or spending time with relatives. Alternatives are evaluated similarly using expected costs and benefits.
- **Investment** size. Investments are all the things that you bring into the relationship to make it work—they include spending money, putting in effort, spending your time, making compromises, property that you bought together, shared children, and even shared memories.

The theory predicts that if the sum of these three factors is high, then the relationship will continue, even if one particular component of commitment is absent. For example, people are likely to remain in a relationship if they have made a considerable investment in it and if there are no good alternatives, even if satisfaction from the relationship is not high. Similarly, relationships will likely break apart if the sum total of the three factors (collectively called commitment) is not high, even if one of the components is very positive. For example, think of a relationship where the partners' satisfaction with each other is high, but at the same time they have not invested much in the relationship and they can both see attractive alternatives.

The investment model of commitment suggests that people who are unhappy often remain in a relationship because they have already made an investment in it and they do not want to lose the investment. For this reason, according to Rusbult, people use a variety of "**maintenance mechanisms**" to try and keep the relationship alive. For example willingness to sacrifice, forgiveness, and creating positive illusions about their partner.

Activity

Design your own theory of romantic relationships. It should be a theory that places primary importance on cognitive factors. Split into groups and limit yourself to 10 minutes.

Present your theory to other groups.

ATL Communication, Social, Self-management

SAT − ALT + INV = COM

↓

stay/leave

▲ Figure 5.18 Investment model of commitment
Satisfaction − Alternatives + Investment = Commitment

Chat with AI

Consider performing an analysis of a real-life situation through the lens of Rusbult's investment model with the help of AI.

Rusbult's model is not very well represented in popular writings, so we should give AI some helpful context. Try the following sequence of steps:

- Use a prompt like this: "What do you know about Rusbult's investment model of commitment? It is one of the cognitive explanations of romantic relationships in psychology. Could you provide a summary of this model in around 200 words, outlining its main propositions?"
- Look at the output and edit it if you find some inaccuracies or if you want to align it better with what you know from this textbook. "I have corrected your summary a little to better align it with what I am studying. From now on, let's use this as our understanding of Rusbult's model."
- Once this is done, give AI a real-life situation: "Here is a real-life situation related to romantic relationships. Could you provide an analysis of it from the point of view of Rusbult's investment model that we defined earlier?"
- Continue your interaction with AI, asking it questions for clarification to further understand how the model can be applied to practical situations.
- You could even ask this: "If you were a relationship counsellor who based their work on Rusbult's model, what ways would you suggest to the partners to save their relationship in this situation?" See if the AI output corresponds to your own intuition and life experience.

Activity

It is an important skill to be able to read original research articles.

Find Rusbult's (1980) article online. It is called "Commitment and satisfaction in romantic associations: a test of the investment model". The full text is available on the website of Northwestern University. The "role-playing experiment" that is described here is Experiment 1 in the article (pp. 176–180).

Read the full description of the study in the article. Have you discovered any additional details that add to your understanding of the investment model?

ATL Research, Self-management

Empirical support for the investment model: a role-playing experiment

Rusbult (1980) conducted a "role-playing experiment". Participants had to read a four-page essay where the protagonist was deciding between remaining in their current relationship and breaking up to pursue another relationship. Participants were asked to place themselves in the position of the protagonist (Robert for male participants, Sarah for females).

For the sake of this example, we will assume here that the participant is female. She was asked to identify with Sarah and the choice was between remaining in the current relationship with Robert or breaking up to pursue a relationship with John.

Three independent variables were manipulated in this experiment:

1. Relationship cost: high (Robert had moved 1000 miles away and they will only be able to see each other once a month) versus low (60 miles away, see each other once or twice a week).
2. Perceived quality of the alternative: high versus low (manipulated by changing the way the essay described John's appearance and intelligence).
3. Investment size: high, medium, and low. In the low investment condition, the essay said that Sarah and Robert had only been together for a month before Robert had to move to another location. In the medium investment condition, Sarah and Robert had been together for one year. In the high investment condition, Robert's father was actually Sarah's boss, so she felt like she would have to quit her job in the event of breaking up with Robert.

Commitment to the relationship was measured through a questionnaire with sample items like this: "How likely is it that you will pursue a relationship with John?"

Table 5.5 presents the descriptive results of this study.

	High cost		Low cost	
	Poor alternative	Good alternative	Poor alternative	Good alternative
Low investment	5.71	5.05	6.10	5.29
Medium investment	6.05	5.45	6.56	5.26
High investment	6.31	5.81	6.60	6.05

▲ Table 5.5 Results of the study. Values in the cells show mean commitment. It can range between 1 (lowest) and 9 (highest). (Modified and adapted from Rusbult, 1980, p. 179)

We must of course investigate the statistical significance of these differences before any conclusions can be made. Analysis revealed that the effects of investment size and quality of the alternative were indeed significant. In both high-cost and low-cost conditions (see Table 5.5), as investment increased, so did commitment. Additionally, commitment was lower if there was an attractive alternative.

It is worth noting that no differences were found in this study between the male condition and the female condition, suggesting that the results are equally applicable to both sexes.

The value of this study is that it is experimental: components of the model were actually manipulated by the researcher, making it possible to have cause–effect inferences. The obvious limitation in the study is that it is a hypothetical situation and, although participants were asked to put themselves into Sarah's (or Robert's) situation, they still answered questions about a hypothetical character in an imagined scenario. Additionally, although some statistically significant differences were revealed, you will notice from the table of results that the absolute differences are not that large (average commitment scores ranged between 5.05 and 6.60 in various conditions).

Empirical support for the investment model: a longitudinal study

To overcome limitations of the role-playing experiment, one needs to conduct a study that is related to participants themselves (i.e., not hypothetical).

Rusbult (1983) conducted a longitudinal study of the investment model. Longitudinal studies are especially valuable in this area because theories of relationships try to predict how relationships will develop over time.

The participants were 17 couples recruited via flyers (i.e., volunteers). Participants completed two types of questionnaire:

1. An initial questionnaire in which participants described their perception of the relationship at the very start.
2. Twelve identical relationship questionnaires completed every 17 days.

All questionnaires were designed to measure the components of the investment model: rewards, costs, alternatives, investments, satisfaction, and commitment.

Activity

Conduct your own role-playing experiment. Keep it simple—manipulate one variable and investigate how it affects commitment. Write a hypothetical scenario describing the relationship between two partners. Prepare two versions of this scenario to manipulate your dependent variable (e.g., long distance versus short distance). Do not feel limited by the investment model—investigate your own hypothesis. Create a short questionnaire to measure commitment (it could be one question).

After discussing the proposed experiment with the "ethics committee" (which in this case could be your teacher) and obtaining approval, conduct the study with your classmates. Split them randomly into two groups and give them the two versions of the scenario. Compare results. Report back to the class.

ATL Research, Communication, Self-management

Chat with AI

Can generative AI act as an ethics committee? Most certainly not, and it would be a bad idea to use it for this purpose. However, drawing your attention to some of the aspects of proposed research that may pose a difficulty from an ethical point of view would be very useful.

Provide a brief description of Rusbult's longitudinal study to your favourite generative AI. Have a conversation with it regarding the potential ethical issues in the study.

"Here is a description of a proposed research study to investigate the role of satisfaction, perceived quality of alternatives, and investments on levels of commitment in a romantic relationship. If I were to submit this proposal for approval to an ethics committee, which aspects of this study could be flagged as problematic and why? Could you have a conversation with me about this, asking me leading questions to help me understand how I can make this research more ethical?"

All questions were measured on a nine-point Likert scale. Participants were also asked to briefly describe anything unusual that happened in their relationships since the last time they completed the questionnaire. If they broke up, they were asked to describe the circumstances surrounding it.

The results of this study supported the predictions of the investment model. As predicted, higher levels of commitment were associated with greater satisfaction, larger investments, and poorer alternatives.

The **predictive validity** of the investment model was also investigated. Predictive validity is the ability of the model to correctly predict how the modelled phenomenon will develop over time. This study gave researchers an opportunity to compare three groups of participants:

1. "stayers" (those who remained in a relationship until the end of the study)
2. "leavers" (those who chose to break up)
3. "abandoned individuals" (those who were broken up with).

The results showed the following:

- The group of "stayers" demonstrated a dynamic in which the quality of alternatives declined over time and all other factors (satisfaction, investment size, the level of commitment) increased. A possible explanation for the decline in the perceived quality of alternatives is that partners in a relationship actually start perceiving alternatives as less attractive. Another possible explanation is that other people, knowing that the individual is in a relationship, do not show interest as much and therefore do not present themselves as plausible alternatives.
- In the group of "leavers", things were the other way around: satisfaction, investment, and commitment declined, whereas the quality of alternatives greatly improved.
- The group of "abandoned individuals" showed an interesting trend that Rusbult called "**entrapment**": with fewer rewards and more costs, their satisfaction was lower than in "stayers". However, they continued investing heavily in the relationship, sometimes even more than the group of "stayers". They remained "trapped" in their relationship and continued investing in it despite growing costs and declining satisfaction, until they were abandoned by their partners.

Limitations and challenges of research

Although these studies, by Rusbult's own words, provide a "fairly good" support for the investment model of relationships, they also highlight a number of issues that are inherent in most research studies in this area.

1. Most of this research relies on self-report measures. It is difficult to think of something more objective. How else do you measure partners' attitudes to each other, their commitment, and their satisfaction with the relationship? A related issue is participants' ability to make judgements about abstract constructs such as commitment and investment. For example, when you ask participants about the "costs" and "investments" in the relationship, will they make a reliable distinction between the two? Will all participants understand the distinction in the same way?
2. Most of the research in this area is correlational. Sometimes interpretations of correlational evidence may be speculative. For example in some

subsamples of Rusbult's study (e.g., men and "stayers"), there turned out to be a positive correlation between costs and commitment (the greater the costs, the stronger the commitment). Obviously, this finding does not fit well into the theoretical model. In an attempt to explain contradictions, Rusbult proposed several possible interpretations. One such interpretation suggests that costs for these groups of participants serve as a type of investment: "previous costs incurred with little reward may lead to increased determination to make those costs 'pay off' in the future" (Rusbult, 1983, p. 114). Such interpretations are just that—interpretations that are not directly supported by evidence. Another possibility is simply that the theory is incorrect. It requires triangulation of the results of multiple research studies to slowly construct a theoretical model that is true beyond a reasonable doubt.

3. There is the problem of repeated measurements. Since we are interested in the dynamics of how a relationship develops, we need to measure the same parameters (e.g., satisfaction) several times. But then how do we know that one measurement does not impact the results of subsequent measurements? Could it be that participants' responses on the second questionnaire are influenced by the mere fact that it is the second time they are completing it? To add even more twist, could it be that participants' behaviour itself is influenced by the fact they complete a questionnaire about it? It is not unreasonable to expect that the mere fact of saying on the questionnaire that they are "very satisfied" with the relationship or very committed to making it long-term can create a sort of self-fulfilling expectations where participants will behave in certain ways towards their partners because they have provided these responses on the survey.

TOK

Would you say that triangulation is equally important in other areas of knowledge, such as natural sciences, history, and mathematics?

Conceptual analysis

Perspective

Cognitive models of romantic relationships suggest that cognitive interpretation plays a role in how relationships develop over time—for example, whether or not partners stay committed to each other. Note: it is not common in research to directly compare the contribution of biological and cognitive factors in maintaining a relationship. It is not on the research agenda to find out which of the groups of factors is "most important". There is a wide recognition that they are both important in their own right.

A more relevant link to the concept of perspectives is perhaps the existence of different cognitive models of romantic relationships. Each of the models has gained some empirical support. How do we decide that one theoretical model is a better description of reality than another?

Causality

The problem with research in this area is that all theoretical models are causal, but most empirical research is correlational. This creates the situation in which many conclusions are somewhat speculative and there could exist various alternative explanations for the same findings. There do exist experimental studies (such as Rusbult's role-playing experiment), but they are limited to hypothetical scenarios and lack ecological validity. The problem is partially solved by conducting longitudinal studies, but even those are not without problems.

Measurement and bias

In this section, we considered a number of typical limitations and challenges that characterize research in this area. Research is dominated by self-report measures because it is difficult to assess perceived relationship parameters otherwise. Correlational data also dominates the field. There is bidirectional ambiguity and a large number of mediating variables, which may mean that multiple theoretical interpretations of the same finding could be possible. Repeated measurements present the problem of self-fulfilling expectations. Research is dominated by studies from Western countries, but it is reasonable to expect that cognitive appraisal of relationships can be heavily influenced by culture.

Change

Relationships change over time and the study of cognitive factors can contribute to our understanding of this change. Cognitive variables are much more powerful than biological variables in explaining the long-term dynamics of relationships. We have seen here that there may be distinct groups of individuals characterized by how factors such as satisfaction, alternatives, and investment interplay in a process of cost–benefit analysis to influence their decision on whether or not to continue the relationship.

Responsibility

Research in this area is not invasive like some studies of biological factors of behaviour, but it commonly involves questionnaires around sensitive topics (participants' personal lives). It is more important here than in many other areas to ensure confidentiality and the right to withdraw. Ethical considerations here also overlap with methodological considerations. For example, if participants are not certain that their responses will be strictly confidential, there is a high chance that the responses will not be entirely honest, or that participants will be influenced by response biases such as social desirability.

5.9 The role of communication and language, and strategies for improving relationships

Inquiry questions

- What can be considered effective and ineffective patterns of communication in a relationship?
- How can partners who are on the verge of break-up be helped?
- Can we predict relationship problems at an early stage?

What you will learn in this section

Key learning:

- Interpersonal communication is worth studying because positive interpersonal relationships are a major source of well-being and happiness.
- People use interpersonal communication to fulfil goals (self-presentational, relational, instrumental). Every message in communication contains both content and relational information. Communication consists of verbal and non-verbal messages, it varies in effectiveness, it can be symmetrical and asymmetrical, and it is inevitable.
- A crucial role of communication is in resolving interpersonal conflict. Having an experience of constructively resolving conflicts may be more valuable than simply managing to avoid them. There are different styles of conflict resolution in relationships which can be categorized along two dimensions: uncooperative versus cooperative, and direct versus indirect.
- Negative reciprocity tends to be self-reinforcing and it is important to avoid it by noticing its early signs and recognizing non-productive communication patterns (such as "gunnysacking" or "button pushing").
- Early signs of relationship break-up due to communication problems can be identified. Such is the model of "four horsemen of relationship apocalypse": criticism, contempt, defensiveness, stonewalling. Research in this area is based on longitudinal studies in which correlational research is combined with "microanalytic observation of behaviour" in a laboratory.
- In order for relationships to be improved, patterns of communication need to be fixed. This may be achieved by teaching partners how to recognize early signs of problems and replace negative patterns of communication with more positive ones. Comprehensive intervention programmes may be designed and an example of this is Gottman Couples Therapy. This is based on a gradual enhancement of trust and commitment by a step-by-step approach of teaching partners to engage in positive communication (such as openly expressing respect and appreciation of each other).



Exam tip

We are combining two topics in one section: "The role of communication and language" and "Strategies for improving relationships". This is because the two topics are closely related to each other: strategies for improving relationships are based on replacing negative communication patterns with more positive ones.

- Studies have demonstrated the effectiveness of such programmes even in cross-cultural contexts, although there are many possible sources of bias and conclusions must be made with caution.

Key terms: interpersonal communication, verbal, non-verbal, symmetrical, asymmetrical, self-presentational, relational, instrumental, conflict resolution, styles of conflict resolution, negative reciprocity, four horsemen of apocalypse, criticism, contempt, defensiveness, stonewalling, microanalytic observation of behaviour, Sound Relationship House theory

In a wider context

Biological variables (e.g., chemical messengers) may play a role in the onset of relationships or creating the feeling of attachment between partners. Cognitive variables (e.g., cognitive interpretation of costs, benefits, and investments) may play a role in how partners perceive their relationship and how satisfied they are with it, affecting its stability and longevity. However, the mechanism of a relationship is, without a doubt, communication. Communication is how we "run" the relationship on a daily basis, how we initiate contact, resolve conflicts, and move from superficial acquaintance to intimacy.

Importance of interpersonal communication

Human beings are fundamentally social creatures who depend on each other for survival and well-being. This is why communication is central to being a human.

Today the importance of research on personal relationships is widely recognized, but until recently public opinion was not in its favour. Two of the early pioneers of psychological research of love and marriage relationships were Ellen Berscheid and Elaine Hatfield (formerly Elaine Walster). Reportedly, a senator in Wisconsin publicly criticized them in 1975 for their research on love (Guerrero, Anderson, and Afifi, 2021). His argument was that scientists can never understand the mystery of love and even if they could, nobody would be interested in their findings.

However, for many people, personal relationships are much more important than any other aspect of life. In fact, relationships are the most common topic of discussion among people, surpassing other topics such as politics or climate change. Positive interpersonal relationships are a major source of well-being and happiness. Unsurprisingly, positive interpersonal relationships serve as a powerful protective factor against health problems and mental disorders. Bookstores and mass media are full of self-help literature on how to navigate interpersonal conflict, find the right partner, and build a trusting relationship. This makes scientific research even more important as it provides the necessary checks and balances for the unsubstantiated claims in popular media.

Principles of interpersonal communication

Guerrero, Anderson, and Afifi (2021) distinguished the following six principles of **interpersonal communication**:

1. It consists of **verbal** and **non-verbal** messages.
2. Every message contains both content and relational information. Beyond the content that they transmit, messages also tell people something about the

relationship that they share. Think about the following question: "What are you doing tonight?" You would agree that this message is much more than a simple request for information.

3. People use interpersonal communication to fulfil goals: (i) **self-presentational** (creating and conveying a certain image of ourselves), (ii) **relational** (establishing friendships and romantic ties, sharing affection), (iii) **instrumental** (asking for assistance, influencing someone's attitudes and behaviours).
4. Communication is inevitable: it is impossible to not communicate. Think about receiving a text message from someone, reading the message but not messaging back. This in itself is a message that the sender will interpret in a certain way.
5. Interpersonal communication varies in effectiveness. Meaning may be lost and distorted when it is coded in a message, sent, and then decoded by the receiver.
6. Interpersonal communication can be **symmetrical** or **asymmetrical**. An example of symmetrical communication is: "Will I see you later today?"—"Yes, looking forward to it". An example of asymmetrical communication is: "I love you"—"Thank you").

Discussion

There are a lot of self-help books available for those whose relationship is not working. What do you think of such literature and how do you think it compares to getting professional help from a marriage counsellor?

Under what circumstances would you consider reading a self-help book? Do you think they have a positive value?

ATL Self-management

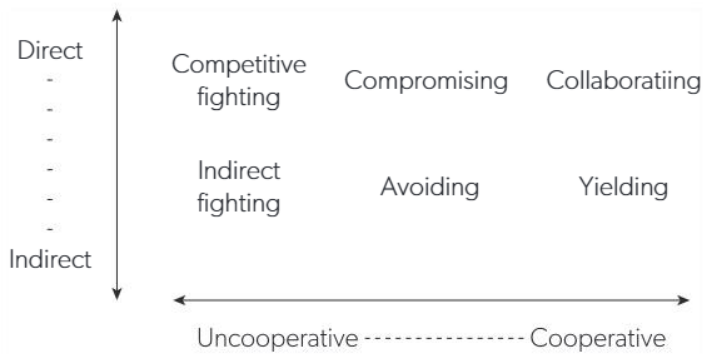
Communication and coping with conflict in relationships

Interpersonal conflict is a common part of relationships. It is not necessarily a bad thing: depending on how we resolve the conflict, we can emerge from it having a stronger relationship than ever before. Having the experience of positive and constructive **conflict resolution** is important.

Guerrero, Anderson, and Afifi (2021) distinguished six **styles of conflict resolution** in relationships (similar typologies were also proposed by other authors). These six styles are positioned along two dimensions, as shown in Figure 5.19.

1. Uncooperative (focusing on one person trying to win the argument) versus cooperative (focusing on the goals of both people and trying to find a compromise).
2. Direct (engaging in conflict and talking about issues) versus indirect (avoiding the conflict).

The six styles of conflict resolution are: (1) competitive fighting, (2) compromising, (3) collaborating, (4) indirect fighting, (5) avoiding, (6) yielding.



▲ Figure 5.19 Styles of conflict resolution according to Guerrero, Anderson, and Afifi (2021)

Although indirect and competitive styles of conflict resolution have negative impacts on relationships, they are also the ones that are most commonly used. This may be due to the well-known principle that aggression tends to escalate: once an act of aggression or direct opposition has been performed by one person towards another, the other person may feel compelled to reciprocate. This idea that "aggression begets aggression" is known as **negative reciprocity**.

Chat with AI

There are many typologies, or "styles", of conflict resolution proposed by various authors. Some of these are evidence-based and published in peer-reviewed journals, some self-help books discuss strategies that are speculative.

One of the key limitations of generative AI (at least at the time of writing) is that it has no mechanism of separating empirically justified information from popular speculative accounts. It is trained on massive amounts of texts from the internet, so everything that has been published on the internet has contributed to the training of it. With topics such as types of neuroplasticity or social identity theory, it is not much of a problem, but with the role of communication in interpersonal relationships things are very different. Therefore, your large language model (LLM) might have been trained on some highly debatable material.

In situations like this, it is advisable to give AI a clear context. For example, consider the following composite prompt:

"Here is a brief explanation of six styles of conflict resolution in relationships, as proposed by Guerrero, Anderson, and Afifi (2021): [insert an explanation of the model in general as well as the specific styles]. Here is a description of a series of interactions during conflict between romantic partners: [Insert a description or a transcript of romantic partners having a conflict; you could make something up or copy and paste a dialogue from movie subtitles]. Could you help me analyse this conversation from the point of view of the six styles?"

There are many "methods" that people may use to trigger negative reciprocity and further escalate the conflict. Examples include:

1. "gunnysacking" (not discussing issues as they arise, but storing them in a metaphorical "gunnysack" and dumping them on your partner all at once when the situation presents itself)
2. "kitchen sinking" (repeating the whole collection of your old arguments as you get into a new argument; bringing up old issues that have already been discussed)
3. bringing third parties into the argument (e.g., attacking your partner's family or comparing your partner to other people in a hurtful way)
4. "button pushing" (intentionally saying or doing something that you know will be a trigger for your partner)
5. empty threats (e.g., saying "I will break up with you" even though you are not actually intending to do so)
6. "mind reading" (e.g., "You never cared about my career").

The four horsemen of apocalypse

Dr John Gottman and Dr Julie Gottman, the founders of the Gottman Institute (known colloquially as the "Love Lab"), conducted extensive research on patterns of communication predictive of break-up or divorce in married couples. With this purpose, couples were invited to have conversations with each other on a variety of topics while their interaction was recorded on camera. Researchers would then watch the recordings and code various communicative behaviours using an observation checklist. This would sometimes be combined with physiological parameters (such as pulse and blood pressure) taken continuously during the interaction. All research was longitudinal: researchers kept track of the relationship status of all participants over long periods of time (Gottman, 1994).

The **four horsemen of apocalypse** is the theoretical model created based on this research. According to this model, there are four key types of communication that, when combined, are strongly predictive of divorce: (1) criticism, (2) defensiveness, (3) contempt, (4) stonewalling.

A **criticism** is a personal attack that assumes that the partner bears the blame for the behaviour, for example: "You always play video games because you don't care about spending time with me". A more constructive way to express the same idea (an antidote to criticism) would be complaining about the specific behaviour instead of blaming the person, and preferably starting with an I-statement, for example: "I feel lonely and kind of anxious when you don't get to spend any time with me because of video games".

Contempt is a kind of hostile humour, mockery, and sarcasm that communicates that you have given up on the relationship and you simply blame the partner for everything that is happening. It may include direct verbal remarks, but could also be non-verbal—for example, rolling your eyes when your partner speaks (so that other people can see). The opposite of contempt is respect, and the Gottman Institute argues that it is healthy to not only respect your partner, but to express this respect in communication.

Defensiveness is protecting oneself against personal attacks, often by shifting blame to the other party. When you are attacked, it is only natural to expect that you will be defensive. However, in most real-life conflicts both parties share at least partial responsibility for the problem. A constructive way of conflict resolution that helps avoid the vicious cycle of escalation would be to accept a part of this responsibility and avoid shifting the blame on the other.

Stonewalling is a serious sign of relationship problems and usually occurs at a stage when conflict has occurred frequently and repeatedly and both partners have "given up" and withdrawn from interaction. It is usually the result of very intense emotions. Because emotions are running high, people may say serious things in the moment that they do not mean and will regret later.

Activity

Take several transcripts of conversations in which romantic partners are having a conflict. You could write them yourself with your classmates. Alternatively, you can find them online, taken from your favourite books or movies. If that is too problematic, then you can simply ask your favourite generative AI to generate some hypothetical conversations (instruct AI about the expected length and ask it to make them different from each other).

Now divide into groups and write out any signs of the four horsemen (criticism, contempt, defensiveness, stonewalling). Compare your observations with other groups. Did you all agree in your interpretations?

If you are alone, you could also ask your generative AI to do the same task and compare if the AI's results are the same as yours.

ATL Research, Social, Self-management

Research of conflict resolution in communication

Gottman and Krokoff (1989) conducted longitudinal observational studies of couples. Their study was aimed at predicting the changes in marital satisfaction over time. They call the method they used in the study the "**microanalytic observation of behaviour**". Participants in the study were married couples who had been married for 24 years on average.

Discussion

To what extent do you think the situational nature of the measurement is a potential source of bias in studies like this? Participants come to the laboratory and have a conversation for 15 minutes and then we correlate parameters of that conversation with long-term outcomes such as the probability of divorce. To what extent is the 15-minute conversation representative of all other conversations that these partners have with each other in real life? To what extent are all their conversations the same? Don't they change from one situation to another?

ATL Research, Thinking

Marital satisfaction questionnaires were given to the couples twice: at the start of the study and three years later (Time 1 and Time 2). Couples were recorded on video as they discussed an area of disagreement in their marriage. Additionally, they discussed another area of disagreement at home and were required to record this discussion. All recordings were transcribed and analysed by coders using coding schemes. Examples of behavioural categories used in these coding schemes include: agree, humour, smile, laugh, excuse, deny responsibility, criticize.

Here are some interesting results.

- Overall the pattern of correlations supported the hypothesis that defensiveness, stubbornness, and withdrawal are predictive of both distress at the present moment and deterioration of satisfaction over time (you would have noticed that these variables are related to three of the four horsemen of apocalypse).
- For the wife, positive verbal communication was associated with marital satisfaction at present, but it also predicted deterioration of marital satisfaction over time.
- For both partners, conflict engagement was associated with marital dissatisfaction at present, but it also predicted improvement in marital satisfaction over time.
- Observations based on home recordings and recordings in the laboratory were similar in the conclusions they suggested. This allows us to treat results of laboratory observations as generalizable to real-life situations of interactions at home.

The conclusion of the study is that not all conflict is the same. Only when conflict is handled negatively (e.g., with defensiveness, stubbornness, and withdrawal) does it lead to relationship deterioration. Positively resolved conflicts, while causing some negative feelings in the moment, contribute to a happier marriage in the long term.

Strategies for improving relationships

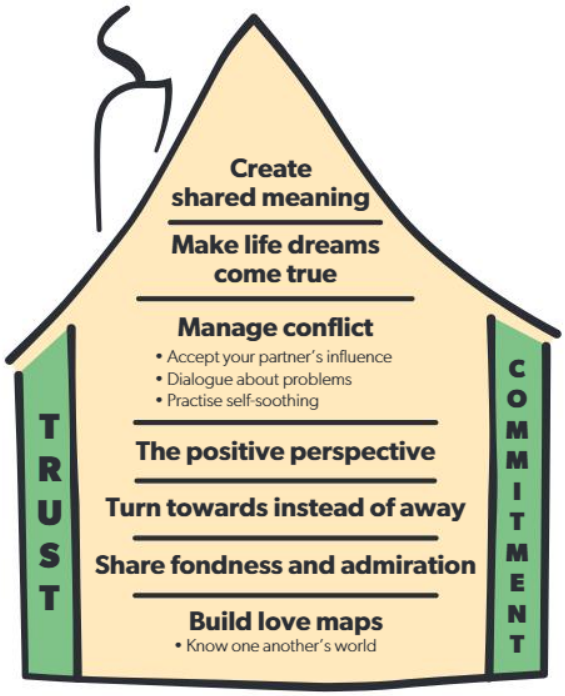
In order for relationships to be improved, patterns of communication need to be "fixed". Partners should learn how to recognize ineffective communication and replace it with effective patterns. Looking at everything that we have discussed in this section so far, this means the following:

- Avoid triggering negative reciprocity by recognizing its early signs. Use more constructive styles of conflict resolution such as compromising and collaborating.
- Recognize signs of the four horsemen of apocalypse and work together with the partner to replace such communication patterns with more constructive ones.

Many of these methods have been "packed" into intervention strategies, an example of which is the Gottman Couples Therapy.

The therapy claims to be based on interventions that are grounded in research. The Gottmans referred to the theory behind their therapy as the **Sound Relationship House theory**. According to this model, there are two "weight-bearing walls" that play a key role in holding a couple together at all times, along

with seven "floors" that represent the elements that a couple can use to further improve their relationship. If you get to the top of the house, you have a high chance of having a long-term "sound relationship".



▲ Figure 5.20 Gottmans' Sound Relationship House

The two "weight-bearing walls" are trust and commitment. Trust enables partners to rely on each other. Commitment enables them to have a sense of certainty that their intentions are serious and any problems are only temporary.

The seven "floors":

1. Building love maps—on this floor couples get to know each other's personalities.
2. Sharing fondness and admiration—on this floor couples clearly express their respect and appreciation for each other.
3. Turning towards, not away—on this floor couples learn to recognize when the partner needs attention and cater to their needs.
4. Having the positive perspective—on this floor couples make positive attributions and learn to recognize that circumstances, and not only their partner, may be responsible for a problem.
5. Managing conflict—here couples learn to deal with conflict constructively by acknowledging each other's feelings and having an open discussion.
6. Making life dreams come true—on this floor partners learn to support each other in their dreams and aspirations.
7. Creating shared meaning—on this floor couples create a joint world where they have rituals and stories that have a shared meaning for the two of them.

Therapy typically involves an initial assessment (individual interviews with each partner and questionnaires), design of an individual therapy plan, a series of sessions with specific assignments and home tasks. Therapy is also combined with informational interventions (educating the clients about evidence-based

Activity

Role-play marriage counselling. Split into groups of three. One person will be the counsellor. The other two will be a married couple. The married couple will need to think of a "problem" that they have been experiencing lately and present it to the counsellor. Make sure to use some negative communication patterns, perhaps even some of the four horsemen of apocalypse.

The job of the counsellor will be to steer the conversation in a way that would help the couple work on their problems. Use the seven "floors" in the relationship building according to the Gottman Couples Therapy as your guiding principle. What would your assignments and methods be, for example, on the second "floor"—sharing fondness and admiration? What exactly would you ask your clients to do? Will they have any home assignments to work on this skill?

It would be good if this could be role-played in front of the whole class, followed by a whole-class analysis.

ATL Communication, Social, Self-management

principles of solid relationships). At the core of the therapy lies increasing the clients' mindfulness about important aspects of relationships as well as positive and negative communication dynamics.

Research of effectiveness of communication improvement strategies

There have been a number of studies that investigated the effectiveness of couples therapy based on the Gottman method. As an example, let us consider Rajaei, Daneshpour, and Robertson's (2019) study of Iranian couples dealing with conflict.

The social problem that triggered this study was the increasing number of divorces in Iran in recent years. In addition to that, it is common for couples to experience emotional divorce—this is when you stay married officially, but become cold to each other and live separate lives, even though you may continue sharing the same home. This is especially pertinent in Iran where divorce is often stigmatized due to traditional cultural values and religious beliefs. Such a unique cultural context makes the study more valuable because it puts to test cross-cultural generalizability of the Gottman method (the effectiveness of which was mostly established in research using Western samples).

The sample in the study consisted of 14 married couples (volunteers) that sought relationship therapy in a mental health clinic in Iran. Couples were randomly assigned to the experimental group and the control group.

- Participants in the experimental group received eight sessions of couples therapy based on the Gottman method. They were given questionnaires at the start of the therapy (Week 1) and at the end of it (Week 8).
- As participants in the experimental group went through therapy, those in the control group were wait-listed. However, in Weeks 9–17 of the study they also underwent the same therapy.



▲ Figure 5.21 An Iranian couple

There are some obvious limitations in this study. The sample in the study was very limited: only 14 heterosexual couples from the same location within Iran. Importantly, these couples sought therapy voluntarily. We can argue that this in itself makes these couples more Westernized than others. One of the authors of the study was also the therapist delivering the sessions, which makes researcher bias a possibility. The study was not blinded: the therapists knew which couple belonged to which condition.

Conceptual analysis

Perspective

One might say that communication is a sociocultural phenomenon. However, cognitive variables seem to be closely intertwined with communication. Communication is a series of turns, and at each turn one of the partners decides what to say—this decision is a cognitive process. To be fair, biological variables are also present at all times. For example, in an argument emotions run high, and emotional regulation is associated with a massive release of chemical messengers. As always, the biological, cognitive, and sociocultural perspectives are different ways of looking at the same process, but we need all three ways to fully understand it.

Causality and measurement

When we investigate the role of communication in human relationships, we want to make cause–effect inferences. However, it is hard to imagine how we could manipulate communication as an independent variable. By splitting participants randomly into two groups and asking one group to have a constructive conversation and the other to have an aggressive conversation? Of course this will not work because people would be simply playing a role and attacks would not be taken personally. Moreover, if communication influences relationships, it is through the recurring patterns that appear in multiple situations, including at home. It is for these reasons that we are limited in what we can do to quasi-experiments (comparing pre-existing groups) and correlational studies.

Correlational studies can be concurrent and longitudinal. A careful examination of the obtained correlational patterns combined with triangulation of evidence from multiple studies brings us closer to speaking about the "influence" of communication on relationships with confidence.

Bias

If a research study is biased, then our conclusions will be biased also. Most, if not all, of the measurements are based on self-report (apart from the obvious outcome variables such as break-up or divorce). There exist multiple measures of the same construct. For example, there are several widely used questionnaires of marital satisfaction, and even Gottman had to use several different coding schemes to analyse observed behaviour to rule out the possibility that different coding schemes produce different results. Most evidence is correlational. Observational studies where communication between partners is observed in a laboratory are promising, but there is always a danger of demand characteristics (e.g., participants will interact in a way that is different to what they normally do at home). Researcher variables such as confirmation bias are difficult to control.

Discussion

Do you think the study of Rajaei, Daneshpour, and Robertson's (2019) is emic or etic? Why? Give reasons for your answer.

ATL Research, Thinking

Change

The study of communication in human relationships is a study of change. We have seen through the research presented here that communication patterns may change over the course of time. How partners treat each other and what they expect from each other is not the same at different stages of a relationship. A huge discovery is that patterns of communication are predictive of whether or not partners will remain satisfied with their relationship. Gottman even went as far as claiming that it is possible for the Gottman Institute to predict with more than a 90% probability whether or not you will get divorced, just by observing the first three minutes of conversation with your partner in their laboratory.

Responsibility

Perhaps the most obvious aspect of ethics in the study of communication in human relationships is the sensitive nature of topics. In a laboratory observation, married couples may be required to have an argument about a topic that they previously identified as sensitive, and be recorded on camera. This should be carefully handled through informed consent, clear opportunity to withdraw, confidential treatment of data, and so on.

There are also ethical considerations that transcend an individual study. Researchers in psychology are ethically responsible for being evidence-driven advocates of social changes. Research of romantic relationships used to be dismissed as non-important. Psychologists know exactly how important this study is to ensuring mental health and well-being, but the rest of society may be under the influence of stereotypes and misconceptions. Therefore, it becomes our responsibility to educate the public. Things are further complicated by a large number of sources of advice, such as self-help books and social media posts, only some of which are based on credible research findings.

Exam-style practice questions

Paper 1 Section A (4 marks)

Explain conformity with reference to one example.

Paper 1 Section B (6 marks)

Football hooliganism is a form of destructive behaviour by some spectators at football events. It typically involves conflict between supporters of different teams. Sometimes football fans form long-standing clubs that engage in rivalry with each other. Some fan groups follow their team from country to country. Conflict may occur spontaneously, triggered by something during a game—for example, a controversial referee ruling. However, in other cases attacks on rival fan clubs are planned in advance. Football hooliganism has become a major problem and today there are strict measures in place to prevent it, including banning certain groups of people from entering a country at the time of a football event.

How can you explain football hooliganism using social identity theory (SIT)?

Paper 1 Section C (15 marks)

A psychologist is investigating human relationships by carrying out a study of the effectiveness of various compliance techniques. With reference to one or more compliance techniques, discuss considerations related to the psychologist's responsibility.

(Concept: Responsibility. Content: Compliance techniques. Context: Human relationships)



6 HL extensions

Overview

There are three Higher Level (HL) extensions in the IB DP Psychology course:

1. The role of motivation in shaping human behaviour
2. The role of culture in shaping human behaviour
3. The role of technology in shaping human behaviour

These extensions are not meant to be separate units, but rather additional "lenses" through which you could look at the material you already know from the four contexts (Learning and cognition; Human development; Health and well-being; Human relationships).

What does the idea of "lenses" mean for exams? It means that you will not have any exam questions directly on the content of this chapter. Material presented here is meant as practice for you to exercise the skills of analysing research and evaluating claims in psychology.

How it will be assessed

HL extension material is assessed indirectly in Paper 3. This is a paper focused on data analysis and interpretation. The paper will include a resource booklet with three to five sources. Each source will describe findings of a research study (they will be either adapted from existing research or created specifically for the exam). The sources will all be related to one of the HL extension areas (i.e., motivation, culture, or technology) in one of the contexts (i.e., Learning and cognition; Human development; Health and well-being; Human relationships). The sources will be from both qualitative and quantitative research and can be experimental or non-experimental. Some sources will include graphs or charts representing the data obtained in the study.

Paper 3 will include four semi-static questions. The first three questions will be based on interpreting the data presented in one of the sources. These questions are designed to assess your understanding of how to interpret data and evaluate a research study. You will not need any knowledge of content outside of the description of research that is provided as part of the resource booklet.

Question 4 in the paper will be a synthesis question. It will provide a claim and ask you to discuss the validity of this claim based on at least three of the provided sources, as well as your own background knowledge in this area.

Structure of the chapter

With all this in mind, we have structured the chapter as follows:

- There are three parts: motivation, culture, and technology.
- Each part opens with a brief overview of the area of research as a whole. Where applicable, we describe the typical research limitations and challenges in this area. Since they are "typical", in all likelihood they will apply to any individual study.

- This is followed by one section per each of the contexts (Learning and cognition; Human development; Health and well-being; Human relationships). Therefore, there are 12 sub-sections overall, for example: Motivation in human development, Technology in health and well-being.

Each of the 12 smaller sections has the following structure:

- A brief overview of this area of research in general
- Two or three sources each describing a specific research study and its key findings
- A claim
- "Analysis of the claim"—here we discuss the validity of the claim, making references to the sources as well as (where applicable) our knowledge of the typical challenges of research in this area.

All research studies provided in these sources are taken from existing published research (this will not necessarily be the case in exams where research studies may be "invented" specifically for exam purposes).

Where possible, we have selected the studies such that they:

- represent a variety of research methods, both experimental and non-experimental
- are quite typical for the given area of research. The challenges inherent in the studies are likely to be present in other research in this area also
- provide different perspectives or express different points of view related to the claim
- contribute to the discussion of the claim in a non-repetitive manner.

We have also tried to formulate the claim in all 12 cases in such a way that the analysis of the claim takes a slightly different angle each time. For example, sometimes the claim is not supported by the sources, but sometimes it is. In some cases, it could be said that the claim in its current form is not supported, but a weaker version of the same claim would be. In other cases, it could be said that the provided evidence supports some aspects of the claim but not others.

How to use this chapter

The chapter has been designed this way to provide you with an opportunity to practise evaluating psychological claims based on provided evidence as well as background knowledge.

It is impossible to predict which studies, or even what type of studies, will be used in exams. The claim that you will need to discuss the validity of is equally unpredictable. The best way to prepare for such assessments is to exercise your skills in a variety of circumstances that model various possible scenarios.

Material and discussions presented here are merely examples. By no means should they be taken as a prescribed way of thinking or even as an exhaustive analysis. In each case our analysis of the claim, as well as the given sources, may have taken a different turn or focused on different aspects. Consider this as a scaffold that is designed to show you a suggested way of approaching these questions, with the ultimate goal of enabling you to "improvise" in the exam situation.

6.1 Motivation

What you will learn in this section

Key learning:

Theories of motivation:

- Definition of motivation.
- Different views on the role of cognitive processing in motivation.
- Content theories and process theories of motivation.
- Mechanistic motivation: psychoanalysis and behaviourism.
- The role of choice in motivation.

Self-determination theory:

- Intrinsic and extrinsic motivation.
- Self-determined and control-determined behaviours.
- Mediating variables in the influence of external events on intrinsic motivation: perceived locus of causality and perceived competence.
- Controlling behaviour versus supporting autonomy.
- Psychological needs related to intrinsic motivation: autonomy, competence, relatedness.

Key terms: motivation, mechanistic motivation, choice, self-determination theory (SDT), intrinsic and extrinsic motivation, self-determined and controlled behaviours, perceived locus of causality, perceived competence, autonomy, relatedness

What is motivation?

Motivation is the reason for which humans or animals engage in a particular behaviour. It is the driving force behind that behaviour. The word "motivation" comes from the Latin word *movere* (to move). Interestingly, the word "emotion" also comes from the same word. Motivation is an internal state that provides an impulse for people to engage in a specific behaviour.

The concept of motivation dates back to the very start of psychology and was widely used prior to behaviourism. In fact, behaviourists were hugely dissatisfied with the term and wanted it eliminated from the scientific vocabulary. For them, speaking about "motivation" was speculating about invisible forces that drive behaviour. Speculations are unscientific, they said.

However, unobservable factors (the "black box") made their way back into psychology since it was acknowledged that classical and operant conditioning alone were not enough to explain the behaviour of animals, let alone humans. This creates a bit of confusion: what exactly is motivation? Is it another factor influencing behaviour over and above the triad biological–cognitive–sociocultural? Is it a mental process?

Overview of theories of motivation

There are a large number of theories of motivation.

One dimension along which various theories of motivation differ from each other is the involvement of cognitive processing.

- Some approaches claim that the idea of motivation is applicable to both humans and animals, and that there is a certain continuity in the sense that humans have inherited some animal forms of motivation, but have also developed others. In this approach, motivation does not necessarily have to be something rational or conscious or intentional. For example, a rat running in a maze does not think or build plans or have intentions the way humans do, but the rat is still "motivated" by something.
- Other approaches draw a line between animals and humans when it comes to motivation. They claim that animals may be driven by environmental stimuli (such as electric shocks) and internal stimuli (such as hunger), but they are not motivated. These approaches suggest that motivation is a product of rational deliberation and cognitive processing that only humans are capable of.

Another way to divide theories of motivation is into content theories and process theories.

- Content theories attempt to describe which goals motivational states are commonly directed at. One example of such a theory is Maslow's hierarchy of needs.
- Process theories describe the cognitive and emotional processes that occur in an individual's mind when we experience a motivational state. Expectancy theory, goal-setting theory, and self-determination theory are examples of these.

Any theory of motivation must address two aspects of behaviour: energy and direction. Energy is a matter of needs: the stronger our needs, the more energy we will have. Direction is what elements of the environment we will target our behaviour at, to satisfy those needs.

Like many other phenomena in psychology, motivation cannot be directly observed and has to be inferred from other observable characteristics.

Mechanistic motivation versus choice

The earliest theories of motivation were theories of drives. Different psychologists placed the key importance on various drives. For example, Sigmund Freud (the father of psychoanalysis) believed that organisms are driven by two basic drives—sex and aggression (Freud, 1916). Hull (1943) created a theory of motivation which asserted that organisms are motivated by four main drives—hunger, thirst, sex, and the avoidance of pain.

Even behaviourism is consistent with basic drive theories because reinforcement may be seen as reduction of drive. A rat is hungry (drive), and therefore it engages in trial-and-error behaviour in search of food. When it finds food, it gets reinforced (drive is reduced). Therefore, this form of behaviour will strengthen in the future.

What behaviourism and psychoanalysis have in common is the insignificant role they ascribe to intention and **choice**. According to psychoanalysis, we are

caught between basic instinctive drives and constraints of the environment, trying to do whatever we can in the given conditions to satisfy the drives. According to behaviourism, our behaviour is a consequence of the learned associations between environmental stimuli and drives.

However, there were some developments that laid the groundwork for a more choice-oriented perspective on human behaviour. "Increasingly, choice and decision-making replaced stimulus–response associations to explain the direction of behaviour" (Deci and Ryan, 1985, p. 6).

Self-determination theory

Self-determination theory (SDT) (Deci and Ryan, 1985) is the belief that people are capable of making choices and that they show "inherent growth tendencies" that drive their own development over and above, and sometimes in spite of, external stimuli.

The theory makes a distinction between **intrinsic and extrinsic motivation**. Extrinsic motivation is when we do something because we expect a reward for it. For example, an office worker who does not really enjoy their job but does it for money. Intrinsic motivation is when we do something because we find the process itself rewarding. For example, young children are intrinsically motivated to play outside, or you may have a hobby that you enjoy doing in your spare time.

The theory also makes a distinction between **self-determined behaviours** and **control-determined behaviours**. This is not quite the same as intrinsic versus extrinsic because it is possible for behaviour to be self-determined but extrinsically motivated.

Self-determined behaviours are driven by choices that people make based on awareness of their needs and goals. Sometimes extrinsic reasons are internalized, where the person actively and purposefully engages in a certain behaviour even though the behaviour itself may not be intrinsically rewarding. For example, a person may not be a huge fan of cleaning their house, but they would make time and put effort into doing it because they have internalized the importance of it. We experience self-determined behaviours as initiated and regulated by ourselves—in other words, the perceived locus of causality is internal.

In contrast, control-determined behaviours are driven by events in the environment or inside the person. These behaviours are determined by demands rather than choices. For example, a student who studies science because their parents forced them to, and not because it was their choice. You might say that such behaviour is extrinsically motivated and not internalized. We experience such behaviours as initiated and regulated by something from the outside—in other words, the perceived locus of causality is external.

According to the theory, there are two important mediating variables in the influence of external events on a person's intrinsic motivation:

1. **Perceived locus of causality** for that behaviour. If the external event promotes a more external perceived locus of causality, it will diminish intrinsic motivation. If the external event promotes a more internal perceived locus of causality, it will increase intrinsic motivation. Events that promote an external perceived locus of causality are said to *control behaviour*. Events that promote an internal perceived locus of causality are said to *support*

autonomy. It is hypothesized that people often respond to events that control behaviour by rebelling against it or, which is a more likely scenario, by complying with the control temporarily, but defying it once it is no longer in full force. Examples of events that support autonomy include giving a choice and providing positive feedback. Examples of controlling events include rewards, deadlines, and surveillance.

2. **Perceived competence.** External events that promote greater perceived competence will enhance intrinsic motivation, and vice versa. When someone succeeds at a task or gets positive feedback, this is hypothesized to increase their perceived competence. Conversely, when someone keeps failing at a task or gets negative feedback, this may decrease their perceived competence.

Another idea that is included in the general framework of self-determination theory (SDT) is the existence of three psychological needs that are related to intrinsic motivation:

1. The need for **autonomy**. Autonomy is the belief that one can make one's own choices.
2. The need for competence. Competence is the belief that one is capable of masterfully and successfully interacting with the environment while performing certain tasks.
3. The need for **relatedness**. Relatedness is having relationships or bonds with people surrounding us.

According to the theory, intrinsic motivation and self-determined behaviour will be maximized when all three needs are satisfied.

6.2 Motivation in human relationships

What you will learn in this section

Key learning:

- What drives people to join groups? Terror management theory as one possible answer.
- Analysis of the claim "the leading motivation for people to belong to a group is to alleviate the anxiety associated with being aware of one's own mortality".

Theoretical overview

There are several theories explaining the motivation that drives people to join groups (Hogg, Hohman, and Rivera, 2008). One of the most influential ones is the terror management theory.

According to the terror management theory (TMT), group membership is motivated by fear of death. Greenberg, Pyszczynski, and Solomon (1986) suggested that there is a unique cognitive ability that sets humans apart from animals—our ability to be aware of our own mortality. This may lead to an experience of paralysing terror. To protect themselves from these terrifying feelings, humans have cultural worldviews that promise them either literal or symbolic immortality (provided that we accept the cultural norms). Many religions have a concept of an afterlife. Outside of religion, we have created meanings around contributing to the greater idea which will live after we die: making a scientific discovery, writing a book that will influence generations, bringing up children to continue ourselves through them, and so on. All of this requires being connected to a cultural institution, belonging to a group.

The theory has generated a great deal of research.

Source 1

Das et al. (2009) sought to provide empirical evidence for TMT by investigating how news reports about terrorist attacks increase prejudice against out-groups.

According to TMT, exposure to news about a terrorist attack will remind the person about their own mortality, triggering feelings of terror. This will lead the person to reinforce their group membership and by extension increase prejudice towards out-groups (in-group favouritism and out-group discrimination). The theory predicts that this will be especially obvious in people with a low self-esteem.

The sample included 100 White European volunteers from the Netherlands. None of the participants were Muslim. The sample was randomly split into two conditions: the experimental group viewed news about terrorist attacks and the control group viewed news about the Olympic Games.

After watching the content, participants completed a word fragment task containing 17 death-related items (e.g., in Dutch, the word fragment doo_ can be completed as dood = dead, doos = box, or doof = deaf). This was used to

Exam tip

Keep in mind that you will never be asked a direct question about motivation. HL extensions are assessed in Paper 3, which is based on a stimulus material that will be provided to you. The studies in the stimulus material will be related to motivation, but the only place where you will be expected to use your "background knowledge" about motivation is Question 4, in which you can support your analysis of the given claim with evidence from the stimulus material as well as your own knowledge.

Motivation is closely linked to the ideas of causality in human behaviour. This means that you already have a lot of background knowledge on the topic, although it was not always explicitly referred to as "motivation". For example, behaviourism may be viewed as a theory of motivation in a broad sense. Self-determination theory is one specific additional content point that is mentioned in the IB DP Psychology Guide, so you should be familiar with it.

However, terror management theory is not a requirement. It is an overarching theoretical framework for the three study sources that have been selected here as an example.

measure the extent to which thoughts about death had been "activated" in their mind. They also completed a self-report questionnaire that was designed to measure prejudice against Arabs.

The experimental group made more death-related word completions than participants in the control group (which means that the experimental manipulation worked) but there was no correlation between death-related thoughts and prejudice towards others. This does not support the theory.

However, during the data collection part of the study, there was a naturally occurring variable: on 2 November 2004, the well-known Dutch filmmaker Theo Van Gogh was murdered by an alleged Islamic extremist. This allowed researchers to investigate the differences in results between those who took part in the study before the murder and after.

Results revealed a significant interaction between thoughts about death and time of participation (before and after the murder). This effect was, as the authors themselves put it, "marginally significant": $t(93) = 1.70$, $p < 0.094$. Specifically, before the murder death-related thoughts were not related to prejudiced attitudes towards Arabs but after the murder the more death-related thoughts participants had, the higher their scores on the measure of prejudice.

Overall, exposure to terrorism news increases the incidence of death-related thoughts. In turn, death-related thoughts lead to more prejudice, but only after a real-life incident. Researchers' interpretation of this is that terrorism news is likely to increase prejudiced attitudes only when it is "psychologically close"—for example, happening in your own country.

Source 2

One of the differences between TMT and social identity theory (SIT) in how they explain prejudice is that social categorization is crucial in SIT but not important in TMT. Suppose a terrorist attack was carried out by members of a radical social group. Social identity theory predicts that this will lead to increased prejudice towards members of that radical group. However, terror management theory predicts that this will lead to increased prejudice towards *any* out-group.

Another study was published in the same paper, Das et al. (2009). It tested the proposition of TMT that thoughts about death increase a person's prejudice against any out-group, not just the group portrayed in the news. With this in mind, the hypothesis was that terrorism news increases prejudice against Arabs in European participants, but also equally increases prejudice against Europeans in Muslim participants.

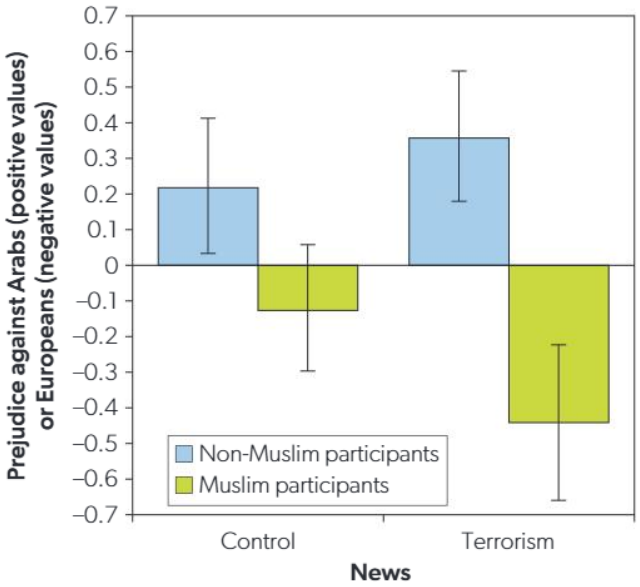
The sample included two sub-groups: Muslim participants (from Morocco, Egypt, Pakistan, Iran, and Iraq) and non-Muslim participants (from the Netherlands, the UK, Belgium, and Germany). The procedure was similar to the previous study, but prejudice towards the out-group was measured using an implicit association test (IAT).

Activity

Refresh your knowledge on what IATs are and how they work. We considered them in more detail in Unit 2.9: Potential for improving cognitive processes.

ATL Research, Self-management

The expectation that thoughts about death would increase prejudice towards any out-group was confirmed: the result was equally significant for Muslim and non-Muslim participants. As seen from Figure 6.1, being exposed to news about terrorism increased implicit prejudice against Arabs among non-Muslim participants, but it also increased implicit prejudice against Europeans among Muslim participants. Thoughts about death cause people to reinforce their group membership no matter what group they belong to.



▲ Figure 6.1 Effects of terrorism news on implicit prejudice. Source: Das et al. (2009)

Source 3

Arndt et al. (2002) predicted that awareness of one's own inevitable death (mortality salience) will lead to increased identification with the in-group when the in-group is framed positively, but disidentification when it is framed negatively.

The study recruited 91 introductory psychology students whose ethnic background was either Hispanic or Anglo-American.

Researchers asked participants to read a "prime"—a news article that described a Hispanic individual in either a positive or a negative light, depending on the condition. The negative news article was about a crime. The positive article was about a charity worker who helped establish social work services in small cities.

After this, participants completed a questionnaire that focused either on their own mortality or a control topic (dental pain).

After this they were asked to indicate their preferences for abstract art that was painted (reportedly) by Hispanic or Anglo-American artists.

Results of the study showed a significant three-way interaction of Prime × Salience × Target ($p < 0.01$). This means:

1. Participants who read the positive prime and whose death-related thoughts were activated evaluated the work of Hispanic artists more highly.
2. Participants who read the negative prime and whose death-related thoughts were activated evaluated the work of Hispanic artists more negatively.

These results fit well into what was predicted.

Note that when death-related thoughts were not activated (in the dental pain condition), these effects were not observed. Therefore, it appears that a negative (or positive) prompt about Hispanic people does not in itself cause negative (or positive) evaluations of Hispanic artists, but when coupled with thoughts about death, it does. All these results also support the initial predictions.

However, there was no difference in these results between Hispanic and Anglo-American participants. Anglo-Americans should not have identified with the Hispanic group in the first place, so why were their judgements about Hispanic artists affected? Researchers explained that positive information about a Hispanic individual portrayed in the prime signalled to the participants that Hispanics and Anglo-Americans shared common values and therefore belonged to a wider in-group of "people who value charity work".

Claim

The leading motivation for people to belong to a group is to alleviate the anxiety associated with being aware of one's own mortality.

Analysis of the claim

Research sources presented here provide some amount of support to the claim.

The most important evidence is that observed patterns of correlations are consistent with the predictions made by TMT: we observe correlations that we are expected to observe if the theory is correct, and we do not observe the ones that are not expected. Additionally, the evidence is strong where the observed patterns of correlations are consistent with TMT but inconsistent with other theories, such as SIT (e.g., in Source 2).

Research studies presented here have used sophisticated experimental designs with several independent variables which allowed researchers to compare various combinations of factors. For example: type of the news article × type of thoughts activated × ethnicity of the artist in Source 3, which is essential to test the theory. Where possible, variables were manipulated by the researcher and the potential confounding variables carefully controlled (e.g., random allocation into conditions and the use of IAT to avoid problems of self-report measures).

All of this supports the claim that thoughts about death play a role in the strength of an individual's sense of belonging to a group. However, support for the "leading motivation" part is less obvious. Results suggest that when thoughts about death are more salient, a sense of belonging to a group is stronger.

However, this does not necessarily support the stronger claim that getting rid of such thoughts is the main, or leading motive. There was no direct comparison with other possible motives.

There are two key weaknesses that span across the data presented in these sources.

First, researchers have accepted some correlations as consistent with the theory when the correlation was almost significant but did not cross the accepted threshold of statistical significance. For example, the "marginally significant" results in Source 1. The accepted boundary of statistical significance is $p < 0.05$. It may be argued that the researchers should not have accepted their result as significant in the first place. This may be a manifestation of a researcher's confirmation bias.

Second, interpretation of results in the sources presented here is post hoc: first researchers obtain the results, then they interpret these results through the lens of the theory. They highlight parts that provide support for the theory and try to provide an explanation for any discrepancy. Again, this makes data analysis vulnerable to confirmation bias. A more rigorous approach would be to plan in advance exactly what comparisons will be made and predict exactly what results are expected, then compare these expectations against actual data (similar to how it is done in pre-registered clinical trials).

An illustration of the second point is the murder in Source 1. It was an unplanned variable. If it had not occurred, the study would have been completed with an insignificant result: it would have to be accepted that terrorism-related news did not cause any change in prejudice. To what extent should we trust the results of a study that seems to have stumbled upon a serendipitous finding that (luckily) seems to provide support for the author's hypothesis? Similarly, the piece of evidence that did not fit in Source 3 was no observed difference between Hispanic and Anglo-American participants. Researchers seemed to provide an explanation, but the explanation is post hoc and we should actually treat this finding as something that contradicts the theory.

Overall, there is certainly evidence that supports the claim, but it would be more appropriate given the existing evidence to make the claim less strong and more operationalized. That is, activation of thoughts related to death increases an individual's self-reported identification with their group, especially when self-esteem is low and the group is viewed positively.

TOK

Post hoc is a Latin phrase, meaning "after this" or "after the event". It is sometimes used to refer to the process of generating hypotheses or explanations after the data has already been observed. Find out how the idea of post hoc explanations is similar to or different from the following:

- post hoc ergo propter hoc (logical fallacy)
- a priori and a posteriori knowledge
- ex post facto
- ad hoc.

6.3 Motivation and mental health

What you will learn in this section

Key learning:

- The role of self-determination (autonomous motivation and supporting autonomy) in treatment.
- Analysis of the claim "autonomous motivation improves the effectiveness of treatment for depression".

Theoretical overview

"Common factors" in psychotherapy are variables that influence the effectiveness of treatment irrespective of the specific therapeutic techniques used during the sessions. It is these factors that have been shown to explain a large percentage of the therapy's effectiveness (Wampold, 2001).

Self-determination theory with its concept of autonomous motivation appears to be especially insightful in terms of suggesting possible common factors that may explain therapy outcomes.

Research in various fields has converged on the idea that when the environment is supportive of autonomy, people become more intrinsically motivated, which in turn leads to various positive outcomes.

Compare the following two statements in response to the therapist's question "What brings you here today?":

1. "My wife thinks that I need to go to therapy, so I am here. She is being ridiculous but she keeps insisting."
2. "I have been experiencing some problems in relationships with my wife lately, and I realized that I may not be the man I used to be. I wonder what I could do to try and restore our relationship."

You could expect better and faster therapeutic progress with the second type of client because that person is self-motivated and takes responsibility for goals and actions.

Source 1

Zuroff et al. (2007) studied the role of autonomous motivation in therapy of depression.

The study compared depressed patients who were being treated with CBT (cognitive behavioural therapy), IPT (interpersonal therapy), and PHT–CM (pharmacotherapy with clinical management). They had been recruited as volunteers for another bigger research programme. It should be noted that some participants terminated the treatment early, also dropping out of the study.

The following measures were used in the study:

1. Interviewer-rated and self-reported scales of depression.
2. A questionnaire to assess autonomous motivation for treatment.
3. A questionnaire to measure the extent to which the therapist was perceived as autonomy-supportive.

Measures of depression were administered before the start of therapy and after its completion (at least 10 sessions). The other measures were administered after the third session.

In the PHT–CM condition, patients were prescribed an antidepressant medication and provided with 10 sessions of "minimally supportive therapy". No special therapeutic techniques were used during these sessions.

The following results were obtained in the study:

- Autonomous motivation was a significant predictor of success: the chance of remission after 10 sessions for patients with high autonomous motivation was twice as high as the chance of remission for patients with average autonomous motivation, and four times higher than patients with a low level of it.
- Autonomous motivation itself was significantly predicted by the perceived autonomy-supportiveness of therapists.
- The positive effect of autonomous motivation did not show any significant difference between the three treatments (CBT, IPT, and PHT–CM).

It is especially interesting to note that all treatment conditions in this study were rigorously "manualized"—that is, all interactions with the client strictly followed the prescriptions of a clinical manual. This shows that even when the therapy follows a step-by-step manual, the relationship between the therapist and the client may vary, defined both by the therapist's behaviour and the client's expectations.

Activity

Design your own method or technique to enhance the client's autonomous motivation during therapy for depression. If you were the therapist, what exactly would you do or say to the client to achieve this?

ATL Communication, Thinking

Source 2

Wilhelmsen et al. (2013) conducted a qualitative study of the motivation to continue therapy among participants in a computerized CBT programme.

Participants in the study were 14 individuals in Norway with mild to moderate symptoms of depression, volunteers in a support programme. Only three failed to complete the full treatment course. The study used a programme called MoodGYM, which was a web application with five interactive modules. Patients completed the modules at home, but had brief face-to-face consultations with the therapist in between modules. During the consultation, the therapist introduced the next module and discussed the patient's motivation to continue.

Semi-structured interviews lasted between 40 and 70 minutes, were recorded digitally and transcribed. A range of open-ended questions were used to evoke participants' narratives.

Inductive content analysis was applied to the interview transcripts. Firstly, the transcripts were divided into "meaning units". They were then further condensed to identify emerging themes and sub-themes (see Table 6.1).

What participants say		What participants talk about	What participants refer to
Meaning unit:	Condensation:	Subtheme:	Theme:
1: I would have left the program anyway, because the modules were not suitable for me.	Dropped out because the modules were not suitable.	Did not identify with the program.	Identification.
2: No, to me the greatest motivation I suppose was the woman I live with. It was. And for my own part, a wish to have a better everyday life […].	The greatest motivation was his wife and to have a better everyday life.	Partner is important.	Belonging/ important others.

▲ Table 6.1 An example of how themes were derived from meaning units. Source: Wilhelmsen et al. (2013)

The following sub-themes relevant to therapy motivation were identified:

1. Gaining control of one's life. Some participants felt that they gained more control over their own life by taking part in this programme, and they experienced it as motivating: "I was actually doing something instead of sitting, waiting, and feeling distressed."
2. Hope for recovery. Some participants expected that participation in the programme would help them recover, and this expectation was also experienced as motivating.
3. Competence and autonomy. It was suggested by the therapist that participants could skip some parts of some modules as long as the overall programme was still being completed. This flexibility was described as motivating because it provided a sense of autonomy and helped them deal with the time constraints.
4. Belonging. This was achieved by sharing the process with people close to them (such as friends and relatives) and being open with them about their successes and struggles.
5. Recognition and self-identification. The course frequently presented a fictional character whose thoughts and behaviour were analysed, and some participants recognized themselves in these descriptions ("Goodness, this is about me, right?"). This realization was described as motivating.
6. Connectedness and expert feedback. Participants reported that the face-to-face sessions were motivating due to their supportive nature. However, some participants also reported that the sessions were too short.

Overall, the SDT claims that in order for intrinsic motivation to be enhanced, the psychological needs for relatedness, competence, and autonomy need to be satisfied. All three have emerged in participants' responses as factors that motivated them to continue computerized CBT.

Source 3

Keeler et al. (2021) studied the effectiveness of engaging in physical activity with peers to treat depression in university students.

The study was conducted in two groups:

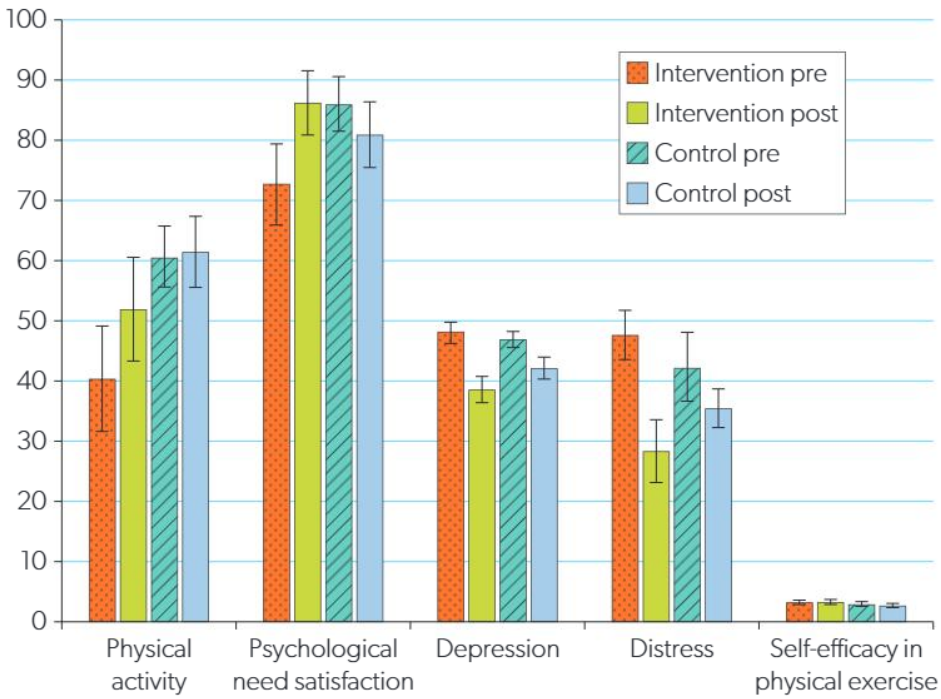
1. The experimental group consisted of 13 college students referred to the programme by a clinician as a supplemental method in the treatment of their mild to moderate depression.

2. The control group included 13 pair-matched individuals recruited from large introductory classes in the same university. Pairs were matched based on self-rated depression, gender, age, and ethnicity.

The peer-assisted physical activity programme offered to participants was designed to meet the three psychological needs postulated by SDT: autonomy, competence, and relatedness. A peer assistant was assigned to each student. The intention was to be physically active as equals rather than act as a trainer and a trainee. Participants could therefore vicariously learn from their peer and receive encouragement. Peers were instructed to offer choices, but avoid giving any direct instructions to create the impression of a workout buddy rather than a personal trainer. The intervention lasted 10 weeks.

Questionnaires were used to measure variables relevant for the study: the level of physical activity, self-efficacy in physical exercise, psychological need satisfaction (with the three scales of autonomy, competence, and relatedness), depression, and distress.

A *t*-test revealed no differences between the two groups at the start of the study. There was no difference between pre-test and post-test measures in the control group. However, the experimental group demonstrated a significant improvement from pre-test to post-test: lower depression and lower distress. Satisfaction of psychological needs in the experimental group improved in the expected direction, but did not reach statistical significance.



▲ Figure 6.2 Results of Keeler et al. (2021)

Claim

Autonomous motivation improves the effectiveness of treatments for depression.

Analysis of the claim

The claim can be read in two ways:

1. Autonomous motivation is a mediating variable in the effects of various other factors on the effectiveness of treatment. Clients whose pre-existing levels of autonomous motivation are high benefit more from various treatment programmes.
2. There is a causal link between autonomous motivation and effectiveness of treatment. If therapy is conducted in an autonomy-supportive way, the client's autonomous motivation will increase and so will the probability of remission.

Additionally, the claim implicitly contains the statement that various types of treatment benefit equally from the client's autonomous motivation.

The sources considered here provide some support to all these statements, but the credibility of evidence varies.

Activity

What is the difference between a cause and a mediating variable? Make two lists:

1. evidence provided in the sources that suggests that autonomous motivation influences (is a cause of) effectiveness of treatment
2. evidence suggesting that autonomous motivation is a mediating variable in the influence of some other factor on effectiveness of treatment.

Read the analysis presented here and see if it overlaps with your list.

Next, turn to your partner and explain the difference between a cause and a mediating variable in less than one minute, giving one example.

ATL Thinking, Self-management, Communication

The study in Source 1 found that clients with a higher level of autonomous motivation have higher rates of remission at the end of the therapy, no matter which therapy was used.

The study also found a correlation between autonomous motivation and the perceived autonomy-supportiveness of therapists. It was interpreted in terms of autonomous motivation being "predicted" by the therapist's autonomy supportiveness. However, we need to remember that both these variables were measured concurrently after the third session, so speaking of prediction or causality is not appropriate in this case.

Overall, the study does support the mediating role of autonomous motivation. However, the causal link between the behaviour of the therapist and the client's autonomous motivation is not supported.

One of the limitations of this study is that some participants terminated the treatment early, which leads to a situation in which the sample of those who completed the study was not exactly the same as those who began it. It remains to be seen if the study has been significantly affected by experimental mortality.

In Source 2, researchers used inductive content analysis without restricting themselves with a theoretical framework, but the resulting themes identified from the transcripts were later related to the three psychological needs proposed by the SDT. There seems to be considerable overlap between the themes that emerged from the transcripts and these theoretical needs. This may mean that the way clients in a computerized therapy interpret their lived experiences matches the theoretical framework of SDT.

It would be more informative if there was some indication of frequency or prominence of the identified themes in participants' responses. It is also important in inductive content analysis that several independent researchers perform the theme extraction and compare results, to ensure credibility.

One of the limitations of this study is that participants who choose to take part in a trial are self-selected and therefore motivated by definition. On a related note, out of 14 participants only three did not complete the course of treatment. This did not give enough representation to the voice of non-completers and did not allow a full investigation of how they experienced therapy.

TOK

To what extent can quantitative and qualitative research be integrated within the same research study? Or is it always two separate studies with subsequent comparison of results?

In the study described in Source 3, the control group was selected from a pre-existing group using a matching procedure. This cannot be considered random allocation into conditions. Therefore, we cannot say with certainty that the independent variable was manipulated by the researcher, making it more appropriate to identify the method of the study as a quasi-experiment.

We can see from Figure 6.2 that both distress and depression decreased in both groups, but the decrease in the experimental group was larger and statistically significant. The chart is a bit misleading for self-efficacy in physical exercise because apparently a different scale was used to measure this variable.

However, a larger problem is the fact that the experimental group took part in quite a complex peer-assisted physical activity programme, whereas the control group carried on with their regular studies. Although the programme was designed based on the principles of SDT, there were many differences between the two groups apart from autonomous motivation. Something about the programme worked, but was it increased autonomous motivation or something else?

This represents a problem that is common for motivation research. It is difficult to manipulate autonomous motivation alone without also affecting something else.

Overall, there seems to be ample evidence in support of self-determination theory suggesting that having autonomous motivation is beneficial for outcomes of depression treatment. Evidence in support of the idea that the client's autonomous motivation may be influenced by the therapist's actions is, as presented in the three sources above, more limited and inconclusive.

6.4 Motivation and human development

What you will learn in this section

Key learning:

- The role of intrinsic motivation in human development.
- Three factors that support intrinsic motivation: optimal challenge, perceived competence, support for autonomous activity.
- Stages of development of internalization: external regulation, introjected regulation, regulation through identification, integrated self-regulation.
- Analysis of the claim "human development is intrinsically motivated".

Key term: internalization

Theoretical overview

Intrinsically motivated development

The relationship between motivation and development is described in Deci and Ryan's self-determination theory quite simply and powerfully: "development is intrinsically motivated" (Deci and Ryan, 1985, p. 115). This means that external stimuli such as punishments and reinforcements are not necessary for it; development will occur even without such stimuli, out of the organism's need to be competent and self-determined.

From this point of view, the correct way to foster healthy development would be to create a supportive environment that provides rich opportunities for choice and self-determination and does not prevent the natural intrinsic tendencies. Perhaps this is why the IB allows you to choose subjects from the six subject groups and choose the level of study within each subject; perhaps this is also why you get several questions to choose from in some papers in IB Psychology, as well as such open-ended components of assessment as a research proposal in the Internal Assessment.

If development is driven by intrinsic motivation, then anything that facilitates intrinsic motivation should also facilitate development. According to SDT, such factors are:

- optimal challenge
- perceived competence
- support for autonomous activity.

Socialization and internalization

As children grow up, they face the conflict between the urge to follow intrinsic motivation and do what they feel like doing and society's mandate to engage in activities that the child does not necessarily find enjoyable. Take school as an example: not every child finds it enjoyable, but there is a social obligation to study.

It is clear that socialization requires extrinsic influences. It is probably the hope of every caretaker that over time their child will internalize the activity, therefore accepting an internal locus of causality, but it starts on the outside.

Overall development of **internalization**, according to Deci and Ryan, goes through the following four stages.

1. External regulation—this is when behaviour is regulated totally by external stimuli (e.g., "I am cleaning up because Mom told me I can only have breakfast once the table is clean").
2. Introjected regulation—this is simply an external rule that has been adopted non-critically (e.g., "I am cleaning up because good children don't leave their room dirty").
3. Regulation through identification—this is when the child accepts the regulation as their own (e.g., "I am cleaning up because I'm good and I like my room to be tidy").
4. Integrated self-regulation—this is when the individual makes flexible choices depending on the situation (e.g., "I typically keep my house tidy because I value being clean and organized, but sometimes when my friends are invited for a party, I find it acceptable to relax my standards a little bit") (Deci and Ryan, 1985; Chandler and Connell, 1987).

Source 1

Chandler and Connell (1987) conducted a series of interviews to investigate if extrinsically motivated behaviours tend to become more internalized in the process of a child's development.

First, they recruited a sample of 45 children aged 5–13 years and asked them to report on activities they enjoyed doing and activities they did not enjoy. This allowed researchers to identify lists of most commonly liked and disliked activities. They also added some "mastery behaviours" to the list—see Table 6.2.

	Liked behaviours	Disliked behaviours
All children	• Building something with toys • Playing a ball game • Playing with a friend • Reading a book	• Tidying up your room • Brushing your teeth • Going to bed on time • Doing what Mom asks without arguing
Ages 7–13	• Playing a board game • Going skating	• Doing homework • Coming home from playing on time
Ages 5–7	• Playing an easy board game • Putting on all your clothes by yourself	• Leaving Mom's things alone • Leaving the poisons under the kitchen sink alone

▲ Table 6.2 Behaviours included in the child interviews. Source: Chandler and Connell, 1987

The same sample of children was then asked about the reasons they engaged in each of these behaviours. The format of the question was "When you do … why do you do it?" and responses were recorded verbatim. Following this, researchers performed content analysis. This exercise resulted in identifying a set of eight categories that researchers further combined into three generic groups of "extrinsic", "intrinsic", and "internalized":

Intrinsic:

1. pursuit of challenge or mastery of skill (e.g., "I play because I want to learn to hit better")
2. pleasure or interest in the activity itself (e.g., "because building is fun").

Extrinsic:

1. anticipation of reward or approval (e.g., " … so my dad will like me")
2. avoidance of disapproval (e.g., " … so my mom doesn't yell")
3. following an explicit rule (e.g., "I am supposed to clean my room").

Internalized:

1. achievement of a self-determined goal (e.g., "I practise so that I can make the school team")
2. avoidance of self-determined negative consequences (e.g., "If I don't clean my room I'll lose the pieces of my models")
3. to do something nice for someone else (e.g., "I want to make Mom smile") (Chandler and Connell, 1987).

Finally, researchers recruited the main sample of 166 children of the same age (5–13 years). They were asked the "why" question about each of the behaviours on the list and their responses were coded into one of three categories using the scheme. The relative proportion of extrinsic, intrinsic, and internalized motivations was tallied separately for "liked" behaviours and "disliked" behaviours, and results were compared across age groups.

Overall, the results of the study demonstrated that intrinsic motivation was predominant when children were describing liked behaviours. However, extrinsic and internalized motivation dominated their descriptions of the reasons why they engage in disliked behaviours (see Table 6.3). This is expected. All differences here are statistically significant.

	Liked behaviours	Disliked behaviours	*t*
Intrinsic	0.67	0.05	30.57*
Extrinsic	0.11	0.36	15.04*
Internalized	0.19	0.49	26.65*

▲ Table 6.3 Differences in motivation between liked and disliked behaviours (proportions do not total to 100 because some responses could not be related to any of the three categories). Source: Chandler and Connell, 1987, Table 4
*$p < 0.001$

When researchers split the sample into four age groups and compared responses, they found the following clear trends:

- The proportion of internalized motivations significantly increased with age, but only for disliked behaviours.
- All other proportions remained the same. For example, there was no change with age in the motivation to perform liked behaviours.

The authors concluded that children's motivation for disliked behaviours gradually becomes more internalized with the course of time.

Discussion

How do these findings apply to your everyday school activities? Think about how your lessons are normally organized.

ATL Thinking, Self-management

Source 2

Zuckerman et al. (1978) investigated whether it is true that people are more intrinsically motivated for activities in which they have greater self-determination.

Participants in the study were 80 university students that comprised 40 randomly formed pairs. There were two conditions:

1. Participants in the "choice" condition had puzzles of six different configurations. They were given 30 minutes to work on them and asked to choose any three of the six. They could decide themselves which puzzles to choose and how much time to spend working on each of them.
2. Participants in the "no-choice" condition were told which three puzzles they will work on and how much time they will spend on it. This was determined based on what was chosen by the partner in their pair. In other words, they worked on the same puzzles and for the same duration as their partner, the only difference being that they were not the ones to choose this configuration.

On another table next to the participant, there were recent issues of three magazines and also two extra puzzle configurations that were not used during the experimental procedure.

In order to measure intrinsic motivation, the experimenter left for eight minutes after the 30 minutes of puzzle solving was completed. The pretext was that the experimenter had to go and get the appropriate questionnaire. The participant was told that they were free to do what they wanted during that time. The experimenter took the six puzzles that were used during the main phase of the study, leaving the participant only with the two extra puzzles that had not been used.

The measure of intrinsic motivation was the amount of time out of eight minutes that the participant chose to spend working on the extra puzzles.

Results of the study revealed that in this choice condition, participants spent on average 260 seconds (4 minutes 20 seconds) working with the puzzles, and participants in the no-choice condition spent on average 165 seconds (2 minutes 45 seconds). This mean difference of 1 minute 35 seconds was statistically significant: $t(39) = 2.15$, $p < 0.05$. Overall, the conclusion of the study was that participants who had some choice over what they do and for how much time were more intrinsically motivated for this task than participants who performed exactly the same activities, but had no control over them.

Discussion

How would you organize a typical school day based on these principles? What changes would you suggest in your school? Can you think of any counter arguments to this approach?

ATL Communication, Self-management

Claim

Human development is intrinsically motivated.

Analysis of the claim

Implicit in this claim is the idea that intrinsic motivation in human development is more important than extrinsic motivation (e.g., operant conditioning).

The sources presented here provide some support to the claim. However, we need to be careful about how far we can generalize the findings of these studies.

The study described in Source 1 provided support to the idea that children's motivation for disliked activities becomes more internalized with age. According to SDT, internalized behaviour is not intrinsic (i.e., it is not motivating in itself), but it is self-determined (i.e., it is determined by internal rather than external factors). The finding that children's behaviour becomes more internalized certainly goes well with the SDT, but it does not necessarily suggest that development is "intrinsically motivated", or that development is not extrinsically motivated. If anything, this finding actually highlights that an activity that is not intrinsically motivated may initially be maintained by external regulation, but over time become internalized. The finding emphasizes the importance of internalization in the process of development, but it is difficult to use it as a basis for any more generic conclusions.

The study described in Source 2 demonstrated that, when we have some amount of choice and autonomy about how to deal with a task, we become more intrinsically motivated for that task. The implication of this finding in developmental psychology is that providing children with a healthy amount of choice and autonomy can enhance their intrinsic motivation to engage in activities. Continuing this line of thought, we assume that higher intrinsic motivation will result in better developmental outcomes because the child will engage with more educational materials. However, this evidence is not sufficient to support the more generic and stronger claim that human development is intrinsically motivated. There was no comparison between the impact of intrinsic and extrinsic motivation on variables of human development in this study.

The strength of the first study (Source 1) could be the way it combined qualitative and quantitative data. Although data collection was done via a qualitative method (interview), data analysis and hypothesis testing were performed quantitatively, by applying an inferential statistical test. Looking at the results of this quantitative analysis, it is interesting that there exist non-zero frequencies of behaviours that do not neatly fit into the framework of SDT. For example, 5% of disliked activities were reported by children as intrinsically motivated (how is this possible?). Similarly, 11% of liked activities were described as extrinsically motivated. These could be random fluctuations caused by factors unrelated to the theory (e.g., the way transcripts were coded), but it is also possible that they reflect important nuances overlooked by the theory.

Another strength of the study is the rigorously organized procedure: for example, a separate sample of children was used to obtain the list of activities and derive coding categories.

The strength of the second study (Source 2) is the rigorous control over potential confounding variables and clear operationalization of constructs. The result was statistically significant with quite a visible effect size: the amount of time spent on the extra puzzles in the choice condition was 58% higher than in the no-choice condition.

Overall, both studies are well-controlled with sufficient measures in place to ensure validity and credibility of findings. The findings provide convincing support for the hypotheses that were intended to be tested. However, the claim "human development is intrinsically motivated" requires a much larger number of studies to be fully supported. In fact, the entire body of research conducted in the framework of SDT and relevant to parameters of human development serves that purpose.

Exam tip

The analyses of the claim presented here and those throughout this chapter are examples of how you could approach similar tasks in an exam situation. We have presented a variety of study types and a variety of possible arguments regarding the extent to which the sources (as well as background knowledge) support the given claim.

Although it cannot be predicted what claim or what sources you will receive on the day of the exam, one preparation strategy could be keeping a list of the "typical" judgements about the claim. For example, here we come to the conclusion that the sources provide support for the claim, but this support cannot be considered sufficient because the claim is very broad. What other "typical" conclusions about the claim can you find throughout this chapter?

6.5 Motivation in learning and cognition

What you will learn in this section

Key learning:

- How does motivation interact with cognitive ability in their influence on task performance?
- Analysis of the claim "Extrinsic motivation, intrinsic motivation, and cognitive ability interact to determine a person's performance on cognitive tasks".

Theoretical overview

Motivation and cognitive ability are two basic determinants of performance in learning and on a job. Large bodies of research have been accumulated about these factors separately. Unsurprisingly, cognitive ability has been shown to be a significant predictor of academic and job success. The same is true for non-cognitive, motivational factors. However, research investigating the two factors at the same time in their interaction with each other is limited.

For example, it is not clear whether the effects of motivation and cognitive ability on performance are independent of each other. It is possible that they interact in some way. For example, that at low levels of motivation cognitive ability does not influence performance but at high levels of motivation it does. One theoretical link between motivation and cognitive ability is attentional resources: attention is limited, and motivation may affect how we allocate this limited resource to various cognitive tasks. If resources are not allocated, cognitive ability will not make much of a difference in an individual's performance.

In any case, clarifying how exactly motivation and cognitive ability are related to each other in the context of learning has both theoretical and practical significance.

Source 1

Kanfer and Ackerman (1989) conducted a study of motivation and cognitive abilities in a field experiment with a sample of 322 US Air Force trainees. They used an Air Traffic Controller (ATC) simulation. This was a challenging task in which participants had to regulate the landing routines of a large number of planes while considering the direction of landing, queue sequence, constantly changing weather conditions, amount of fuel remaining, and so on. If fuel reached zero, the plane crashed.

Participants received points for each plane successfully landed. Points were deducted for each technical error or for a crashed plane.

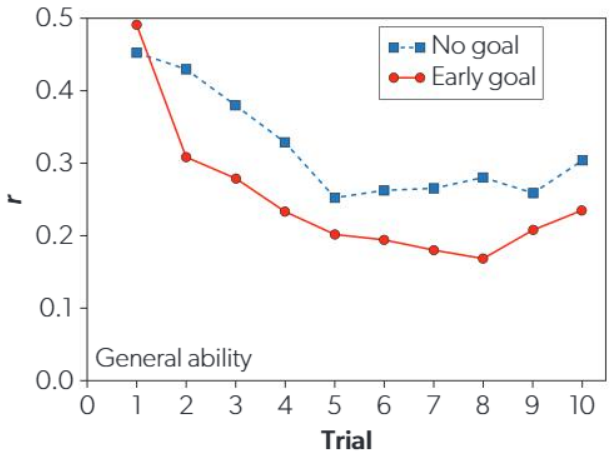
They were split into two conditions:

1. No goal—participants had one practice trial and nine 10-minute task trials (10 overall). They were instructed to "do your best".

2. Early goal—participants were given a specific and challenging goal for trials 2, 3, and 4. They were also given (during those three trials) a periodic opportunity to check their progress relative to the goal. For example: "Based on your current performance, you will attain 80% of your goal". For the remaining trials, these participants were instructed to "do your best", similar to the other condition.

Researchers also measured general mental ability obtained from an IQ test.

Results showed that at trial 1 (before the experimental manipulation), the correlation between ability and task performance was identical in both groups. Following that, the correlation started dropping, perhaps as a reflection of the fact that task performance depended on other factors apart from cognitive ability (such as motivation or the ability to keep focus for a long time). Notably, the drop was more significant in the early-goal group. The correlation continued to weaken after trial 4 (see Figure 6.3).



▲ **Figure 6.3** Correlations between general cognitive ability and performance on the cognitive task in the two groups (no goal and early goal). The vertical axis shows the correlation coefficient (*r*) between general cognitive ability and task performance during a specific trial.

Source 2

Robinson et al. (2012) studied the effects of intrinsic and extrinsic motivation on attention and memory. The sample in this study included 40 male subjects in the UK (aged 18–35 years).

One of the main measures used in the study was the Newcastle Spatial Memory Test (NSMT). This is a computer task in which participants are presented with up to 12 cups and told that the computer has hidden a marble under one of them. Participants have to turn the cups over to find the marble. Once the marble is found, the game repeats (there are as many trials as there are cups). The computer never hides a marble twice under the same cup. Participants make an error if:

- they turn over a cup that they have already checked on the same trial (WSE = within-search error),
- they turn over a cup under which they had already found a marble during one of the previous rounds (BSE = between-search error).

Participants were also given a questionnaire to measure their intrinsic motivation for the task.

Activity

The studies presented in this section have quite complicated designs. It is important to fully understand how the study was organized in order to make relevant arguments about the sources of evidence. Try drawing a schematic representation of each study.

For example, how many variables were there in the study in Source 1? What was the dependent variable and the independent variable? How many times was each of the variables measured? What was compared?

Draw such schema for each of the three studies presented here, exchange with a partner, and give each other feedback.

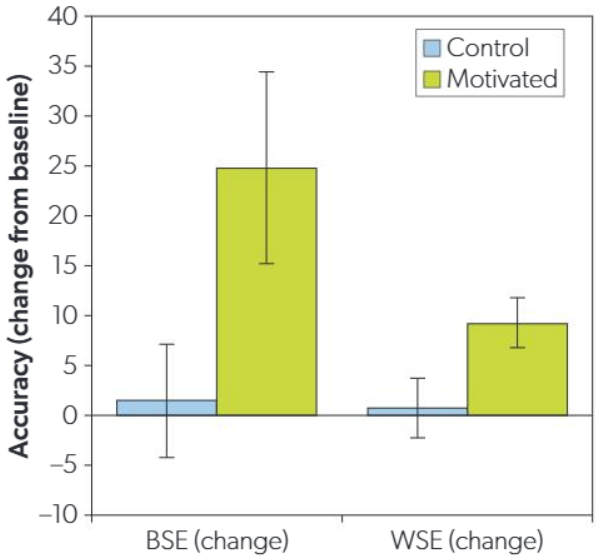
ATL Thinking, Self-management

Participants completed this task twice on the same day: at 1.40pm (the baseline session) and at 3.45pm (experimental session). For the experimental session (3.45pm), they were randomly allocated into one of two conditions:

1. The non-motivated group received a fixed payment.
2. The motivated group began with a certain reward and lost £0.07 for every error they made.

Analysis of data revealed the following results:

1. Performance on the task improved significantly between sessions in the motivated group ($t(16)=2.58$, $p = 0.02$) but not so much in the control group.
2. The number of errors decreased in both groups, but the decrease was significantly larger in the motivated group.



▲ **Figure 6.4** Change in working memory from morning to afternoon session. Higher values indicate greater improvement at session 2.

Additionally, researchers found a significant correlation between the measure of intrinsic motivation (questionnaire) and task performance: the more intrinsically motivated participants were at baseline, the fewer errors they made on the marble search.

Discussion

Suppose student A is naturally talented, but does not put much effort into studies, whereas student B consistently puts in a lot of effort but lacks the baseline ability. In your own experience, how do their chances of being academically successful compare to each other? Does it differ depending on their age?

ATL Communication, Thinking

Source 3

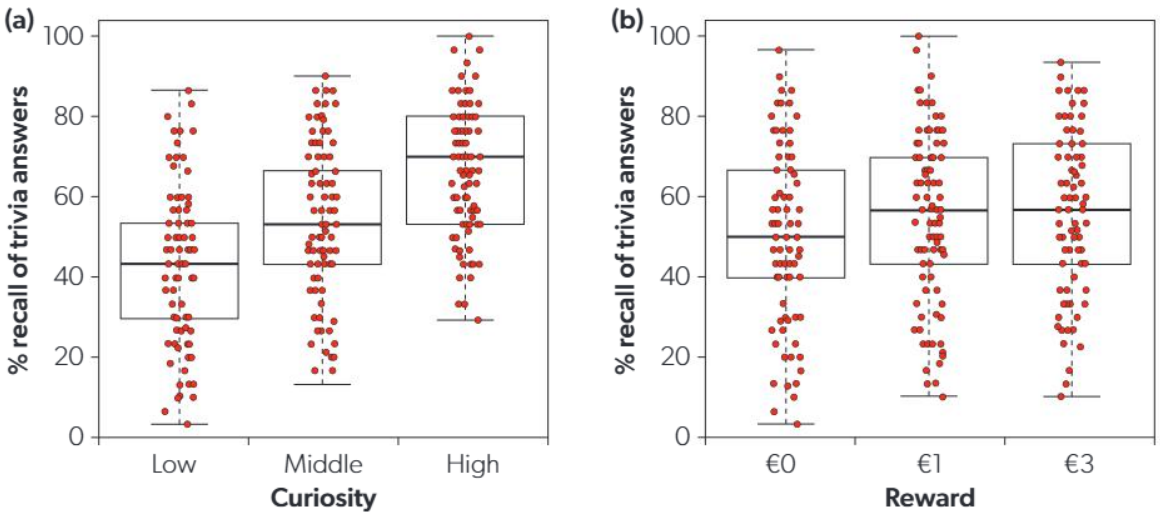
Duan et al. (2020) studied the effects of intrinsic and extrinsic motivation on memory formation. Participants were 35 young Dutch people. They were paid €64 plus a bonus depending on their performance on the memory task.

The study consisted of three stages:

1. Screening: participants read 770 trivia questions and rated each question on: (a) whether they already knew the answer, (b) their curiosity for getting to know the answer. Based on screening, each participant ended up having an individually selected set of 270 trivia questions that they did not know the answer to, 90 for each of the three levels of curiosity.

2. Study: participants were presented with 270 individually chosen questions together with the corresponding answers. Each question was presented for 4 seconds together with the monetary reward for this question (€0, €1, or €3), then the answer was presented for 2.5 seconds. During the whole time, participants were in a brain scanner and their brain activity was being measured.
3. Test: a day after the study phase, participants were given a recall test for the studied 270 trivia questions. Questions were presented one by one in the same order as they were studied, and participants had to type in the response without a time limit. Participants then received a bonus based on their performance and were debriefed.

Results clearly showed the effect of curiosity (intrinsic motivation) on memory: low curiosity = 43% correct responses, medium curiosity = 54%, high curiosity = 67% ($p < 0.001$). The effect of anticipated monetary reward (extrinsic motivation) on memory was also significant, although not as prominent: low = 52% correct responses, medium = 55%, high = 57%. However, the interaction of these two factors was not significant, suggesting that their effect on memory performance is additive.



▲ Figure 6.5 Curiosity, reward, and per cent of correctly recalled trivia answers

Additionally, results of brain scans revealed that effects of curiosity and monetary rewards were regulated by distinctly different brain regions. Higher levels of curiosity were associated with more activation in the middle temporal gyrus and inferior parietal lobule. Anticipated monetary reward was associated with activation in the dopaminergic pathway (the brain's "reward system"—we came across this system in Unit 5.7, in Fisher, Aron, and Brown's study of romantic love on page 388).

Claim

Extrinsic motivation, intrinsic motivation, and cognitive ability interact to determine a person's performance on cognitive tasks.

Analysis of the claim

The idea of interaction between two factors in their influence on behaviour implies that the influence of factor A on behaviour depends on the level of factor B.

Activity

We already encountered several examples of statistical interaction in this book—see Unit 4.6: Environmental factors in mental disorders. Can you graphically represent all the possible combinations of the three effects? For example:

- no main effects but a significant interaction effect
- no interaction effect but significant main effects
- one significant main effect and a significant interaction effect.

ATL Research, Communication

For example, this would be the case if the influence of A on behaviour is positive when B is low but negative when B is high. Statistically, any combination of the following three effects may take place:

- the main effect of A on behaviour
- the main effect of B on behaviour
- the effect of the interaction between A and B.

If both A and B affect behaviour but there is no interaction between them, such influences are called additive. Then A and B contribute to that behaviour independently.

The claim therefore suggests that the effect of cognitive ability on cognitive task performance will be different depending on the level of motivation.

With this in mind, the sources presented in this unit contain the following findings:

- The correlation between cognitive ability and performance on a cognitive task is higher when the level of motivation is low and lower when the level of motivation is high (Source 1). This could suggest that motivation and cognitive ability interact. However, the study used correlation coefficients and the effects of interaction were not tested directly. Only one type of motivation was investigated—extrinsic motivation of obtaining scores.
- Source 2 investigated the correlation between motivation (extrinsic and intrinsic) and performance on a memory task. Results showed that both extrinsic and intrinsic motivation can improve task performance. Interaction between the two types of motivation was not investigated and neither was the interaction between motivation and cognitive ability. In fact, there was no measure of cognitive ability. Therefore, these findings are not relevant to the claim.
- Source 3 directly tested effects of the interaction between intrinsic and extrinsic motivation (curiosity and reward) and found them to be non-significant, although there were significant main effects of the two factors separately. To corroborate this, it was shown that the two types of motivation are associated with distinctly different patterns of brain activation. This suggests that there is no interaction between intrinsic and extrinsic motivation and that their effects on cognitive task performance are additive. Cognitive ability was not measured in this study.

Overall, there is little doubt that the separate main effects of intrinsic motivation, extrinsic motivation, and cognitive ability on cognitive task performance are significant. They may contribute to task performance in an additive way. However, evidence of interaction between any two of these three factors is limited.

Sources 2 and 3 provide reliable evidence that the main effects exist. Source 2 shows that the addition of extrinsic motivation results in around a 25% reduction in the number of between-search errors. In Source 3, the box-and-whisker plots show a clear trend of improvement in the number of correct responses on the recall task depending on curiosity, especially with the transition from medium level to high level of curiosity.

In the sources presented here, only Source 1 suggests that there could be an interaction between motivation and cognitive ability in their effect on cognitive task performance. Indeed, when participants are given an early challenging goal, the correlation between their cognitive ability and task performance weakens and continues weakening. However, two points must be raised:

1. Although a difference in the strength of the correlation was observed between the two groups, the source presents no evidence that this difference is statistically significant.
2. A weaker correlation does not necessarily imply interaction of factors. An additive effect of motivation and cognition could also produce a similar result. In the no-goal condition, performance on the Air Traffic Controller task depends only on the participant's natural ability. In the early-goal condition, performance will also depend on how much effort participants put into learning the task. Cognitive ability alone will no longer be sufficient to explain the observed variation in task performance, hence weaker correlation.

6.6 Culture

What you will learn in this section

Key learning:

The role of culture in human life:

- Sociocultural factors.
- Cultural differences.
- Cultural dimensions.
- Acculturation.

Typical challenges of research in this area:

- Limited samples.
- Using country of origin as an approximation of culture.
- Effects of culture on cognition are not static.
- The problem of deciding whether or not a cognitive process is universal.

Key term: cultural dimensions

Overview

Culture is what makes us human. It plays a fundamental role in human cognition and behaviour. Some scientists even claim that it is impossible to understand any single phenomenon in psychology without considering it in its cultural context.

A clear illustration of the role culture plays in our lives are case studies of feral children—individuals who have been brought up with no exposure to human culture. Feral children are clearly different, both cognitively and socially. Moreover, such cases demonstrate that it is impossible for such children to gain some of the key human skills even if they are brought into human society at a later time in their life.



▲ Figure 6.6 Dina Sanichar

There have been a number of documented journalistic accounts of children raised by animals. One example is six-year-old Dina Sanichar, who was spotted with a pack of wolves by a group of hunters in Uttar Pradesh, India, in 1872. It appeared that the boy had been raised by the wolves. The hunters decided to "help" the boy and bring him back to human civilization. The wolves were killed and the boy was taken to an orphanage. Sanichar walked on all fours, growled like a wolf, ate raw meat, and sharpened his teeth on bones. Although not mute, he could only produce animal sounds. Reportedly, orphanage workers patiently taught him how to live like a human. However, after 28 years in education, Sanichar showed little improvement. He was able to stand and walk and started eating from a plate, but he could never speak beyond producing animal sounds and he continued to sniff the food before eating. The only truly human habit that Sanichar picked up was smoking tobacco. He became a chain smoker and died of tuberculosis at the age of 34 (Mathur, 2018).

There were many instances in this book where we considered the role of social and cultural factors in human cognition, development, mental health, and relationships. See for example the following units:

- Chapter 2: Learning and cognition—Unit 2.7: Cultural factors in cognitive processes
- Chapter 3: Human development—Unit 3.3: Sociocultural factors in development, Unit 3.4: Enculturation of social norms, Unit 3.8: Attachment
- Chapter 4: Health and well-being—Unit 4.5: Cultural differences in mental disorders, Unit 4.7: Prevalence of health problems
- Chapter 5: Human relationships—Unit 5.2: Cultural dimensions, Unit 5.3: Acculturation, Unit 5.9: The role of communication and language, and strategies for improving relationships.

Typical challenges of research in this area

Despite the impressive body of evidence accumulated in research studying the role of culture in human behaviour, there are some fundamental constraints and limitations. Let us now consider some of them.

First, an overwhelming majority of research studies are based on a comparison of North American and East Asian participants. The differences between these two groups theoretically should be the most visible (seeing how they rank opposite on the most prominent **cultural dimension** identified by Geert Hofstede, individualism–collectivism). However, it is important to include other cultural groups and other cultural dimensions of difference in our research efforts in order to avoid one-sided findings and conclusions.

Second, in most studies the country of origin is used as an approximation of culture. However, this is not necessarily the case. Not all people have acquired the norms and values of their culture to the same extent. There may be considerable variation within a country—for example, people belonging to different SES (socioeconomic status) groups, age groups, or even occupations. Finally, with increasing globalization and access to mass media, there is more and more mutual penetration between cultures.

Third, the effects of culture on cognition are not static. They can be weakened or amplified depending on the current mental state of the participant. For example, it has been shown that feeling happy strengthens the culturally specific cognitive style: when in a happy mood, Koreans tend to be even more holistic in their cognitive style, whereas Americans tend to be even more analytical (Ji and Yap, 2016). Priming participants with various cues may also have an effect—for example, bilingual individuals in Hong Kong may use different decision-making strategies about one and the same problem depending on which language the problem is written in, English or Cantonese (Briley, Morris, and Simonson, 2005).

Fourth, the problem with a research area being dominated by studies conducted on a particular population (WEIRD research) is that we may start to think of a cognitive phenomenon as being universal, when in fact it is not. For example, if study after study shows us that children learn better when information is categorized, it is tempting to jump to the conclusion that human memory is categorical in nature. Not enough cross-cultural studies means not enough



Discussion

The acronym WEIRD is commonly used to refer to research conducted in societies that are Western, Educated, Industrialized, Rich, and Democratic.

Do you think it is indeed weird that most of the research in social sciences is conducted using these samples?

ATL Research, Communication

attempts to challenge this idea and truly test its validity. However, if we do conduct cross-cultural research and it reveals that the organization of memory in some societies is not categorical, the conclusion may be that organization of memory is a product of culture rather than a universal phenomenon.

Elaborating on the fourth point, the scientific implications of deciding whether or not a cognitive process is universal are huge.

- If cognitive processes are universal, then we should be seeking an explanation for them based on how the human species evolved, their potential survival value and their determination by brain structure.
- If cognitive processes are a product of culture, our theory would be very different. We would be looking at cultural transmission practices (e.g., education) and trying to understand how cognition may be shaped by a person's life experiences.

The truth is probably somewhere in the middle, with some cognitive processes (or some components) being universal and others produced by culture. Understanding where exactly the boundary lies has enormous scientific and practical implications. Quite simply, if something is a product of culture then we have a degree of control over it; if not, then it cannot be changed.

Finally, in any given research study the variable that we measure will always exhibit some central tendency and some variability. For example, if we investigate at what age children develop an understanding of another person's beliefs (theory of mind), some children in the sample will always pass the test a little later and some a little earlier. If we believe that cognitive processes are universal, we will look at the central tendency (e.g., the mean age) and perhaps explain the variation away as random fluctuations. However, if we believe that cognitive processes are culturally determined, we will focus on the variation itself and try to identify factors that determine why some children develop it earlier and some later. This means that the answer (universal versus a product of culture) does not just emerge from the results of a study. The expected answer will reflect how we have organized the study. If we want the answer to be unbiased, we need to plan both types of research.

6.7 Culture in human relationships

What you will learn in this section

Key learning:

- Are there culturally universal patterns of relationships?
- Analysis of the claim "patterns of formation and maintenance of friendships and romantic relationships depend on the cultural environment".

Key terms: universal patterns, cultural relativism, cultural differences

Theoretical overview

In Chapter 5: Human relationships, we considered the role culture plays in shaping human behaviour. Unit 5.2: Cultural dimensions and Unit 5.3: Acculturation were especially relevant when we looked at group behaviour—that is, how the group may influence the behaviour of an individual in that group.

Another aspect in that unit was interpersonal relationships, which refers to the dynamics of relationships between two or more people. We focused on friendships and romantic relationships as a key example. In Unit 5.9: The role of communication and language, and strategies for improving relationships, we discussed some evidence suggesting that patterns of communication in a relationship may not be universal across cultures.

Here we will continue analysing cross-cultural evidence regarding various other aspects of friendships and romantic relationships, such as its formation (attraction) and maintenance (examples of which are expressions of appreciation and experience of jealousy).

The big overarching question in this area of research is whether or not it is possible to identify culturally **universal patterns**: core behaviours that would be characteristic of all humans, even if differences in more superficial aspects exist. A competing point of view would be that of **cultural relativism**: that patterns of relationships are specific to each culture and deriving universal "laws" is impossible.

Source 1

Karandashev et al. (2020) conducted a cross-cultural comparison of sensory preferences in romantic attraction. Participants in the study were 2740 students (mean age 22 years, approximately two-thirds women). Data was collected in six countries: USA (Middle West, South East, North East, Hawaii), Jamaica, Portugal, France, Russia, and Georgia.

The authors divided the countries participating in this study into three clusters, using the following wording:

- less modernized societies with more pronounced survival values
- more modernized societies with more pronounced values of self-expression
- societies between these two extremes.

Societies in the first cluster included Russia, Georgia, and Jamaica. Societies in the second cluster included regions of the USA. France and Portugal were placed in the third cluster.

Participants completed a survey that included 54 questions about visual, auditory, tactile, kinesthetic, and olfactory preferences regarding romantic partners. They used a five-point Likert scale ranging from "not important" to "most important" to rate such statements as "This person has expressive eyes" and "This person's singing is nice". When answering the questions, they were asked to think about their present partner, past partner, or a desirable future partner.

Results of the study showed that relatively stable biological characteristics of appearance (e.g., body, skin, and smell) are rated as more important in the societies belonging to the first cluster (less modernized, survival values). Conversely, sensory parameters that are more socially determined (such as speaking, hair, dress, dancing) were rated as more important in regions of the second cluster (more modernized, values of self-expression).

Interpreting their results, the authors suggest an evolutionary explanation: in less modernized societies individuals may be more concerned about survival and therefore the importance of those physical parameters that indicate the mate's good health is higher. On the other hand, in more modernized societies physical appearance may be taken more as a sign of social status and personality expression (makeup, expressive speaking, and so on).

Discussion

What do you know about evolutionary explanations in psychology? What is their value? What are their key limitations?

ATL Communication, Self-management

Source 2

Bello et al. (2010) compared methods for expressing appreciation in friendships and romantic relationships between participants in the USA and China.

The sample consisted of 200 university students (79 from the USA and 121 from China). Participants completed a questionnaire in class. The questionnaire instructed participants to think about their romantic partner or best friend and then list the things that they have said or done to express their appreciation for this person.

The lists were analysed independently by two researchers who arrived at a set of categories, then achieved consensus through discussion. These categories were then used to code each of the items mentioned by each of the participants. The inter-rater agreement reached 86 per cent. Once responses were coded this way, researchers used simple frequency counts for each category.

Examples of categories used in this study are given in Table 6.4.

Methods	Examples
Verbal methods	
Expressions of thanks	"I appreciate what you do for me"; "I love you"
Compliments	"You are the best"; "You are so sweet"
Assurances	"I tell him I will live my life with him forever"
Non-verbal methods	
Physical affection	"I hug him"; "I hold her hand"
Favours/help	"I cook for him"; "I help her with school work"
Emotional support	"I listen when she talks"; "I'm there for her"
Sacrifice	"I stay up with him while he is studying"

▲ **Table 6.4** Examples of methods of appreciation expression. Adapted from: Bello et al. (2010), Table 1, p. 298

Results of the study revealed the following patterns:

- The total frequency of appreciation expression in the USA sample was more than double compared to the Chinese sample. These differences existed in both verbal and non-verbal behaviours.
- USA participants relied equally on verbal and non-verbal methods of showing appreciation.
- However, it was more common for Chinese participants to express their appreciation indirectly and non-verbally rather than directly and verbally.
- Although the effect of culture in this study was significant, the effects of gender and the type of relationship (romantic or friendship) were not. This means that the results of the study applied equally to friendships and romantic relationships, and to male and female participants.

Appreciation expression is an important component of communication between partners that contributes to the stability of relationships. Apparently, any conclusions about effective or ineffective patterns of communication in relationship maintenance should consider culturally accepted patterns.

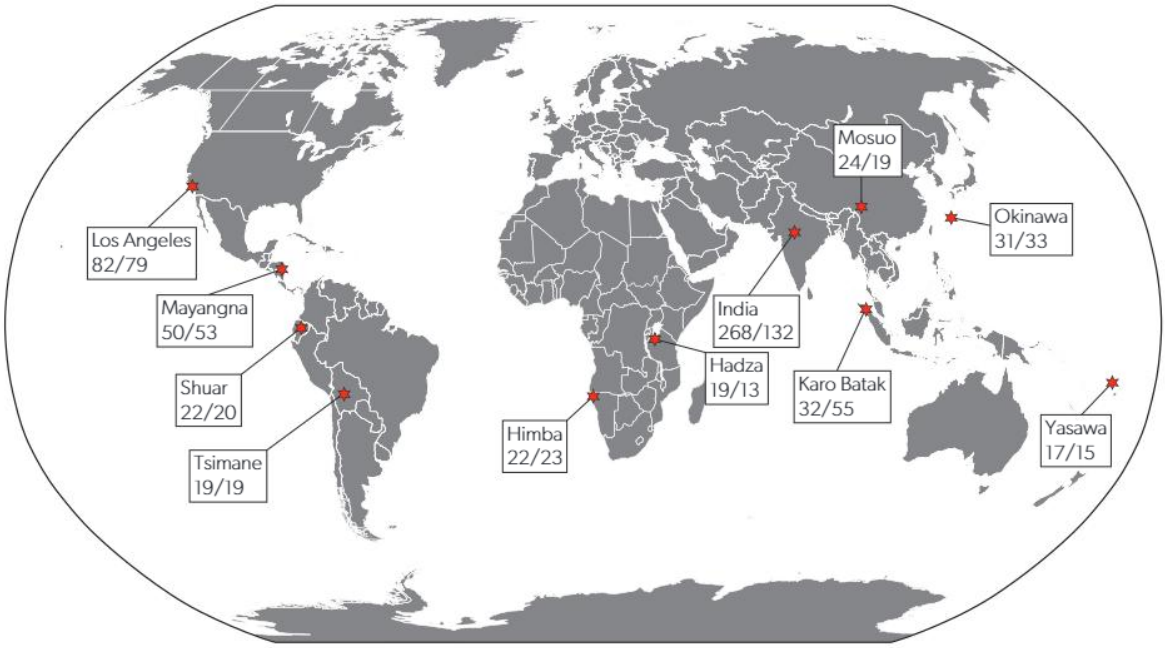
Source 3

Scelza et al. (2020) studied cross-cultural variations in jealousy response.

Hypothetically, the cost of infidelity is different for men and women. Men face what is called paternity uncertainty (they are not sure they are a child's father) and a prospect of taking care of a child that is not biologically theirs. Women face the threat of losing resources to support the child (i.e., the man will allocate money to another family).

If this reasoning is correct, then the following predictions can be made:

- Men will be more upset by physical infidelity, whereas women will be more upset by emotional infidelity.



▲ **Figure 6.7** Geographical location of study participants (the numbers represent sample size: male/female)

- Jealousy response in a given society will depend on paternal investment—that is, the extent to which fathers are expected to invest in their wives' children. There should be higher ratings of jealousy (both physical and emotional, both for men and women) in societies where the level of expected paternal investment is high.

The study was conducted in 11 populations (see Figure 6.7).

Participants were given a survey that consisted of two parts. In the first part, participants were given scenarios of physical and emotional infidelity and asked to rate them using a Likert scale ranging from "very bad" to "very good". In the second part, they were given a forced-choice vignette where they were asked to choose whether physical or emotional infidelity was worse.

Cultural variables such as the expected amount of paternal investment in a given culture were established in prior research using other surveys with representative samples from these societies.

The following finding emerged from the first part of the survey:

- Increased levels of paternal provisioning were associated with more negative ratings of male, but not female infidelity—both physical and emotional.

The following finding emerged from the second part (forced-choice scenario):

- Men choose physical infidelity to be worse than emotional infidelity more frequently, as compared to women. This statement was true for 9 of the 11 populations used in the study. In the two cultures that differed: the Tsimane (Bolivia) where both men and women found physical infidelity to be more upsetting than emotional infidelity, and the Yasawa (Fiji) where women found physical infidelity slightly more upsetting than men.

Chat with AI

Remember that you can use generative AI to help you learn how to effectively analyse sources of evidence in relation to a given claim. Here is a typical prompt starter that you might want to consider:

"I am given a claim [provide the claim] and a source of evidence [provide one of the sources]. The task is to analyse the claim with reference to the source. To what extent does the source of evidence support the given claim? Could you help me achieve an in-depth analysis and help me clearly formulate my conclusions? Do not just give me an answer—ask me questions and provide hints that will help me discover some new aspects on my own. Here is what I have so far: [provide your draft analysis]."

Overall, results of this study suggest that jealousy response is to a large extent universal across cultures. Male participants are more upset with physical infidelity, whereas female participants are more upset with emotional infidelity. Where cross-cultural variation is observed, it can usually be explained away by variations in expected paternal investment.

Claim

Patterns of formation and maintenance of friendships and romantic relationships depend on the cultural environment.

Analysis of the claim

The claim is formulated in a way that does not make it clear which of the following positions is implied:

- All relationship patterns are culturally specific and there are no universal patterns (cultural relativism)

  or

- There are culturally universal patterns of formation and maintenance of relationships, but within that there is also a degree of cross-cultural variation.

Assuming that what is meant is the second, less categorical claim, the sources presented here certainly provide evidence supporting this statement. At the same time, one needs to be aware of the common limitations inherent in this evidence, and be cautious about making unjustified generalizations.

Evidence presented in Source 1 suggests that there exist **cultural differences** in what romantic partners find attractive in each other. However, several considerations limit the validity of this evidence:

- There could be differences between the two opposite "clusters" other than the ones that researchers decided to focus on (i.e., modernized versus less modernized). Certainly countries such as Russia, Georgia, and Jamaica are different from the USA in more than one way.

- The evolutionary explanation for the findings is justified, but it is somewhat speculative. There is nothing in the data that can directly support this explanation.
- The study relies on self-reporting. It is possible that cultural differences manifest in the way participants respond to the survey questions but not so much in the characteristics of attraction itself. For example, it could be that participants in the USA are actually equally attracted to physical characteristics, but they are reluctant to say so on the survey because of the movement against objectification of men and women (which is not so prevalent in other countries).

Evidence in Source 2 suggests that appreciation in friendships and romantic relationships is also expressed differently (in the USA and China). It was a mixed study in which the qualitative component included content analysis of participant response and the quantitative part was the comparison of response frequencies between cultures. The strength of the study was a high level of inter-rater agreement, which suggests that the coding scheme was credible. Another finding that could be considered a strength is that cultural differences outweighed differences between types of relationships (friendship versus romantic) and gender. This increases our confidence that the observed difference indeed exists and cannot be explained away by a simple error of measurement. A limitation, again, is the self-reported nature of this data.

Evidence in Source 3 suggests that there may be cultural differences in the way that people experience and express jealousy. However, it is notable that not all predictions were equally supported by data. For example, the prediction that men should be more upset by physical infidelity than women was only supported in the forced-choice scenario but not the first part of the survey, and not for two countries out of 11.

TOK

How much evidence is enough evidence?

How would the answer to this question depend on the area of knowledge?

Another limitation of the study (and all similar studies) is its correlational nature. When we observe that the way cultures vary along the paternal investment continuum is related to the way cultures vary in jealousy responses (e.g., cultures where expected paternal investment is higher also tend to be the ones where jealousy response is more obvious), we need to remember that these cultures also differ in many other parameters. Any of these other parameters may turn out to be a third variable that influences both jealousy and paternal investment.

Overall, the sources presented here do provide some evidence that supports the claim. However, this evidence is limited in various ways (as is any research in psychology) and the three sources alone are not sufficient to fully support the claim. Additionally, the sources presented here tap into some, but certainly not all, aspects of "patterns of formation and maintenance" of relationships. Broad generalizations from this evidence alone cannot be made.

6.8 Culture and mental health

What you will learn in this section

Key learning:

- Cultural expression of symptoms and somatization.
- Patient variables and clinician variables in diagnosis and treatment.
- Analysis of the claim "cultural factors must be considered when diagnosing and treating mental disorders".

Key terms: cultural expression of symptoms, somatization, patient variables, clinician variables

Theoretical overview

The influence of culture in mental health can manifest in different ways.

One of these is **cultural expression of symptoms**. We say that cultural differences affect the expression of symptoms when representatives from two different cultures have the same mental disorder but differ in the way they experience it or express it in observable behaviour.

For example, people from some cultures exhibit **somatization**, where they report symptoms of mental disorders (e.g., depression) as physical symptoms rather than mental distress. They may complain about headaches, insomnia, fatigue, and stomach upset. Somatization has commonly been ascribed to collectivist societies like China or India. Presumably the reason somatization occurs is the cultural stigma existing in such societies around mental illness: if you complain about being "sad", people may say that you are simply trying to avoid work. However, having a physical problem such as stomach upset may be a legitimate reason to seek help and get time off work. It is important to understand that somatization is not simply intentional deception. People actually interpret and experience their mental disorder as physical problems: they actually believe that what they have is a stomach ache.



▲ Figure 6.8 Expression of symptoms

Differences in prevalence rates may be a result of differences in how symptoms of mental disorders are expressed and presented to psychiatrists. You might call such differences **patient variables** because they stem from the fact that patients behave differently. However, it is also conceivable that patient behaviour is actually not that different and it is the psychiatrists who introduce cultural biases in the interpretation of their patients' behaviour. You might call such differences **clinician variables**. In reality, both patient variables and clinician variables probably take effect.

Finally, response to treatment may also be different in different cultures. If the patient and the clinician perceive the treatment process differently (e.g., due to their cultural backgrounds), then treatment may be ineffective. For example, there is lack of compliance when a patient is prescribed medication but does not take it, or when a patient attends therapy sessions but does not engage in a serious manner.

TOK

How do we know the true nature of things?

If patients can present their mental disorder as something else, how can we see through this and identify the real problem?

How similar or different is it in psychology as compared to how the same problem is approached in other areas of knowledge?

Source 1

Lin, Carter, and Kleiman (1985) reviewed the clinical records of Chinese, Filipino, Vietnamese, and Laotian patients in primary care in the USA to determine the presence of somatization. A distinction was made between refugees and immigrants. Approximately half of the patients had been born and raised in Vietnam and had been forced to migrate as refugees due to the Vietnam War (1955–75). A second group of patients had lived in Mainland China, Taiwan, or Hong Kong, and it was their conscious choice to emigrate to the USA.

Somatization was defined as vague somatic symptoms such as headache, abdominal pain, dizziness, and insomnia in the absence of a clear cause. Somatization was diagnosed in 35% of patients. Refugees were more likely to have somatization than immigrants. Patients with somatization were more likely than patients with physical disorders to have a large household size and lower levels of education. Both these variables (household and school) point to how "traditional" the cultural background of the patient is. They were also more likely to be less proficient in English (this probably relates to lack of social support in their new country).

The authors concluded that somatization was one of the most important clinical problems in Asian refugees and immigrants. The more "traditional" their society was, the more they seemed to be prone to somatization. Refugee status also seems to be an important contributor to somatization: people who were forced to immigrate displayed more somatization than people who left by choice.

Source 2

Payne (2012) showed 239 therapists four specially designed clinical videos and asked them to make diagnostic judgments. In two videos, an actor played the role of a depressed man with classic symptoms of Major Depressive Disorder. In the other two videos, the actor "displayed" culturally expressed African American symptoms of depression. The latter were established on the basis of prior studies that identified the key differences in the symptoms of depressed African American patients compared with depressed White patients. The actor was either African American or White and each of the actors displayed classic symptoms of depression in one video and culturally expressed African American symptoms in the other. The two actors were of similar age and physical appearance and they were dressed identically.

Culturally expressed African American symptoms of depression were determined based on multiple publications and documents and included, for example:

- Irritable mood, increased hostility, and increased agitation (as opposed to depressed mood and sadness)
- Decreased overt expression of depression (less likely to say "I am depressed")
- Insomnia (as opposed to hypersomnia in "classic" depression).

There are two competing hypotheses to explain cultural differences in clinical diagnosis.

1. The clinician bias hypothesis—this suggests that African American and White patients exhibit similar depression symptoms, but clinicians mistakenly judge these symptoms differently because of personal prejudices, cultural ignorance, and so on.

2. The cultural variance hypothesis—this suggests that African American and White clients express their symptoms differently and clinicians are insensitive to such cultural differences.

The study did not directly support the clinician bias hypothesis, as no significant differences were found between clients of either race if they presented the same symptoms (e.g., African American clients presenting classic depressive symptoms were not misdiagnosed more often than White clients presenting the same symptoms). However, clinicians misdiagnosed depression more often when culturally expressed African American depressive symptoms were presented by clients of either race. It was concluded that it is not race itself that produces bias in diagnosis, but culturally specific expressions of symptoms (which clinicians seem to be unaware of). In other words, racial bias in diagnosis exists, but it is not caused by overt racism. The study suggests that clinicians can be better trained to recognize culturally specific expressions of symptoms, which will potentially decrease bias and increase the validity of diagnosis.

	Classic symptoms	Culturally expressed symptoms
African American patient	Not often	Often
White patient	Not often	Often

▲ Table 6.5 How often were patients misdiagnosed? Results from: Payne (2012)



▲ Figure 6.9 Screenshots from Payne's study (2012)

Source 3

Kinzie et al. (1987) examined 41 depressed Southeast Asian patients who underwent long-term treatment of depression with antidepressants in clinics in the USA. Levels of antidepressants in their blood were examined and no detectable medicine level was found in 61% of the patients, while a therapeutic level of medicine was only found in six patients (15%). This means that there was a high incidence of non-compliance with the treatment. Apparently, patients were not conscientious about following the prescription. There were some interesting inter-group differences as well. For example, compliance with treatment was significantly higher in Cambodian patients than in Vietnamese patients.

Reasons for low levels of medication in the blood may include patients' reluctance to take medication due to social stigma associated with taking psychiatric medicine. However, cultural attitudes towards authority may lead such patients to maintain the appearance of strictly following the clinician's guidelines. This may be done in an effort not to offend the clinician.

However, after a discussion of the problems and benefits of antidepressant treatments was held with the patients, rates of compliance improved (among the Cambodian patients also). The researchers concluded that cultural and educational factors may influence compliance with the treatment of depression. They claim that the level of antidepressant in the blood should be determined in Southeast Asian patients shortly after the start of the therapy. If low levels are detected, a doctor–patient discussion should be held about beliefs surrounding medication and its effects.

Discussion

If culture makes it difficult for patients to recover, to what extent is it possible to create a "micro-culture" in the therapist's office that would counteract the effects of the "bigger" culture?

ATL Thinking, Self-management

Claim

Cultural factors must be considered when diagnosing and treating mental disorders.

Exam tip

Read the claim very carefully and think about every word. This will greatly affect how you approach your analysis in Paper 3. For example, this claim states that cultural factors "must be considered". What does this mean? Does it imply that culture influences mental disorders or not necessarily? Try to uncover the assumptions and implications of the claim and make them explicit in your writing. This will help you approach the analysis in a systematic, structured, and focused way.

Analysis of the claim

The sources presented in this unit support the claim in three different ways. They show that cultural factors:

1. can affect the way depression is presented to the psychiatrist
2. influence the way a psychiatrist interprets the patient's behaviour
3. play a role in the extent to which the patient will be compliant with the prescribed course of treatment.

All three considerations are crucial in diagnosis and treatment. For example, culturally specific expression of symptoms may lead to misdiagnosis and mistreatment.

However, no study is perfect, and we should be aware of the crucial limitations in available evidence. Many such limitations represent challenges that are common in this entire area of research.

Somatization is not easily separated from alternative explanations. What we observe is that in some cultures, patients who may have depression express it through somatic symptoms and seek help from a medical doctor rather than a psychologist. The alternative explanations should be considered:

1. These patients actually do not have depression. Prevalence rates of depression in China are actually lower than in the USA.
2. These patients have depression and they experience it as such, but they consciously choose to misrepresent the symptoms to the psychiatrist, for example, due to cultural stigma.
3. These patients have depression, but they experience it as a somatic illness. They genuinely believe that what they have is a physical problem.

Depending on which explanation we choose, diagnosis and treatment may be different. But often available research evidence is consistent with more than one explanation. For example, evidence presented in Source 1 does not contradict any of the three explanations. Refugees and migrants from more traditional societies could experience their symptoms more somatically, but they could also consciously misrepresent symptoms. Differences were observed in "vague somatic symptoms": it is possible that there are no actual differences in depression among the groups involved in the study.

Studies need to be designed in such a way that allows us to eliminate some alternative explanations and accept others. For example, the study in Source 2 was specifically designed to test two alternative explanations (the clinician bias hypothesis and the cultural variance hypothesis) against each other. With this end, they had to manipulate two independent variables (race of the patient and type of symptoms presented) and create four videos: (1) African American patient presenting classic symptoms, (2) African American patient presenting culturally specific symptoms, (3) White patient presenting classic symptoms, (4) White patient presenting culturally specific symptoms. Since there were no differences

between groups (1) and (3) and between groups (2) and (4), researchers eliminated the clinician bias hypothesis. Of course, validity of the evidence in this study largely depends on the quality of the videos and on how successful the researchers were at controlling possible confounding variables.

The study presented in Source 3 suggests that Southeast Asian patients who underwent long-term treatment of depression with antidepressants in USA clinics do not show very high rates of compliance with treatment (they fail to take medication as prescribed). The reported result (61% of patients with no detectable level of medicine in the blood) seems like a significant amount. However, the source does not explicitly provide a comparison statistic for USA patients in USA clinics—what is their rate of compliance? Additionally, it has been reported that there are differences in compliance rates even within the sample of Southeast Asian patients (e.g., Cambodian and Vietnamese). It would be important to know exactly how large these differences are. If the variation observed within the sample of Southeast Asian is larger or even the same as the variation between Southeast Asian and USA patients, then we cannot make the claim that there is a difference between Southeast Asian and USA patients. This points to another common problem in research of mental health: selective reporting of results.

In summary, the claim that cultural factors must be considered in the diagnosis and treatment of mental disorders is fully supported by the evidence presented in the three sources above. However, we need to be aware of some common challenges of research in this area before making any definitive conclusions.

6.9 Culture and human development

What you will learn in this section

Key learning:

- To what extent are Western theories of development (e.g., Piaget) applicable to other cultures?
- Analysis of the claim "patterns and sequence of human development are unique in each cultural context".

Theoretical overview

A feature of Western theories of intelligence such as that of Piaget's is a sharp division between how broadly intelligence is defined and how specifically it is measured. The definition of intelligence in a broad sense is successful adaptation to the environment. This definition implies cultural differences: since environments are different, intelligence should also be different. However, the measurement of intelligence was typically reduced to a number of highly specific tasks, such as being able to tell if the amount of water in two glasses is the same.

This discussion closely links to Unit 3.2: Stage theories and continuous models of human development, Unit 3.3: Sociocultural factors in development, Unit 3.7: Theory of mind, and Unit 2.7: Cultural factors in cognitive processes.

Source 1

Vinden (1996) carried out a study of theory of mind in children of the Junín Quechua community. Theory of mind is the cognitive ability to understand other people's beliefs and intentions. Western children typically develop this ability between the ages of three and five years.

Theory of mind rests on the key understanding that people act upon their mental representation of the world rather than the world itself. It also allows you to realize that how things are and how they appear to be (to yourself and others) are not necessarily the same.

Junín Quechua is a native community living in the high Andes mountains of Peru. They have their own language and traditionally theirs is an oral culture with no writing. This is changing now as they are increasingly coming into contact with the larger society in Peru through schooling and trade, but their culture still has a large impact upon their lives.

In the language of Junín Quechua, concepts related to other people's mental states (such as "thought", "belief", and "denial") are referred to indirectly. For example, "What would he think?" in Junín Quechua will be expressed as "What would he say?", and instead of "deny" they would use "say no". Therefore, Junín Quechua do not have a vocabulary for mental states in their language.

Activity

Junín Quechua have a folk tale about a hungry fox who sees a large round cheese in the middle of the lake. However, it only looks like cheese but actually is a reflection of the moon. The fox asks an owl to help, but the owl refuses saying that it is not a cheese. The fox then attempts to swim and drowns. In the Junín Quechua version of the story, there is never any reference to the thoughts and beliefs of either the owl or the fox.

Researchers recorded a literal translation of this story and presented it to a sample of Canadian graduate students, asking them to write it down as they remembered it. All participants used at least one reference to a mental state, such as "he thought he saw a cheese" (Vinden, 1996).

Can you retell this short story to a partner without using any words referring to mental states?

ATL Communication, Self-management

Results of the study showed that Junín Quechua children fell behind Western children in terms of their performance on theory of mind tasks. For example, one task required participants to interact with an ambiguous object: a sponge that looked like a rock. The experimenter showed the object to the child and asked, "What is this?" After receiving the child's response, the experimenter then handed them the object and said, "Look at this carefully. Take it and squeeze it. You can see that it's not a rock but a sponge." The experimenter then took the object back and tested the child's understanding: "Does it look like a sponge or does it look like a rock? What is it really? Is it a sponge or is it a rock?" Finally, the experimenter asked a false belief question: "Shanti hasn't touched it. If he just looks at it, what would he think it is? Would he think it is a sponge or a rock?" Shanti had been introduced to the participants as a "child just like you" and represented by a cardboard cut-out. The word "think" in the native language literally translates as "say"—"What would he say it is?"

Results demonstrated that even those Junín Quechua children who understood appearance and reality (i.e., they correctly said that the object is a sponge and correctly said that the object looks like a rock) failed to produce the correct answer to the false belief question. They said Shanti would think that the object is a sponge. This only started to change gradually in older children (aged eight years and above).

The study demonstrates the immense influence that culture may have on cognition through language.



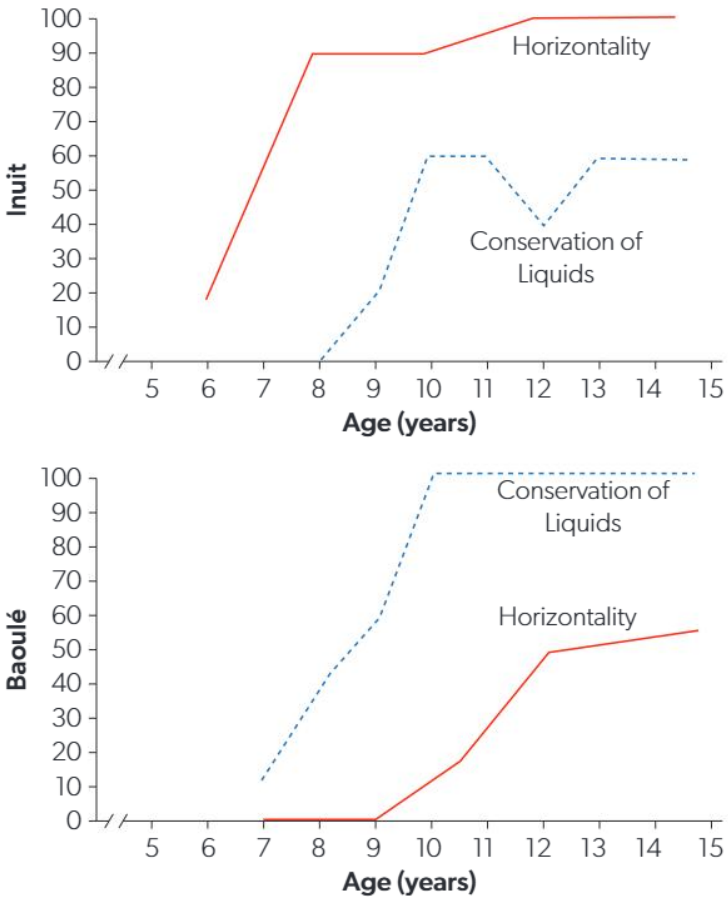
▲ Figure 6.10 Quechua children dressed in traditional clothing

Source 2

Dasen (1984) used a number of Piagetian cognitive tasks to compare performance of children in two cultural groups: the Inuit (Alaska) and the Baoulé (West Africa). The Baoulé people are agricultural: they grow crops, store them, and exchange them on the market. In contrast, the Inuit are a hunter-gatherer society.

Two of the tasks that were used in this study were the conservation of liquids and horizontality.

1. Conservation of liquids is the well-known Piagetian task where the researcher first shows the child two glasses with equal amounts of water, then pours the water from one of the glasses into a taller and narrower glass, and asks the child which of the two glasses now contains more water. Children at the preoperational stage of development typically say that the tall glass contains more water because they lack the understanding that it is the same amount of water in a different shape, and they are unable to mentally reverse the operation.
2. Horizontality refers to the set of tasks where the researcher shows the child a bottle half-filled with water and asks the child to note that the water surface is horizontal. The researcher then asks the child to predict (e.g., to draw or to show with a gesture) how the surface will change if the bottle gets tilted. The right answer, of course, is that the water will remain horizontal, but up until a certain age children believe that the water surface will get tilted as well.



▲ **Figure 6.11** **Percentage of children giving correct responses on the conservation of liquids task and the horizontality task.**

The sample included 10 children from each age group in each of the societies. Results of the study are presented in Figure 6.11.

It can be seen from these results that the Baoulé children acquire an understanding of conservation before they understand horizontality. It is the opposite for the Inuit children. Generally, conservation of liquids is well understood by the Baoulé but never fully understood by the Inuit children, and the opposite is true for horizontality.

The explanation could be that people value and develop those skills and concepts that are useful in daily activities to successfully adapt to the cultural environment. Precise quantitative comparisons may be important in agricultural societies who grow crops and exchange them on the market and not so important for hunter-gatherers. Conversely, spatial skills and abilities exemplified by the horizontality task may be more valued in nomadic societies that have to find their way in space.

Dasen (1984) believed that these differences are a matter of how children are traditionally educated in a given society.

Exam tip

This is the kind of situation in which your background knowledge on Piaget's theory of cognitive development may be useful to provide an in-depth analysis of the claim in relation to the given sources. The relevant ideas are Piaget's stages of cognitive development and the sociocultural studies that challenged the notion that these stages are always fixed. One could also bring up the debate between biological maturation and learning as a leading factor of cognitive development.

As you see, the provided sources already contain all relevant information, but if you are aware of a broader context in which these findings may be placed, consider mentioning it where relevant.

Claim

Patterns and sequence of human development are unique in each cultural context.

Analysis of the claim

Implicit in these studies was a comparison with the "standard" developmental milestones registered previously in samples of Western children.

It was Piaget's influential belief that cognitive development of human children follows the same genetically predetermined sequence of biological maturation, meaning that "cognitive structures" are expected to mature in a particular order and at a particular pace. Observing substantial cultural variation (e.g., a reversal of stages) would contradict this theory and suggest that development is determined by sociocultural rather than biological factors. This links back to a discussion of stage models versus continuous models in developmental psychology.

Vinden (1996) found that Junín Quechua children develop theory of mind abilities (such as understanding false beliefs) at a later time as compared to Western children (eight years old versus three–five years old). This finding casts doubt on the idea that the sequence of cognitive development is biologically predetermined. However, one needs to be careful when making sweeping generalizations for the following reasons:

- The study shows that the development of theory of mind in Junín Quechua children is slower than in Western children, but it could still follow the same sequence (in terms of which cognitive structures mature before theory of mind abilities, and in which order).
- Although at a later time, the children in the study still developed theory of mind abilities and passed the standard tasks despite the absence of mental states vocabulary in their language.
- It should always be kept in mind in studies like these that performance of a child on a standard developmental cognitive task may depend on many factors implicit in the experimental situation. Examples include their ability to understand the experimental instruction, whether or not they find the task relatable, and appropriateness of the task to the cultural context. It is possible, for example, that children do understand false beliefs, but up to a certain age are unable to express this understanding in language. The use of technology such as eye tracking could reveal if this is true, but such technology is rarely used in cross-cultural studies.

Source 2 (Dasen, 1984) is another example of a study that takes "standard" developmental cognitive tasks and applies them in a culturally unique sample. Related to the claim that is being analysed, this study contains two crucial findings:

1. The sequence of development was reversed in one culture compared to the other: conservation first versus horizontality first.
2. One of the cognitive milestones was never fully achieved in each culture (conservation of liquids in the Inuit and horizontality in the Baoulé).

TOK

No law is truly universal. There always exist some random fluctuations caused by a variety of factors, as well as some exceptions. This is especially true for human sciences.

Where is the line between a "universal pattern with some cultural variation" and "no universal pattern"?

However, before we jump to the conclusion that human development is "unique in each cultural context", we should consider the following:

- The reversal observed in this study occurred within a single stage of cognitive development (according to Piaget)—the preoperational stage. Piaget believed that cognitive structures of the concrete operational stage, for example, cannot develop prior to cognitive structures of the preoperational stage. However, he acknowledged that some variation of sequence is possible within one stage.
- The author explains the observed differences by the differences in daily cultural practices (agricultural society versus hunter-gatherers), but this part is a speculation that is not directly supported by data. There could be other explanations.
- Similar to Source 1, we should bear in mind that the performance of children on these tasks could be affected by their understanding of the instruction, their language, the dynamics of their relationship with the interviewer, and other factors. It would be important to know details such as what language the interviews were conducted in, whether or not the same person conducted interviews in both cultures, and so on.
- Only 10 children participated in the study per age group. From the measurement perspective, if we were to add confidence intervals to the plots, the confidence intervals would be quite large due to the small sample size.

One common problem to be remembered in studies like this is the imposed-etic bias. Stages of cognitive development, the ideas of conservation and horizontality, theory of mind, and Piagetian tasks themselves are all constructs and measurements that were designed in the Western cultural and scientific framework. Simply testing the performance of children in other cultures on these same tasks may bear the danger of bringing in constructs and measures that are fundamentally alien to the culture and do not reflect how people in that culture think and what it means for them to be "intelligent".

Overall, the studies presented in these sources provide support for a somewhat weaker claim: that there exists some cultural variation in cognitive development—at what time certain cognitive abilities are achieved and in what sequence. This variation may be related to the unique cultural context of a given society, such as their daily experiences, schooling, and language.

6.10 Culture in learning and cognition

What you will learn in this section

Key learning:

- Some cultural differences in cognitive processes: field dependence and field independence, holistic and analytic cognitive styles.
- Analysis of the claim "culture influences the way people cognitively process information".

Key terms: field dependence, field independence, holistic cognitive style, analytic cognitive style

Theoretical overview

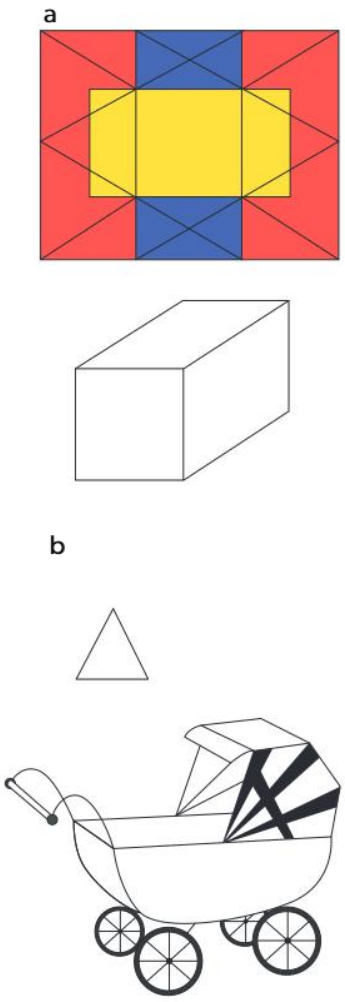
You would imagine that the most basic and fundamental cognitive processes, such as sensory perception, should not be influenced by culture. After all, no matter what culture you grew up in, you can still tell the difference between horizontal stripes and vertical stripes, for example. However, as it turns out, this is not exactly the case and there do actually exist differences between cultures even in such fundamental things as sensory perception.

Witkin and Berry (1975) suggested that a person's cultural background has an influence on their **field dependence/field independence**. Field dependence is when your perception is easily affected by the surrounding context. Conversely, field independence is when you can easily dissociate from the surrounding context and focus your perception on the object of interest.

For example, in one of their tests known as the Embedded Figures Test (EFT), researchers showed participants images of complex objects and asked them to find a less complex "figure" (or object) hidden somewhere in the image (see Figure 6.12).

Based on research, R.E. Nisbett (2003) argued that Europeans and North Americans are more field-independent in their perception and tend to think analytically, focus on the target object and its features, whereas East Asians (e.g., Chinese, Japanese, and South Koreans) tend to be more field-dependent and think more holistically taking into account the context.

A related concept is **holistic** versus **analytic cognitive styles**. Holistic cognitive style implies stronger attention to context, a focus on the whole, and how parts within the whole are related to each other. This is more characteristic of collectivist societies. Analytic cognitive style implies a focus on separate details and less attention to the context and the surroundings. This is more characteristic of individualist societies.



▲ Figure 6.12 Sample items from the Embedded Figures Test (EFT)

Discussion

In what ways do you think your culture has influenced your cognitive processes, such as perception, memory, and decision-making?

ATL Self-management

Why do these differences exist? One theory that gained prominence explains that cognitive styles are a product of the structure of society. Namely, societies whose members traditionally have been more dependent on one another may have evolved holistic cognitive styles with a heightened attention to context and relationships. For example, agricultural societies that live in closely knit communities and depend on trade. On the contrary, nomadic societies that have traditionally relied on hunting may have evolved to be less sensitive to contexts and more sensitive to the target goal.

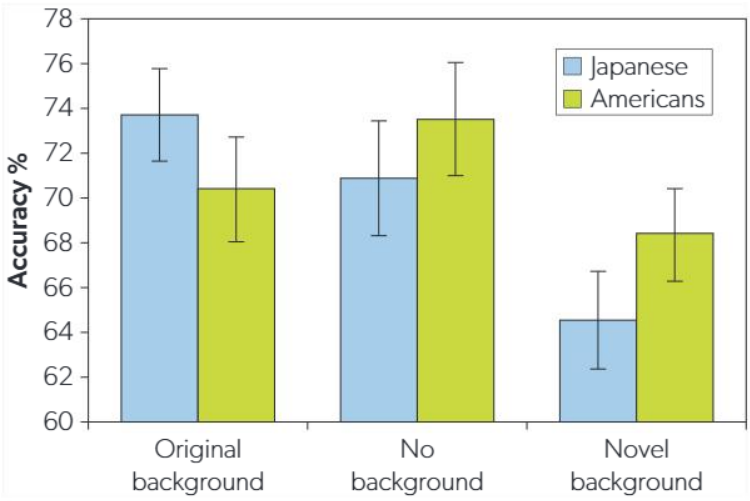
Source 1

Masuda and Nisbett (2001) conducted a study in which they presented a group of North American and a group of Japanese students with animated vignettes of underwater scenes. Each vignette included a few obviously visible fish as well as several smaller and less prominent details such as rocks, shells, snails, frogs, newts, bubbles, and smaller fish. These smaller details appeared to be in the background of the "main scene". The colour of the water was also different in each vignette.

All participants watched the videos in highly standardized conditions (e.g., the same type of monitor in both venues, a chair with a fixed chin rest to ensure that the distance between the viewer's eyes and the screen was exactly 15 inches, etc.). After watching, participants were given pictures of 90 objects—45 previously seen in the videos and 45 novel objects that did not appear in the vignettes. Some objects had the original background (e.g., the same colour of water as seen in the video), some objects had the background altered. For each object, the participants had to indicate whether they had seen it in a video ("yes" or "no") as well as indicate their level of confidence on a scale ranging from 1 (not at all confident), to 7 (extremely confident).

The expectation was that the recognition by Japanese participants would be more influenced by the background of the object: if presented with the original background, Japanese participants would more easily recognize the object than if the background had been changed. Conversely, it was expected that recognition by American participants would be relatively unaffected by whether or not the background remained the same. Results of the study confirmed these expectations. You can see in Figure 6.13 that Japanese participants recognized the objects much more accurately when the objects were presented against their original background (as opposed to a new background), whereas this difference was statistically non-significant for the American participants.

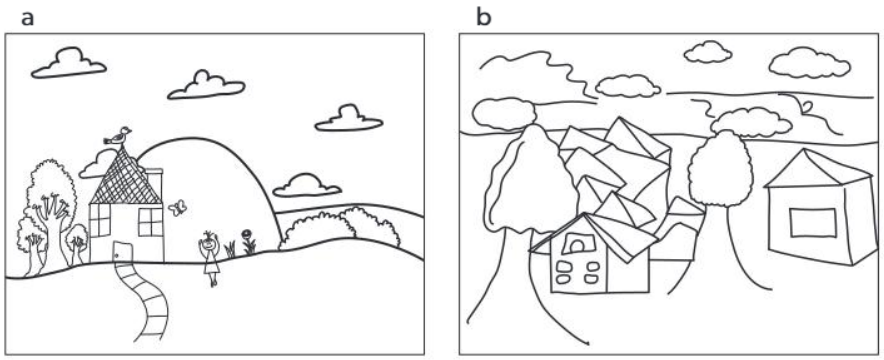
Interestingly, one alternative explanation that the authors of the study considered was that Japan consists of islands surrounded by ocean and therefore, the Japanese may be expected to be more familiar with sea-related visual stimuli. What if the differences in the study may be explained by this familiarity rather than a deep difference in patterns of perception? To eliminate this possible confounding variable, Masuda and Nisbett (2001) conducted a follow-up study which had a similar set-up, but the stimuli this time consisted of animals and scenery typically found in North America. Findings were replicated: the perception of Japanese participants was still much more sensitive to context. This alternative explanation was therefore ruled out.



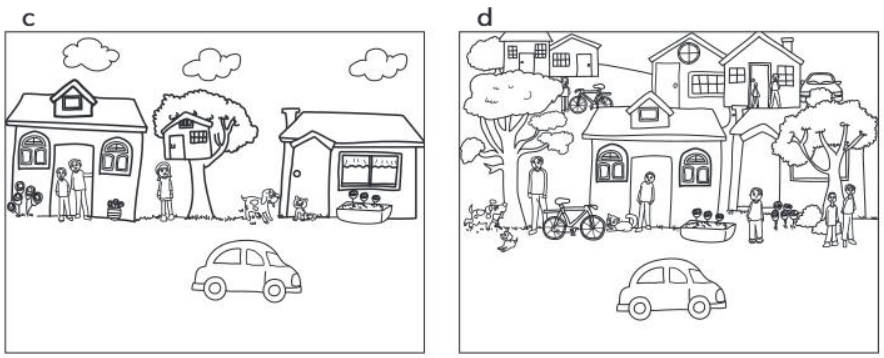
▲ Figure 6.13 Results of Masuda and Nisbett (2001)

Source 2

Senzaki, Masuda, and Nand (2014) observed similar differences in Grade 2 children's artwork. They found that Japanese children tend to draw the horizon higher and include more contextual details in their drawings, as compared to Canadian children. See an example in Figure 6.14.



Study 1: A Canadian child's landscape drawing (a) and a Japanese child's landscape drawing (b) at Grade 4.



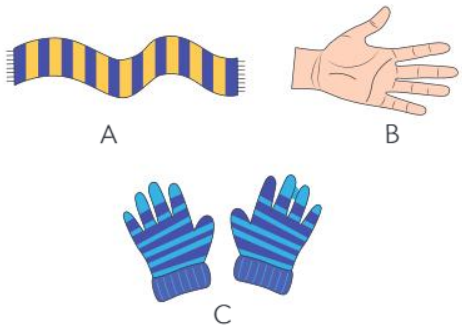
Study 2: A Canadian child's collage work (c) and a Japanese child's collage work (d) at Grade 4.

▲ Figure 6.14 Examples of drawings of Canadian and Japanese children

TOK

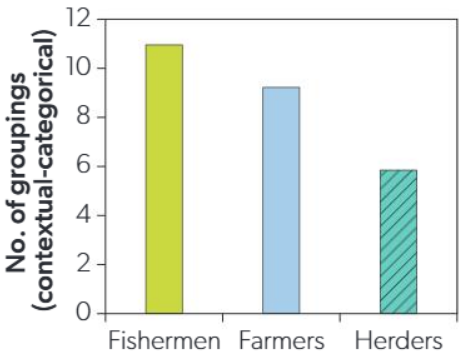
Do you think it is possible to conduct a true experiment in which culture would be the independent variable manipulated by the researcher? Imagine that such an experiment has been conducted (no matter how absurd it is). Describe this experiment.

a



What goes with this? A or B

b



▲ **Figure 6.15** An example of a grouping task (top) used in Uskul et al. (2008) and results of the study (bottom). Herders chose responses such as B (hand) less frequently than farmers and fishermen.

Source 3

Researchers believe that people with an analytic cognitive style tend to categorize things taxonomically—for example, grouping a cow and a chicken together because they are both animals. In contrast, people with a holistic cognitive style tend to categorize objects based on their belonging to the same context. For example, putting a cow and grass in the same group because cows eat grass.

A study by Uskul et al. (2008) found that Turkish farmers and fishermen tend to have a more holistic cognitive style than Turkish herders. Farmers and fishermen live in large communities with a high degree of being dependent on each other. In contrast, herders do not require much cooperation and are more dependent on individual decision-making. In the study, participants were given 18 triads of three objects (such as a glove, a scarf, and a hand) and asked to indicate which two of the three went together. Two of these objects always shared a contextual relationship (e.g., glove and hand) and two objects always shared a category (e.g., glove and scarf). As a measure of the dependent variable researchers subtracted the number of categorical groupings from the number of contextual groupings. For example, if a participant grouped 7 of the 18 triads categorically, that participant would be assigned the score of 11–7 = 4. The larger the number, the larger the preference of the participant for contextual groupings (therefore, indicating a holistic cognitive style).

Results showed that all three groups preferred contextual grouping, but this preference was more pronounced in farmers and fishermen as compared to herders see (Figure 6.15). The difference was statistically significant at $p < 0.005$.

Claim

Culture influences the way people cognitively process information.

Analysis of the claim

Broadly speaking, the studies presented in the three sources support the given claim because they report the existence of cultural differences in such cognitive processes as perception, memory, and thinking.

However, there are two considerations that outline the limits of this support.

First, the nature of the discovered differences was that of styles but not performance. For example, both the Americans and Japanese were successful at recalling the objects they had perceived (Source 1), but within that, North American participants were better at recalling objects with no backgrounds and Japanese participants were better at recalling objects within the original context. Both herders and fishermen in the study in Source 3 used contextual reasoning more often than categorical reasoning. However, in one group this difference was larger than in the others. Therefore, in the phrase "the way people cognitively process information", we should read "the way" as "how", but not "how well" or "how much" or "how deeply".

Second, that culture "influences" cognitive processes is a cause–effect inference which is not directly supported by the fact that there exist cultural differences. This is because any observed difference between pre-existing groups may be determined by any factor that differentiates these groups. For example, it was

observed that Canadian and Japanese children have different approaches to drawing landscapes (Source 2). However, Canadian and Japanese children are different in a number of ways, not limited to the predominant cultural values in these societies (collectivist versus individualist). They may be exposed to different typical landscapes when they grow up, or they may differ in where they spend most of their time (outside or inside). The point is, if we do not manipulate the independent variable, then it is difficult to know what exactly the observed inter-group differences should be attributed to.

There are also some considerations more specific to the nature of these particular studies.

In Source 1, the difference between accuracy of recall between the original background condition and the novel background condition was statistically significant in the Japanese participants but not in the North American participants. Having said that, it is important to notice the magnitude of difference. The bar graph reported in Source 1 is somewhat misleading because the *y*-axis does not start at 0, causing the differences to appear larger than they really are. The difference between conditions in the Japanese group, although statistically significant, is only approximately 9%, whereas this difference is approximately 2% among Americans (insignificant). Similarly, in the study reported in Source 3, although the result was highly statistically significant at $p < 0.005$, we can see that the difference between fishermen and herders is 5 points on the scale that may range from –18 to +18, which may not be considered that much of a difference in an absolute sense.

The study reported in Source 2 may suffer from a lack of construct validity. Results were interpreted as evidence that Canadian and Japanese children have differences in the cognitive process of sensory perception. However, the fact that they draw pictures differently may be a manifestation of something else. It cannot be ruled out that they perceive the world similarly, but they choose to draw it differently due to some existing cultural conventions.

The study reported in Source 1 sets a good example in this sense because the authors consider a possible alternative explanation of their findings and conduct a follow-up investigation specifically designed to test this alternative explanation. This is the way research programmes in this area should be organized—systematically testing all possible alternative hypotheses until we are left with one explanation that appears to be the most credible.

In summary, evidence presented in these three sources is not enough to support the claim in its entirety, but it does demonstrate that there exist cross-cultural differences in the way people process information (especially in regard to memory and thinking).

Activity

Imagine you had unlimited time and resources to conduct any psychological research, limited only by your imagination. We have concluded here that evidence from the three sources is not enough to support the given claim in its entirety. Could you imagine a study (or a series of studies) that would be enough? What would be the requirements that such studies would have to meet?

ATL Thinking, Research

6.11 Technology

What you will learn in this section

Key learning:

Technology in human life:

- Digital technology: the internet, smartphones, social media, video games.
- Arguments for the negative psychological effects of technology.
- Arguments for the positive psychological effects of technology.

Key challenges of research:

- It is not technology itself that produces an impact, but the way it is used.
- Some types of digital technology may have more positive effects than others.
- Effects of digital technology may depend on the age of the user and the time when it was introduced.
- The same technology may have both positive and negative effects.

Key considerations in designing research:

- Need to combine controlled laboratory studies with research conducted in real-life settings.
- Need to supplement self-report measures with more objective data.
- Need to investigate both transient short-term effects and long-term effects that leave a permanent impact.

Key terms: digital technology, short-term and long-term effects

Overview

At this time, human existence is unthinkable without technology. However, going back in human history, this statement would probably have been true at all times. When early humans started using tools such as fire and stone weapons for hunting (e.g., spear, knife), this changed their entire existence and eventually led to the emergence of civilization. These tools have allowed humans to become what they are. They have changed us: modified our ways of life, changed the skills required to be successful, pushed the development of cognitive functions. In a broad sense, tools like fire or a stone knife may also be referred to as technology.

The focus of this section will be on **digital technologies** such as the internet, smartphones, social media, and video games. We are interested in how digital technology changes the way we are: our cognitive processes, our mental health and well-being, the way we establish and maintain relationships with others, the way children develop.

There are many issues that are debated in this area, but the overarching debate is probably whether or not the influence of technology has been mostly positive or mostly negative.

Those on the "negative side" claim that digital technology makes us stop using our human abilities, and that we eventually lose them. For example, having easy access to the internet means always having information at our fingertips. We may rely on this so much that we lose the ability to remember information, or the ability to separate useful information from that which is less useful. With the invention of the "search bar", we may have lost the kind of cognitive ability we need when we, for example, look for necessary information in a library. With the invention of generative AI, some believe that students will soon lose the ability to think for themselves: in any difficult situation they will just ask their AI assistant. In human relationships, proponents of the "negative view" maintain that digital technology provides a surrogate which makes us addicted and makes us lose our social skills and interpersonal empathy. Social media removes the necessity to initiate contact: where children used to invite each other to play together, they now "like" each other's photos or reposts. Dating apps remove the necessity to nurture intimate relationships: where we used to overcome our shyness to approach a person and invite them on a date, we now swipe right or left and let the algorithm do its job. They say that digital technology makes our children more impulsive, less patient, and less attentive. They claim that social media creates addiction and anxiety and leaves a lasting negative impact on our mental health.

Those on the "positive side" do not deny that negative trends exist, but they claim that negative outcomes are associated with misuse or abuse of technology. Their argument is that, if used in a healthy way, the advantages of digital technology outweigh its costs. For example, they acknowledge that the internet has reduced the requirement for us to remember things and we may be losing this ability. However, they argue that the internet has opened up new opportunities for exercising advanced cognitive skills: critical analysis of information, thinking about credibility of sources, triangulation of evidence. The invention of generative AI makes it easy for us to redirect questions to an AI assistant instead of thinking for ourselves. However, it may also present an opportunity for us to practise asking correct questions, formulating questions in a way that is likely to give us what we are looking for.

Maybe, they say, it is actually fine for us humans to delegate some of the low-order cognitive skills to a machine as long as we can focus on the development of more complex, higher-order cognitive skills. We have delegated to machines such things as lifting heavy objects and quickly getting from one place to another, yet nobody is complaining that later generations of humans are becoming worse runners and less efficient at lifting heavy weights. They see positives in the area of human relationships as well. Perhaps technology can make us more empathetic by increasing our exposure to personal stories, and increase our interpersonal skills by making it easier to initiate contact for subsequent face-to-face encounters. In mental health, scientists and therapists are actively creating AI chatbots and applications for e-therapy—this bears the promise of making mental health services accessible to a much wider range of people, including those from geographical regions where mental disorders are stigmatized, as well as those from low SES communities.

It is difficult to determine which of the positions is closer to the truth. The debate is ongoing. Technology is changing very rapidly, so much so that it is difficult for scientific research to keep up.

However, it is worth remembering some key challenges that researchers in this area are facing:

- It could be the case that digital technology in itself is neither positive nor negative, and that what matters is the way technology is used. For example, social media can be used to strengthen and maintain existing face-to-face friendships, or it can be used to make new friends online. It could be that these different uses create different outcomes (Carrier et al., 2015). The internet can be used occasionally when required, or it can become addictive—for example, browsing for the sake of browsing. Again, cognitive and health outcomes in these situations may be different.
- "Digital technology" is a very broad term and it may not be a very good idea to try to make a judgement about digital technology in general. Some types of digital technology may affect us more positively than others. For example, a virtual reality flight simulator for pilots and a video game are similar in principle, but used for different purposes and therefore may have different effects.
- Effects of digital technology may depend on the age of the user and the time when technology was introduced. For example, when social media first appeared, there were few users, so it was hard or impossible to use it to maintain existing friendships. Instead, social media was used to make new friends online, probably knowing that you would never meet them in real life (Howard-Jones, 2011). Today, social media can be used differently. Think about "using telephones"—the way we used telephones 20 years ago and the way we do now are entirely different. Therefore, it makes little sense to speak about the influence of "telephones" on our cognition and mental health across time periods.
- The same technology can have both types of effects: positive and negative. For example, playing violent video games has been the focus of studies looking at game addiction and aggressive behaviour. At the same time, playing such games can improve our spatial reasoning and even make us better at understanding some abstract scientific concepts that involve spatial thinking, such as plate tectonics (Sanchez, 2012).

Taking all this into consideration, we should be careful to design research studies in a way that does not oversimplify the reality of things.

There are some other typical considerations that should be considered when planning research on the effects of technology:

- The effects of technology observed in carefully controlled laboratory conditions may not be the same as those existing in real-life conditions. For this reason, both carefully controlled laboratory experiments and field studies (correlational and observational) are important.
- When we use some types of technology (e.g., social media), our perception of time is distorted. This needs to be considered when we rely on self-reporting to measure, for example, daily screen time.
- Longitudinal studies are especially important to make a distinction between **short-term effects** (which may be transient) and **long-term effects** that leave a permanent impact.

6.12 Technology in human relationships

What you will learn in this section

Key learning:

- How have digital technologies (such as smartphones and social media) affected our interpersonal relationships?
- Analysis of the claim "using technology for communication has both positive and negative effects on romantic relationships".

Key term: digital technology, mediating variables

Theoretical overview

How has the emergence and proliferation of **digital technologies** affected our interpersonal relationships? Modern communication is massively mediated by technology: emails, text messages, social media apps, and other solutions allow us to communicate without a face-to-face encounter. Technology has also provided new ways of communication that have not existed before—for example, online communities and group chats where your messages or media can be immediately visible to multiple individuals. People are more connected than ever before, but at the same time more separated than ever.

In 2015, more than 50% of adolescents in the USA reported that they had developed friendships online, and that their preferred method of communication with friends was text messaging (Pew Research Center, 2015, as cited in Okdie and Ewoldsen, 2018).

As is the case with other areas of research, speaking about the positive and negative influences of digital technology on interpersonal relationships (on the whole) does not make sense. The influence may be positive and negative, depending on a number of contextual factors. Therefore, a detailed study of these contextual factors and **mediating variables** is more important than looking for a universal answer.

For example, it has been demonstrated in some studies that using mobile phones while in the process of face-to-face communication diminishes our ability to connect with each other meaningfully. Conversely, some studies have shown that individuals who have been rejected by peers and ostracized can benefit from interactions on social media and satisfy their need to belong, potentially preventing mental health issues (Okdie and Ewoldsen, 2018).

Exam tip

Review Chapter 1: Research methods and data interpretation to refresh your knowledge of surveys and questionnaires. When can we say that a survey is a research method in its own right (quantitative research study) and when is it a data collection tool used within some other research method?

The research method in Morey et al. (2013) was a correlational study. But variables for this study were measured by means of self-report questionnaires.

Sometimes studies like these would be referred to as a "survey", but this only highlights the fact that data was collected through self-report. Do not get confused by this usage of the word "survey".

Activity

Review what you have learned about attachment styles from Unit 3.8: Attachment.

ATL Self-management

Source 1

Morey et al. (2013) investigated the role of communication technology in young adults' romantic relationships and the mediating role of attachment styles. The study was an online survey of two cohorts of dating undergraduates. The measures on the survey included the following:

- A questionnaire of attachment style (secure, anxious, avoidant)
- Frequency of in-person communication versus technology-mediated communication
- A self-report measure of quality of romantic relationship: relationship satisfaction, intimacy, support, and conflict.

Previous research had demonstrated that the attachment style acquired in childhood is likely to influence our romantic relationships in adulthood (Hazan and Shaver, 1987).

Four types of communication were assessed in this study: telephone calls, email, social network sites, and texting.

Patterns of correlations were examined and the following results were obtained:

1. More frequent texting was related to greater relationship satisfaction and greater self-reported intimacy among participants with an avoidant attachment style, but unrelated among participants lower in avoidance.
2. More frequent use of communication through social network sites was associated with greater intimacy and support, but only for people with an anxious attachment style— this association did not exist for individuals with low attachment anxiety.

The last finding is somewhat counterintuitive because the use of social networks presumably can support feelings of jealousy and make individuals even more anxious about their partner. However, public declarations of commitment (e.g., posting pictures together with a partner to the public) can serve as a reinforcement of relationship status, thus helping people overcome their anxiety. The study did not include any specific measures to test this hypothesis.

Overall, the study "illustrates how attachment can help to explain why the use of specific technology-based communication channels within romantic relationships may mean different things to different people, and that certain channels may be especially relevant in meeting insecurely attached individuals' needs" (Morey et al., 2013, p. 1771).

Source 2

Vaterlaus et al. (2018) studied the perceived influence of technology on adolescent romantic relationships. They used a purposive sample of school students and college students. School students were included because access to technology as well as establishing romantic relationships were their current experiences. College students below age 30 were included because they could provide a retrospective view on the topic (how media and technology has influenced the relationships they have established).

The study used an online survey. The survey included the following open-ended question: Do you think that interactive technologies (e.g., mobile phones, social networking, video chat, email) influence school romantic relationships? Why or why not?

Researchers conducted a thematic analysis of responses. First, two researchers independently read all responses and developed coding categories in discussion with each other. The answers were then coded by the same two researchers independently using the developed coding categories. Their agreement in the categories reached 90%. Discrepancies were resolved through discussion. A third researcher checked the coding. Credibility checks were also conducted by sharing the results with some participants and asking them to validate the conclusions.

Participants generally agreed that technology facilitates all cycles of a relationship: social networks are a place where initial contact occurs, texting is the most common way to get in touch and resolve ongoing issues. One participant claimed that the amount of technology-mediated communication in adolescent romantic relationships is "about 60 per cent-ish". Technology is also used to end a relationship, and it is much easier to do than in person.

On the other hand, some participants noted that technology-assisted communication may become dissociated from real life: "If the couple begins talking through text and not face-to-face, it is sometimes harder to talk to them in person if you don't spend enough time with them. They become a separate person... over text than they are in person, and that is definitely not good in a relationship" (Vaterlaus et al., 2018, p. 664).

Participants also maintained that technology in communication facilitates openness. This is because people do not immediately see the consequences of their utterances, so they say things over technology that they would not easily say in real life. One participant (female, 19) said: "People text things that they would never say in real life ... and then it makes things awkward when you see them in real life" (Vaterlaus et al., 2018, p. 662).

On the negative side, there was a concern that the use of technology was leading to a decline in face-to-face communication. Simply put, technology may have created a situation in which romantic partners avoid talking about serious matters in person and would rather discuss them over text. A higher percentage of college students shared this perspective as compared to school students, which suggests that these problems may be more obvious in retrospect.

Source 3

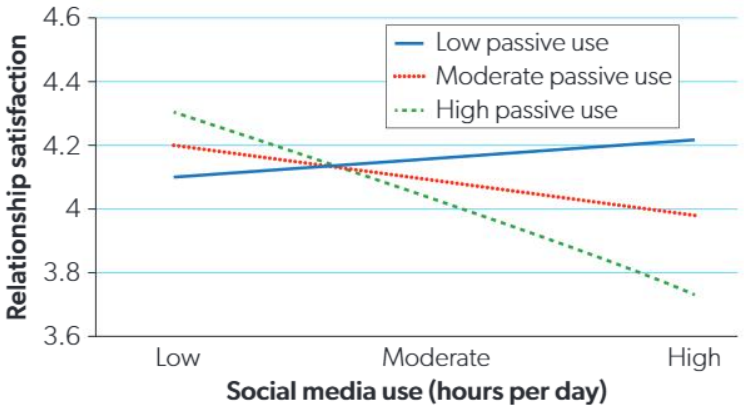
Quiroz and Mickelson (2021) investigated the correlation between frequency of social media use and relationship health, and how this correlation depends on type of social media use (active versus passive). To use social media actively means to share information, post content, and interact with others. Passive social media use implies simply browsing content posted by others. The two uses are very different. Theoretically passive social media use would be more harmful to one's well-being because it encourages social comparison. This damages self-esteem. Active social media use can promote, at least theoretically, establishing and maintaining meaningful relationships.

The study involved an online survey of a sample of 432 adults in the USA who had been in a romantic relationship at least three months (53% dating, 15% living together, 32% married). The average age of participants was 28 years. Participants received either course credit or monetary compensation (US$1). There were three attention checks included in the survey, and participants who failed more than one of them were excluded from analysis (27%).

The quality of romantic relationships was measured by two self-report scales: satisfaction and commitment. Self-report measures with Likert scales were also used to assess how frequently participants use social media sites (hours per day), which sites they prefer using, and how exactly they tend to be using them (posting content, communicating with friends, or merely following others and reading updates).

Results of the study revealed the following:

- When considered in general, number of hours per day spent on social media was negatively correlated with relationship quality (both satisfaction and commitment).
- However, there was a significant interaction between frequency of use and type of use (passive or active) in their influence on relationship quality. More specifically, in the sample of participants who tended to use social media actively rather than passively, the correlation between frequency of use of social media and quality of romantic relationships was either positive or insignificant. Conversely, in the sample of participants who preferred to use social media passively, there was a negative correlation between frequency of use and relationship quality. This effect was especially pronounced in female participants (see Figure 6.16).



▲ Figure 6.16 Results of Quiroz and Mickelson (2021): women's hours per day of social media use and relationship satisfaction, depending on passive use

Discussion

What is your personal experience with using communication technology to establish and maintain relationships (be it romantic relationships or friendships)? Conduct a semi-structured interview with a partner. Have you found any results that add to the three sources presented here?

ATL Research, Communication

Claim

Using technology for communication has both positive and negative effects on romantic relationships.

Analysis of the claim

The statement seems justified by evidence in the sense that the results of research studies are indeed inconsistent. In some cases, studies demonstrated positive effects of technology on some aspects of romantic relationships and negative effects on other aspects. However, as it is formulated, the statement does not do justice to the research findings accumulated in this area. It sounds like we do not know exactly if the effects of technology on romantic relationships are positive or negative. This is not the case. We do know, but the question of whether technology has a "positive or negative" effect on romantic relationships is too simplistic.

Research has demonstrated the importance of mediating variables in answering this question. It is usually the case that effects are negative given one configuration of mediating variables and positive or neutral given some other configuration. Prominent examples of mediating variables that have been discovered in research include an individual's attachment style, type of use of social media, gender, and age.

For example, Quiroz and Mickelson (2021) (Source 3) demonstrated the importance of considering the type of use of social media. Among those who use social media passively, the amount of social media use is associated with lower quality of relationships. However, among those who use social media actively this association becomes neutralized and in some cases, even reversed. This is a clear case of statistical interaction of effects.

Morey et al. (2013) (Source 1) obtained a similar pattern of findings. For example, among individuals with an anxious attachment style, more frequent use of social media for communication was associated with greater intimacy and support in relationships. However, among individuals with low attachment anxiety this association did not exist.

Vaterlaus et al. (2018) (Source 2) in a qualitative study also demonstrate that there are both positives and negatives, and it depends on how exactly technology is used.

Our knowledge of mediating variables is based on a correlational design where the relationship between variables A and B is assessed on different levels of variable C. It is estimated how much of a difference variable C creates in this relationship. However, the decision on which variable to test as a potential mediating variable is a theory-driven decision. Data cannot just tell us that a certain variable plays a mediating role. Instead we ask a question about specific data and analyse the data to obtain a response: "I think this one is a mediating variable … let me test it … yes or no."

Overall, the statement is justified by available evidence, but the statement is very superficial. It could be applied to practically any other area in psychology. Research has actually made more progress than this and we are able to say, with some degree of confidence, if the influence of technology on romantic relationships will be mostly positive or mostly negative in a given situation.

6.13 Effects of technology on health and well-being

What you will learn in this section

Key learning:

- Problems of research investigating the influence of screens on mental health: long-term and short-term effects, mediating variables, curvilinear relationships.
- Analysis of the claim "using social media has a negative impact on students' mental health".

Key terms: curvilinear relationship, mediating variables, short-term and long-term effects

Theoretical overview

Social media is often blamed for depression and other mental health problems, especially among adolescents. One of the reasons behind this could be the fact that the USA and the UK witnessed a sudden surge in the rates of depression, anxiety, and self-harm among teenagers in 2012–13. The most obvious cause was social media, which also increased in use around the same time. It reached the point at which two thirds of teenagers were using it on a daily basis (Haidt and Allen, 2020).

Here are a few typical problems existing in research investigating the influence of digital technology ("screens") on mental health:

1. "Screens" may include a lot of things: video games, educational apps, completing school assignments by typing rather than handwriting, social media use. All these things may have different effects on well-being. It is debatable if investigating the correlation between the total amount of time spent on screen (irrespective of how exactly it is used) and mental well-being is good practice.
2. Like other areas in which technology influences behaviour, the relationship may be **curvilinear**. For example, it could be the case that greater screen use is associated with better well-being, but only to a certain point, after which the association gets reversed.
3. Similar to other research areas, careful consideration of potential **mediating variables** is important. If mediating variables are not documented and not considered, research findings can be contradictory.
4. **Short-term and long-term effects** of technology on well-being may be different, but knowledge of both is equally important. Short-term effects are easier to investigate under controlled conditions, including true experiments. Long-term effects have to be studied in real-life situations in which strict control of variables is impossible.

Source 1

Hunt et al. (2018) is one of the few studies that used a true experiment for the study of social media and its influence on mental health. Participants were 143 undergraduate students from a university in the USA. They were required to have certain social media accounts and to own a particular brand of phone.

In this study, participants were required to send researchers screenshots of their phone battery usage each night for a period of four weeks. This screenshot provided information about the number of minutes each application had been actively open on the screen in the past 24 hours.

At the start of the study (Week 0), participants filled out a baseline survey. The survey included questionnaires that measured the following parameters, among others: depression, loneliness, social support, self-esteem, autonomy, self-acceptance.

At the end of Week 1, students completed the survey again, before being split into two conditions:

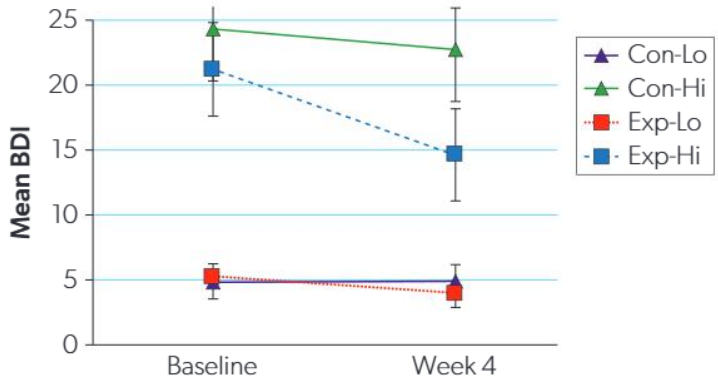
1. The control group was told to continue using social media as usual.
2. The experimental group was told to limit their usage of certain social media apps to 10 minutes per platform per day.

At Week 2, Week 3, and Week 4, participants completed the survey again.

Results:

- In the group of participants who had low baseline depression: those in the control condition retained the same level of depression after four weeks, but those in the experimental condition showed a statistically significant, but small improvement (reduction of score from 5.1 at baseline to 4.1 at the end of Week 4).
- In the group of participants who had a high level of baseline depression: those in the control condition showed no improvement over time, but those in the experimental condition improved significantly (reduction of score from 23.0 at baseline to 14.5 at the end of Week 4).

These results are presented in Figure 6.17.



▲ Figure 6.17 Results of the study: depressive symptoms by condition and baseline depression, BDI = Beck Depression Inventory; CON-Lo = control group, low baseline depression; CON-Hi = control group, high baseline depression; EXP-Lo = experimental group, low baseline; EXP-Hi = experimental group, high baseline

Discussion

What factors do you think are important to consider when studying the impact of social media on mental health? The ones used in the sources that we are discussing here are: the pre-existing tendency to compare yourself with others and the pre-existing level of depression. Is there anything else you think is likely to impact the extent to which a student would be negatively affected by social media use?

ATL Thinking, Research

A similar pattern of results was obtained for one more variable in the analysis: loneliness.

Researchers concluded that "experimentally limiting social media usage on a mobile phone to 10 minutes per platform per day for a full three weeks has a significant impact on well-being" (Hunt et al., 2018, p. 763).

Source 2

Kleemans et al. (2018) studied the effects of manipulated social media photos on body image in adolescent girls.

Participants were 144 girls aged 14–18 years in the Netherlands. They were recruited via snowball sampling. The study was an online experiment that followed a 2×2 design with two independent variables: original versus manipulated photos and low versus high social comparison tendency.

The experimental group was exposed to 10 digitally manipulated photos on a social media site. This included improving the colour intensity and brightness, removing eye bags, wrinkles, and impurities, and making the person look slimmer. All photos were "selfies" of the same teenage girl. The photos varied in their angle and focus: some emphasized the girl's face and hair, others showed her entire body. All photos were given the same number of likes to exclude this as a potential confounding variable.

Participants were instructed to carefully look at all photos, after which they were given a survey.

Participants in the control condition followed the same procedure, but they viewed the original rather than the manipulated photos.

The survey included questions about the participant's (dis)satisfaction with her physical appearance, body shape, weight, attractiveness, and so on. Another scale was used to measure the girls' social comparison tendency—a sample item is "I often compare myself with others with respect to what I have accomplished in life".

Results of the study showed the following:

1. In the experimental condition, girls rated the pictures as prettier and more attractive than did the girls in the control condition.
2. Interestingly, in both conditions (original and manipulated) girls equally disagreed with the statement that the photos paint a better picture than reality.
3. Overall, girls exposed to the manipulated photos reported a significantly lower body satisfaction than girls exposed to the original photos ($r = 0.17$, $p < 0.05$).
4. However, on closer analysis it was revealed that this effect is only significant in the group of girls who had a high score on the tendency to compare themselves with others ($r = 0.25$, $p < 0.01$). By contrast, girls who scored low on social comparison tendency were unaffected by digital manipulation.

Activity

The study of Kleemans et al. (2018) included four groups (2 x 2) and a large number of variables. When findings are listed like this, it may feel overwhelming because there are so many aspects. To more clearly visualize the results of a study, list all the conditions and all the variables in one table and highlight the ones that turned out to be significant. For example, you can achieve this by drawing a 2 x 2 table (original versus manipulated photos, low versus high social comparison tendency). In each cell of the table list all the key variables that were measured in the study: rating of attractiveness of the picture, the belief that photos paint a better picture than reality, and body satisfaction. Now draw connecting lines between cells (and variables) that revealed significant differences.

ATL Research, Self-management

Source 3

Braghieri, Levy, and Makarin (2022) described results of a unique natural experiment: the staggered introduction of Facebook across colleges in the USA in the mid-2000s.

Facebook was created at Harvard University in February 2004 and made available to the general public in September 2006. In between, there was a staggered introduction of Facebook to USA colleges. When a college was granted access to the Facebook network, student sign-up was very rapid and widespread. This creates unique research conditions: we have a number of colleges where all students quickly started using Facebook, each college at a different time. We can look at historical data regarding mental health statistics in these colleges and see if the rapid introduction of Facebook coincided with any changes in mental health status.

The analysis involved data from 775 US colleges. One of the variables was the date when Facebook was rolled out in that college. The other variable came from the National College Health Assessment, a comprehensive survey of mental and physical health of students conducted in the country at that time.

The main conclusion from the study was that the introduction of Facebook at a college was associated with worsening student mental health. The researchers estimated that the effect size of Facebook use on mental health was approximately 22% of the effect of losing one's job. They also performed some sophisticated statistical analyses to estimate what percentage of the increased prevalence of severe depression among students in the last two decades can be attributed to the introduction of Facebook. Their estimate was that Facebook use accounted for 24% of that increase (Braghieri, Levy, and Makarin, 2022).

Claim

Using social media has a negative impact on students' mental health.

Analysis of the claim

The studies presented in the three sources all converge in supporting the claim. However, there are certain limitations that we need to keep in mind that impact the extent to which we can generalize research findings. These limitations are specific to each study.

The study reported in Source 1 provides evidence of a cause–effect relationship because the frequency of media use was manipulated. However, one needs to keep in mind that the manipulation was not entirely "clean". There was no way to prevent students from using social media beyond the allowed 10 minutes. Some participants used more social media than they were allowed to.

The decision to limit usage by 10 minutes per platform, and the study period of four weeks, was somewhat arbitrary. The limit may have been too strict or not strict enough to produce an effect, and perhaps the duration was not sufficient to detect the changes. Only additional studies can clarify this.

The sample used in the study was a convenience sample of students, who owned a mobile phone. Of course, this limits the generalizability of results.

However, the overall conclusion of the study is quite strongly supported—participants show an improvement in self-reported depression and loneliness if they limit their social media use, especially if they had a high baseline level of depression.

The study reported in Source 2 is also an experiment, further strengthening our belief that using social media has a negative impact on students' health. Note: the first independent variable (photos on social media) was manipulated by the researchers by means of random allocation into two groups, but the second independent variable (social comparison tendency) is a pre-existing difference. Therefore, it is used as a potential mediating variable in the causal relationship between exposure to manipulated photos and body image.

An important limitation of this study is that it analysed a very short-term effect of exposure to manipulated photos on body image. It is unclear if this effect is short-lived or accumulates in the longer term.

It would also be interesting to see what the effect is of seeing manipulated photos of someone you personally know on your body image. It could be argued that it is easier to spot digital enhancement in a picture of someone you know in real life, and therefore it will have a weaker effect. However, it could also be argued that we identify more with people from our own network of acquaintances and therefore the effect will be even stronger.

Beyond these considerations, the overall conclusion of the study is well-supported and could be formulated like this: girls exposed to manipulated photos had a short-term reduction in self-reported body satisfaction, but only if they had a pre-existing tendency to compare themselves to other people.

Finally, the study reported in Source 3 provides even further evidence that social media negatively impacts mental health, this time on a much larger scale involving the whole nation. Although the design of this study is not truly experimental, the staggered datasets allowed researchers to be more confident in their cause-and-effect conclusions. The massive datasets used in the study also allowed researchers to quantify the negative impact of social media on mental health: 22% of the effect of losing one's job.

Overall, the studies presented here support the claim, but include some important limitations. For example, they show that the negative impact of social media on mental health only occurs if certain conditions (such as a social comparison tendency) are met.

Discussion

Some of the points we are discussing here link to the typical limitations of research outlined in the "theoretical overview" at the opening of this unit. Which ones? Which of the considerations presented here do you think we can consider "typical" for this entire research area?

ATL Thinking, Communication

6.14 Effects of digital technology on child development

What you will learn in this section

Key learning:

- What makes it challenging for research to find out how the use of devices affects development?
- Analysis of the claim "the use of digital technology has positive effects on the cognitive development of children".

Theoretical overview

Today's children are growing up in a digital world surrounded by devices, and this will influence their development. The problem is it is difficult to establish exactly what the impact is. To make things more complicated, technology is changing rapidly. This means that well-planned research may not be able to keep up with the pace of change to fully understand the effects of the most up-to-date technology.

A common problem that transpires through research in this area is a potential curvilinear relationship between use of technology and developmental outcomes. For example, the association between interacting with touchscreens and a certain developmental outcome may remain positive up until a certain point (e.g., 30 minutes per day), but after that point the relationship may reverse. This could happen because any usage over that amount of time may result in a loss of opportunity to engage in other developmental activities, or due to gradually forming addiction, or some other reason. In any case, it seems very likely that curvilinear relationships exist, but investigating them requires longitudinal studies and multiple measurements, which have rarely been done.

Source 1

Arabiat et al. (2022) attempted a systematic review of the use of interactive technology and child development. They focused specifically on early childhood (up to age seven years) and cognitive development. Interactive technology was defined as on-screen based technology in which the user had to interact with a device to achieve a result. Selection criteria allowed researchers to identify 53 studies.

They found a large amount of inconsistency among studies in the precision with which potential mediating variables were reported. Such factors as amount of exposure to digital technology, intensity, and duration of its use were not always reported in publications. This made it very difficult to come up with aggregated results.

Examples of potential mediating variables are amount of use and the purpose with which technology is used. For example, one study (Axford et al., 2018, as cited in Arabiat et al., 2022) found that the use of a tablet computer was associated with better fine motor skills, but only if the use was limited to applications specifically designed for training such skills. The same study discovered that replacing regular hand use with using a tablet was negatively associated with the development of fine motor skills, but only if it was for more than 30 minutes a day.

Interestingly, only a very small number of studies were longitudinal. There were differences in the overall results between experimental studies and correlational/descriptive studies. For example, many correlational and descriptive studies reported a positive association between interactive technology use and fine motor skills, but many experimental studies reported a negative association or no association.

Researchers concluded that "the result of our review would have been much more meaningful if the dose of exposure to games, videos, computers, or mobile phones had been reported in an informative manner" (Arabiat et al., p. 701).

Source 2

Mayer et al. (2020) studied how reading and writing performance in a nursery school is affected by the writing tool used for practice: pencil, keyboard, or tablet stylus. A stylus on a touchscreen involves a slippery surface that has lower friction than paper, making motor control more difficult.

The sample consisted of 147 German nursery school children aged four to six years. They participated in a seven-week training programme in which they had a series of letter learning games (28 training sessions). Depending on the condition, the training was done by:

- handwriting with a pencil on a sheet of paper
- writing with a stylus on a tablet
- typing on a virtual keyboard on a tablet.

Assessments were performed before training (T1), immediately after training (T2), and four to five weeks after training (T3). Assessments included measures of reading performance, writing performance, and visuospatial skills. Assessment was done on two levels: letters and words. It should be noted that assessments were performed in the same medium as the training. For example, in the keyboard condition in the word writing test, children had to type the word, whereas in the handwriting condition they had to write it by hand.

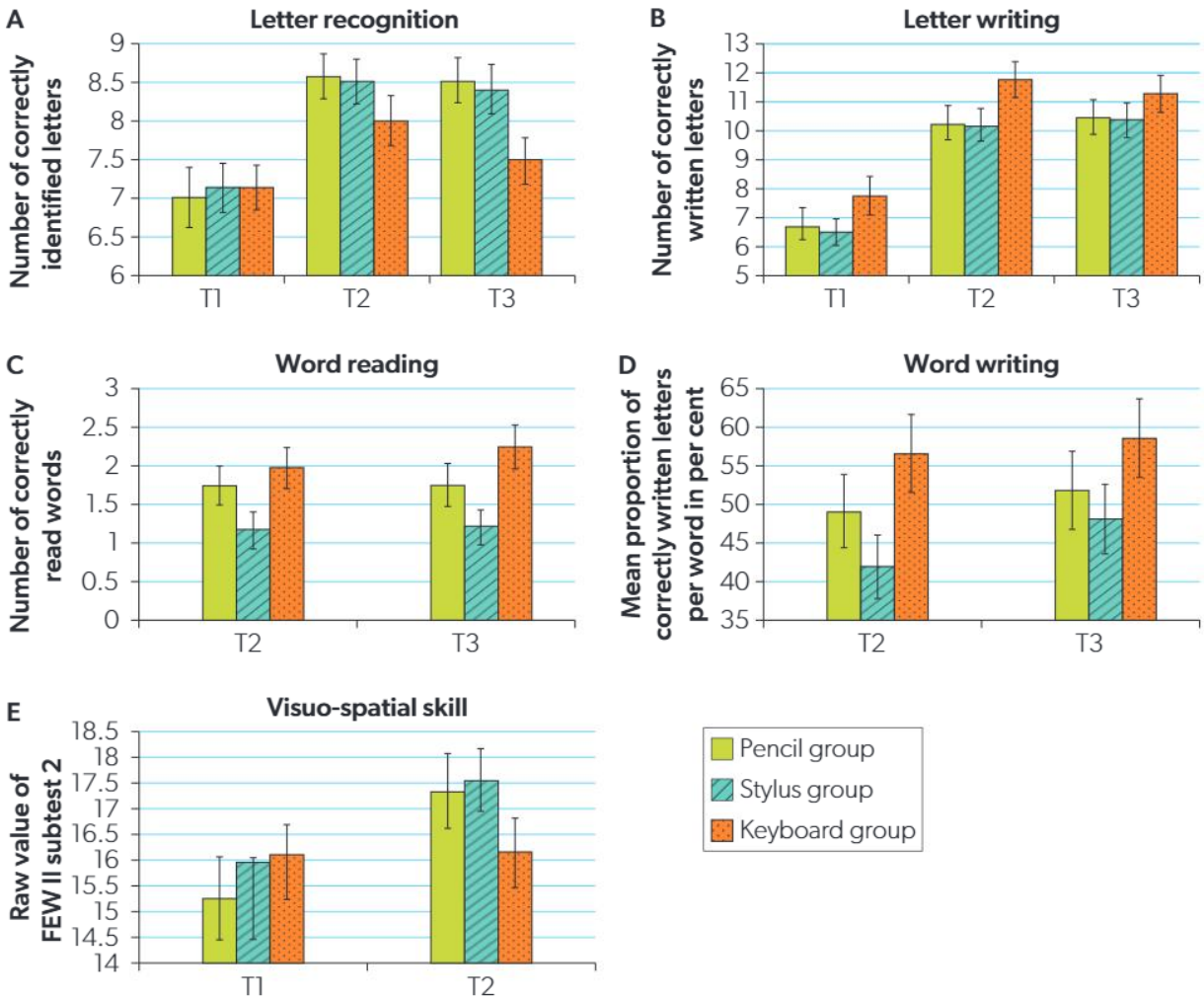
Results of the study showed that the impact on learning was the most positive in the pencil group, followed by the keyboard, and then the stylus condition. More specifically:

- On the level of letters: children in the pencil condition showed superior performance in letter recognition and better visuospatial skills, as compared to using a keyboard. The stylus condition did not differ significantly from either of the other two.

Activity

Draw a schematic representation of Mayer et al.'s (2020) study. There were three experimental conditions, three measurements (T1, T2, and T3), and five variables. Find a way to capture this information as a table, a diagram, or a flowchart. This will make it easier for you to clearly understand the results of the study.

ATL Thinking, Self-management



▲ Figure 6.18 Results of the study. T1 = before training, T2 = immediately after training, T3 = four to five weeks after training. FEW II = a German test of visuospatial skills for children.

- On the level of words: children in the keyboard condition were superior to the stylus condition in word reading and writing. There was no difference between the pencil condition and the keyboard condition.

Descriptive results of this study are given in Figure 6.18.

The overall conclusion from these results reached by researchers was that writing with a stylus on a touchscreen is the least favourable writing tool. Simultaneously, writing with a pencil on paper appears to be the most superior tool, in terms of early literacy and visuospatial skills.

Source 3

Cihak et al. (2016) studied the role of augmented reality in teaching primary school students with ASD (autistic spectrum disorder) to brush teeth. Brushing teeth in this study is taken as an operationalization of performing a "chain task": the kind of mental operation that relies on memory of the correct sequence of steps.

Activity

Review what you know about the interpretation of bar graphs. What do the error bars mean? What does it mean when error bars overlap? Does it matter what value the *y*-axis starts from? Look carefully at the bar graphs in Source 2 and list all the conclusions you can make from them.

ATL Thinking, Self-management

The study included three male students with ASD and a moderate intellectual disability (Allen and Bart, aged six years, and Chris, aged seven years). None of the students at the start of the study could brush their teeth independently.

On the mirror in the bathroom, there was a picture showing a reminder to brush teeth. Students had a tablet. When they pointed its camera at the picture on the mirror and viewed the picture through the screen, the algorithm recognized the picture and automatically displayed a 62-second video clip next to it. The video showed a female student brushing her teeth, with a narrator explaining the sequence of 16 steps in this process.

After watching the video, the students had to brush their teeth. The teacher was there to help them. Researchers recorded the number of steps (out of 16) that the student was able to complete independently without the teacher's help.

The study used a "multiple probe design". All students started the baseline test simultaneously. Then the tablet was introduced to Allen. Researchers kept repeating on consecutive days until Allen reached 100% independent performance. At this point, the tablet was introduced to Bart. After however many trials it was required for Bart to reach 100% performance, the tablet was introduced to Chris.

Figure 6.19 shows the results of the study for all three children. In all three cases, there is a visible "response": an increase in the percentage of independently completed steps immediately or shortly after the introduction of augmented reality.

According to the authors, the study demonstrates that augmented reality is an effective tool for teaching chain tasks to students with ASD.

Claim

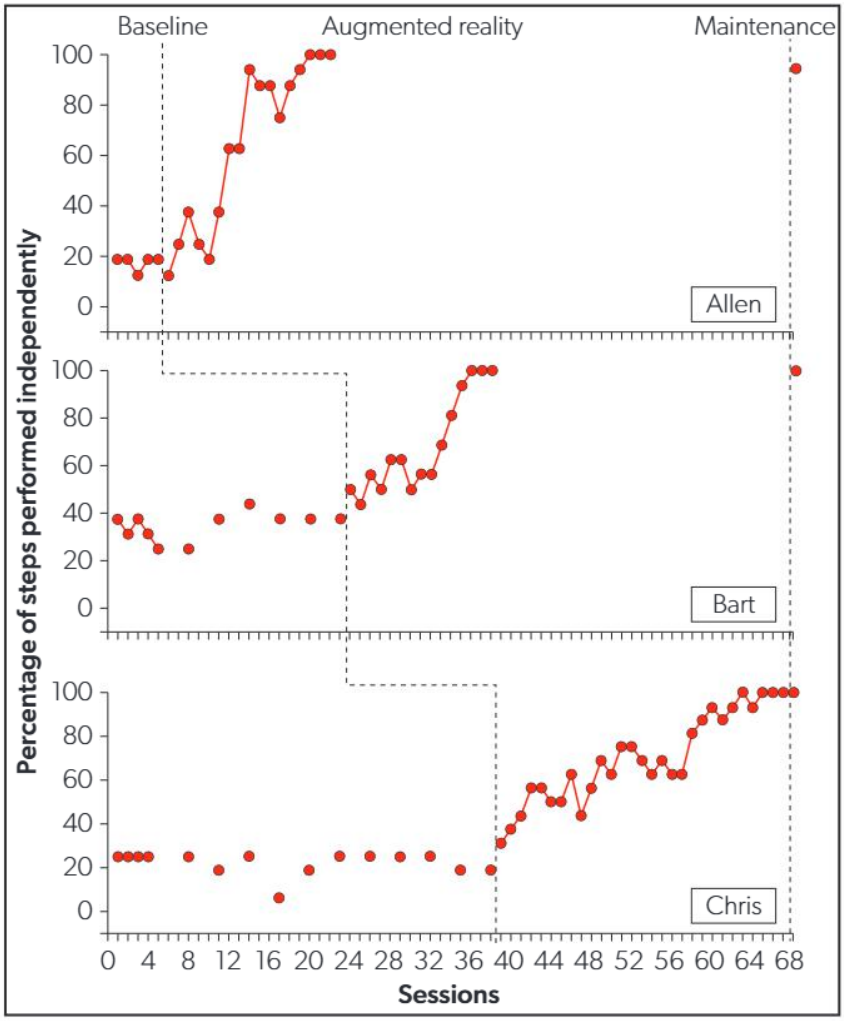
The use of digital technology has positive effects on the cognitive development of children.

Analysis of the claim

The claim is phrased in a misleading way. The studies presented in the three sources provide some empirical evidence in support of this claim. However, the claim ignores the complexity of the issue and the fact that the effects of digital technology on cognitive development may also be negative, depending on circumstances.

A more accurate claim reflecting the findings of these studies would be that the effect of digital technology on cognitive development of children depends on a number of factors in each particular case, such as: the type of technology, the purpose for which it is used, how frequently and intensely it is used, which cognitive process is being studied, and so on.

The study reported in Source 1 is valuable in this respect because it is an aggregation of other research—it demonstrated that the only definitive conclusion that can be made at this point is that the answer to the question cannot be universal and that the drawback of many studies conducted in this area is not properly documenting potential moderating variables.



▲ Figure 6.19 Results of Cihak et al. (2016)

Source 2 serves as a good example of how "inconclusive" research findings may be. There can be two competing theoretical perspectives on the development of early reading and writing skills:

1. One perspective is linked to the cognitive load theory. It suggests that typing is easier. Therefore, it should result in improved performance on writing and reading. This is because memory is less overloaded with the demands of motor control and can be allocated to reading and writing. If this perspective is correct, then the keyboard condition should be superior to the two handwriting conditions.
2. Another perspective (the so-called grounded cognition theory) suggests that physically shaping the form of each letter during handwriting creates an additional motor trace in the child's memory. This facilitates the acquisition of written language. If this perspective is correct, then the handwriting conditions (especially pencil and paper) should be superior to typing.

We can see that results of the study are mixed and do not fully support either of the perspectives. Keyboard training appeared to be superior in some aspects, but handwriting training was superior in others.

Exam tip

One of the descriptors on the Paper 3 rubric is considering different points of view. If you happen to know different theoretical perspectives that could be used to explain the results of a study reported in one of the sources, it could be a great opportunity to discuss such perspectives. Likewise, any different ways of looking at results of the study (or a collection of studies) will serve the same purpose.

Furthermore, the observed differences were rather small, as can be seen from the bar graphs. In addition, the study spanned a relatively short period of time (four to five weeks) and did not probe into more long-term effects of literacy training with the use of technology.

The study of Cihak et al. (2016) (Source 3) seems to have several key limitations (many of which were explicitly acknowledged by the authors).

- First, there is no direct comparison of using augmented reality to other competing conditions, such as video only. It could be the case that using a simple instructional video would have achieved a similar effect, at least for some participants.
- Second, there is a large gap between the very broad construct of "chain tasks" and the very narrow operationalization of brushing teeth. We cannot be certain that brushing teeth is representative of all other chain tasks.
- Third, the study uses a single-subject design. Essentially it represents three single-subject studies published in one article. The graphs are analysed descriptively without using a statistical inferential test between the baseline phase and the intervention phase.

The three boys who took part in the study all showed slightly different patterns of results: for some the improvement was very sharp and quick, for others it was more gradual; the baseline level of proficiency also differed. This might be explained by a difference in the participants' background characteristics that serve as important mediating variables between the use of technology and developmental outcomes.

Overall, it is impossible for one study to provide a conclusive test of developmental hypotheses. Progress in this area heavily relies on our ability to triangulate evidence from multiple studies. Therefore it becomes very important for authors of research to publish their studies in detail and with precision, making it easy to place their research in the bigger context of other evidence accumulated over time.

6.15 Effects of technology on learning and cognition

What you will learn in this section

Key learning:

- The study of distributed cognition: can technology transform the nature of a cognitive process?
- Analysis of the claim "technology is changing our cognitive processes in a way that we become more efficient at processing information with technology rather than without it".

Key term: distributed cognition

Theoretical overview

Studying the role of technology in learning and cognition closely overlaps with the study of how technology affects a child's cognitive development (see the previous section). However, there are at least two angles that can be considered:

1. The study of how technology affects a cognitive process quantitatively—for example, its effectiveness or rate of development.
2. The study of how technology affects the nature of a cognitive process itself. We considered the first aspect in the previous section and we will focus on the second one here.

In the study of "**distributed cognition**" (Hutchins, 1995), learning is considered to be something that is not limited to the mind of one individual. Learning is a characteristic of a complex system composed of individuals and tools and artefacts provided by culture. So preparing for an IB Psychology exam, for example, is a function not of your mind, but of a complex system consisting of your mind, this textbook, your laptop, and the internet. Therefore it is futile to try to understand individual learning outside of this larger distributed system.

From the point of view of distributed cognition, the human and the computer are not two entities interacting with each other, but actually one entity. It is suggested that cognitive processes can "literally spread beyond the boundaries of skull and skin" (Michaelian and Sutton, 2013, p. 2).

Sparrow, Liu, and Wegner (2011) studied the effects of a popular internet search engine on human memory. The starting premise of their study is that the invention of the internet search engine has made it incredibly easy to access collectively stored information—when we do not remember something, we can just search for it online. They conducted a series of experiments, two of which will be described in this unit.

Discussion

Imagine your average school day without the internet. Would you feel deprived of something? Would you say that it would be a feeling comparable to temporarily losing one of the senses (like vision or hearing)?

ATL Social, Self-management

Source 1

In the first experiment by Sparrow, Liu, and Wegner (2011), participants had to answer either easy or hard trivia questions that had a "yes" or "no" answer. After the quiz they were presented with a Stroop task with computer and non-computer terms. A Stroop task is when a series of words is presented to you printed in different colours and you have to name either the word or the colour (for example, the word "red" is printed in a green colour, and you have to say "green").

In the modified version used in this experiment, there were words either related or unrelated to computers, and these words were typed in different colours. Participants had to ignore the meaning of the word itself and name the colour. The authors hypothesized: if participants get a difficult trivia question that they cannot answer, they will immediately think about searching for the answer online and therefore they will be primed to think about computers. This will affect their performance on the Stroop task because they will spend some time processing the meaning of computer-related terms, making their reaction slower.

Results of this experiment revealed that after receiving difficult quiz questions, participants took a longer time to process (i.e., to name the colour of) computer-related words than general words. The average reaction time to a computer-related word in the difficult quiz condition was 712 ms, compared to 591 ms for general words ($p < 0.003$). Researchers interpret the result as suggesting that knowledge questions, and especially difficult ones, make us think about computers. Therefore, we find it harder to disengage from the meaning of computer-related words on a Stroop task.

Source 2

In their second experiment, Sparrow, Liu, and Wegner (2011) investigated how people remembered information when they knew that they could look it up online later. Participants read 40 trivia statements such as "An ostrich's eye is bigger than its brain" or "The space shuttle Columbia disintegrated during re-entry over Texas in February 2003". After reading each statement, participants were asked to type them into the computer. After reading and typing all 40 statements, they had to write down as many as they could remember. Half of the participants were told that what they typed would be saved on the computer. The other half believed that it would be erased. Additionally, half of the participants in each group were asked explicitly to try to remember the information and the other half was not. In other words, it was a 2×2 independent measures design with two independent variables, as shown in Table 6.6.

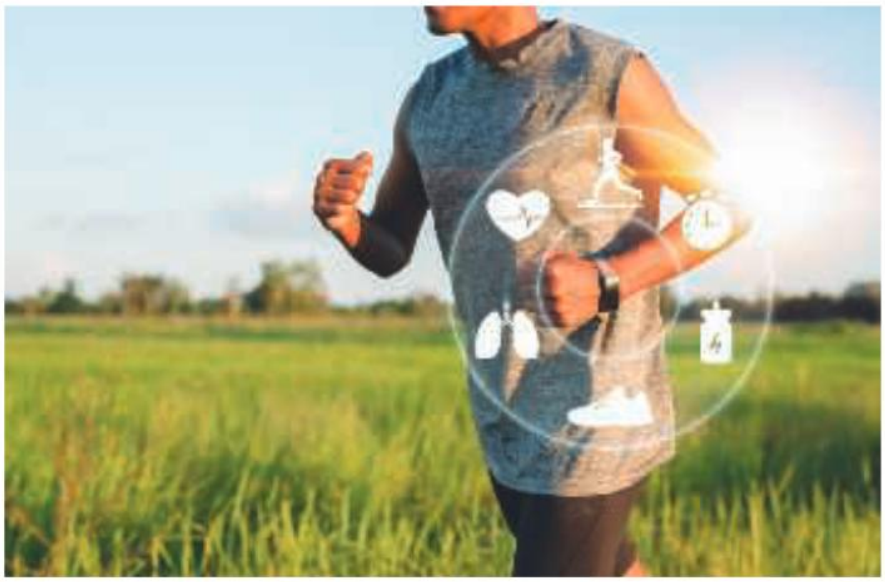
	Information will be saved	Information will not be saved
Told to remember	Group 1	Group 2
Not told to remember	Group 3	Group 4

▲ Table 6.6 Groups used in the study

Results revealed a significant effect of the erased versus saved manipulation: participants who believed that what they type would be erased demonstrated

a better memory of the trivia facts than those who believed that the typing was saved. As researchers explain it, participants did not make the effort to remember when they thought they could look the information up at a later time.

Interestingly, the explicit instruction to remember their information versus no such instruction did not have any significant effect in this study. "Participants were more affected by the cue that information would or would not be available to them later, regardless of whether they thought they would be tested on it" (Sparrow, Liu, and Wegner, 2011, p. 477).



▲ Figure 6.20 A wearable activity tracker

Source 3

Nelson et al. (2020) studied embodiment of wearable technology. Embodiment is the phenomenon when technology is subjectively experienced as a part of one's body, as an extension of one's physical abilities, cognition, and even self.

In the study, 12 participants used a wearable activity tracker for nine months. The study was a series of three semi-structured interviews (every three months). The questions remained the same, but researchers delved deeper each time and clarified if anything had changed.

Researchers used a thematic analysis to identify themes emerging from the data. Two researchers always carried out the coding procedures independently and the level of agreement between them was calculated (it was sufficiently high).

The following themes emerged in the study:

- Adjustment to the wearable tracker. Most participants reported that they had to adjust their normal routines slightly to learn to live with the device (e.g., press a button first thing in the morning to indicate that they have woken up).
- Engagement/reengagement. Over the nine months, participants reported having periods of relatively heavy use alternated with periods of inactivity.

- Wearability. Most participants described the device as comfortable, as something that did not bother them.
- Awareness of the wearable tracker. After several weeks, most participants said that they hardly noticed the device ("it is just there").
- Extending the mind. Some participants described "missing" the device when they were unable to use it. These participants started perceiving the ability to know their activity levels as their own additional cognitive ability, and they felt deprived when they could not use it for some reason.

Overall, the study suggests that participants did indeed experience some extent of embodiment.

Claim

Technology is changing our cognitive processes in a way that we become more efficient at processing information with technology rather than without it.

Exam tip

There are always multiple ways in which you can analyse and evaluate a research study, depending on what your focus is. Your focus may be determined by the exact claim that you are evaluating. In this section, you will see that the claim makes a statement that is more generic than that which can be directly supported by the data, so we have chosen to focus on generalizability and operationalization. It does not have to be this way and your approach can be different.

Analysis of the claim

The three studies presented here all support the idea that technology has a profound impact on the way we cognitively process information. In summary:

- Source 1 suggests that we react to difficult questions by thinking about computers, so our first reaction is to use technology to find out the answer.
- Source 2 suggests that our memory works differently if we know that information will be digitally recorded and accessible at a later time.
- Source 3 suggests that wearable technology over the course of time may become embedded and feel like an extension of our own cognitive processes.

We also know from previous research that many other effects of technology on cognition have been registered. This is not unexpected either. We know that neuroplasticity depends on the activity of the brain. Therefore, the kind of experiences we engage with have to determine the structure of synaptic connections in our brain. When we are surrounded by digital technology from an early age, we engage in these experiences and therefore our neural networks must be reshaped accordingly.

It has been suggested that cognitive processes should not be viewed as something confined to the limits of an individual's mind, but rather should be viewed as a function of a more complex system—a system that is distributed among the individual, tools, and technology being used, other people, and information systems. This idea is known as "distributed cognition" and it is different, for example, from the study of human–computer interaction (HCI) because the latter still assumes that the human and the computer are two separate entities that interact with each other.

Having said that, it is important not to make generalizations that are not justified. Therefore, we should remember several key limitations.

First, there is always a theoretical leap from observed empirical data to the generalized inference. The idea that the human and the computer are two separate entities that interact with each other is a generalized inference. So is the idea that the human and the computer comprise a single entity, a distributed system. The findings of any research study will not directly support either of these generalizations, and in fact it is probably impossible to put them to a direct test. Many findings will be consistent with both these generalizations. What we observe in research is that the use of technology influences cognitive processes in a variety of ways, but whether or not the human and the computer are separate entities, and whether or not the relationship between them is "symbiotic", are interpretations. The claim we are analysing does not make any sweeping generalizations like this and is quite grounded in observed data.

Second, any given research study can only investigate a limited aspect of the relationship between technology and cognition. For example:

1. The study in Source 1 only considers a person's reaction time when processing words related to computers, after being asked a different question. Actual technology is not even involved in the study.
2. Source 2 reports a study that looked at how people remember information when they believe that it would be accessible later (versus no such belief). One could imagine a similar study in which information would be recorded on paper and put in a drawer, rather than recorded in a computer. Would a similar effect have been observed? There was no comparison with such a "paper condition" in the study.
3. The study in Source 3 focuses on participants' experiences around wearing an activity tracker and their responses may be interpreted as containing signs of "embodiment", but there is quite a leap from here to suggesting that our cognitive processes are being reshaped.

Nevertheless, qualitative studies are important in studying the role of technology in learning and cognition because they allow us to understand people's inner experiences and interpretations, something that would never be captured by "impersonal" methods such as the experiment.

All three studies considered here only looked at relatively short-term effects of technology (or thoughts about technology) and none of them measured physiological changes in the brain. Overall, all the sources support the conclusion that cognitive processes such as memory can adapt to the use of technology in certain ways. However, the more general conclusion that cognitive processes are "changing" due to technology can only be based on a much larger body of research including longitudinal studies with the use of brain imaging.

Activity

To further exercise your analysis of the extent to which the given sources support the given claim, practise the opposite exercise. Look at the available sources of evidence and formulate a claim that is fully supported by these sources. It cannot be too broad, but neither should it be too narrow.

How would you rephrase the given claim to make it more aligned with the available evidence?

ATL Thinking, Self-management, Communication



7 Internal assessment

7.1 Overview of the task

The Internal Assessment (IA) in IB DP Psychology is a research proposal to investigate a real-life problem that is of significance to a certain population of interest. There are a few things you need to keep in mind to better understand the nature of this task.

- First, it is a research proposal and not an actual conducted study. This means that you are less limited in what you can suggest. For example, you could plan a longitudinal study that runs over many years, or you could plan a study involving expensive equipment. However, the proposal also needs to be realistic from both the practical and the ethical points of view. In the real world, all researchers submit a proposal for approval to an ethics committee. In addition, researchers submit proposals to get funding for their research. Be realistic in what you are planning: the costs should be justified. Formulate your proposal in a way that is likely to be approved. It is unlikely that you will receive funding for performing expensive fMRI scans on thousands of participants if it does not present any clear advantages over using a simple survey.
- Second, one of the key purposes of this task is to assess your "methodological decision-making". This means deciding on what kind of study would be most effective and appropriate in a given situation. Your study does not have to be perfect (no study is perfect), but you should demonstrate that you have thought the proposal through and that the suggested parameters of your study (e.g., the sample or the data collection tool) are thoroughly justified. To do this, you have to be aware of the advantages and disadvantages of various interview techniques, various experimental designs, and so on.
- Third, the IA task highlights the fact that psychological research can solve practical problems and improve the quality of people's lives. For this reason, your proposal should be designed to target a specific practical problem in a specific real-life context.

The task can be broken down into four major steps corresponding to the four major sections of the final report and the four assessment criteria: Introduction, Research methodology, Data collection, and Discussion.

We will consider these sections one by one, explaining the assessment requirements and also giving some recommendations along the way.

7.2 Criterion A: Introduction

There are three components that should be included in the introduction:

1. A description of a real-life problem and an explanation of its impact on a population of interest.
2. An aim that is relevant and clearly focused on the population of interest.
3. The findings and key conclusions of two pieces of relevant research.

The IB DP Psychology Guide states that the "population of interest" can be selected from within the school community, a local community, or a national community. Accordingly, the practical problem addressed in the research proposal should be of relevance to the population of interest.

Here are some examples to give you an idea about what could be used as a population of interest and a real-life problem:

- (National community). A large number of refugees from a neighbouring country who are experiencing trauma and need psychological help. Not enough help is available due to limited resources, so online peer support groups are created. You are wondering whether such groups actually help, and how effective they are.
- (Local community). A local rural community has a high percentage of elderly people. Younger people are moving to big cities and older people are staying behind. It has been reported that this may lead to feelings of loneliness and loss of the sense of control over one's life. You want to investigate how older people experience being left alone and what the potential protective factors are.
- (School community). A school has recently had incidents of discriminatory behaviour among students. You are wondering what changes can be made to the curriculum or methods of instruction to reduce discrimination.

Once you identify the real-life problem, it should be clearly described. How exactly does it affect the population of interest? What evidence do you have that this is experienced as a significant issue?

Once this is done, formulate the aim or the research question of your proposed research study. Remember: you are not solving the problem itself. You are proposing a research study that may potentially contribute to the solution of this problem. The aim of the study or the research question should reflect that. For example, the aim could be:

1. To investigate the perceived effectiveness of online peer support groups among refugees.
2. To identify potential protective factors that older people report as helping them to cope with the stress of being alone.
3. To test the hypothesis that a specially designed kindness curriculum in the seventh grade will be effective in reducing instances of discrimination among school students in the short term.

Finally, it is also required that your introduction includes findings and key conclusions of two pieces of relevant research. It is definitely recommended that

you do some background reading and you read more than two pieces of research, but select the two that are most relevant in the context of your proposal.

Avoid citing research that is too generic. To give you a better sense of this, Table 7.1 has some examples of relevant versus "too generic" research, linked to the example aims 1–3 on the previous page.

Example of research proposal (1–3 above)	Example of relevant background research	Example of background research that is too generic or less effective
Perceived effectiveness of online social support groups for refugees	A study that found that a large number of refugees in that community use online social support groups A study that found that online peer support groups are perceived to be almost as effective as face-to-face therapy, although the study was conducted in another location on a similar vulnerable population	A study that compares the effectiveness of online therapy (tele-therapy) and face-to-face therapy in general A qualitative study that summarizes the main themes emerging from an interview with refugees about what they find most stressful in their current situation
Protective factors for the elderly to cope with loneliness	A study that suggested that some older people in this community are more resilient and adjust to being alone more effectively than others A study that suggested a possible link between daily exercise and social activity, and perceived quality of life in older people	A meta-analysis that investigates possible factors of loneliness A study that found that feelings of loneliness are related to symptoms of depression
Effects of kindness curriculum on discrimination among school students in the short term	A study suggesting that this kindness curriculum has a positive effect on children's prosocial behaviour A study that found that children who are more prosocial are also less likely to engage in acts of discrimination	A study that tested a theory of discrimination to try to understand why discrimination occurs (e.g., social identity theory) A study of bullying conducted in the same school

▲ Table 7.1 Relevant background research

You are not required to report on the procedure of these studies unless this is important in the context of your own research. The focus should be on presenting the main findings. It would be great if you could read the full article, but in most cases the abstract of the article should be enough to understand the main finding of the study. Abstracts are always openly available even if you need to pay for the article itself.

7.3 Criterion B: Research methodology

This section of your written research proposal should include the following:

- a justification of the choice of research method
- an explanation of the choice of procedure
- an explanation of ethical considerations in conducting the study.

The IB DP Psychology Guide states that you can choose the research method for the IA from the following list:

- Experiment (true, quasi-, or natural)
- Interview (structured, semi-structured, unstructured, or focus group)
- Observation (naturalistic or controlled, overt or covert, participant or non-participant)
- Survey/questionnaire.

The strength of experiments is that they make possible cause–effect inferences, so we conduct them when we try to uncover the cause of some behaviour. Observations are great for studying the behaviour of a group of people holistically when it is the whole range of behaviours you are interested in. Interviews are good for studying things that cannot be registered by observing behaviour, such as people's experiences and interpretations. Surveys/questionnaires rely on self-reports, but make it possible to register a wide range of parameters (both behaviours and experiences) in a standardized and comparable way.

The main point here is that you should choose the research method that would be most appropriate in the context of your proposal, and explicitly explain and justify why this is the case.

When you identify the research method to be used in your investigation, use its full characteristics, including subtypes—for example, "overt participant naturalistic observation". These features should be justified as well. Why did you choose an overt observation rather than covert? Why naturalistic rather than controlled?

What is meant by procedure is "all decisions about planning and carrying out the investigation, including, but not limited to, sampling technique, sample characteristics, design (if experimental), setting and process" (IB DP Psychology Guide, 2025).

Sampling technique and the characteristics of the sample are aspects of the procedure that do not depend on which research method you have chosen: whatever the method, there will be a sample.

Some other aspects of procedure will be unique depending on the chosen research method. Examples include, but are not limited to:

- Experiment: experimental design, measures taken to control confounding variables, method of allocation of participants into groups, blinding (e.g., single-blind, double-blind)

- Observation: setting (where and under what circumstances observation took place), frequency, targets of observation, recording, how the aim of observation was explained to participants
- Interview: group or individual setting, how rapport was established between the interviewer and the interviewee(s), recording and transcript, where the interviews took place
- Survey/questionnaire: how it was administered to participants, the instruction, setting (individual, group, online), how participants were motivated to respond, how much time it took to complete.

Exam tip

A note on surveys/questionnaires and correlational studies in the Internal Assessment

You must have noticed in the requirements for the IA that correlational studies are missing from the list of methods to be chosen from. Does it mean that a research proposal for a correlational study cannot be written?

Our suggestion is that a correlational study can be proposed within the survey/questionnaire category. The focus will be on constructing the survey/questionnaire as the data collection tool. However, you do not have to limit yourself to using the survey as a research method in its own right (i.e., the quantitative descriptive method).

As part of your research proposal, you can plan that the variables measured with the help of your survey/questionnaire will be further investigated by correlating them with each other or with some other measures. You can also plan for various groups of respondents to be compared in terms of the variables measured by the survey/questionnaire.

For example, your research proposal might be a cross-cultural study where you are giving the same questionnaire to participants with different cultural backgrounds, and compare the results. Technically, this is not a quantitative descriptive study because you are making inferences rather than simply describing a population. If you are strict about categorizing this study within one of the research methods, it would probably be a correlational study. However, for the purposes of the IA, you can propose this research study if there is sufficient focus on the survey/questionnaire being used for data collection.

The final element in this section is an explanation of ethical considerations in conducting the study. Ethical considerations should be relevant, and they should be explicitly linked to the investigation. This means that you should avoid making statements that are too generic and can be related to any study whatsoever. For example, "measures were taken to protect participants from harm" or "informed consent was obtained from all participants". Indicate what measures they were exactly and why these particular measures were important in the context of your investigation. Explain what exactly was included in the informed consent and why it was important to gain participants' consent to these things in your study.

7.4 Criterion C: Data collection

The following are the required elements that should be included in the written research proposal:

- the choice of one data collection tool to measure behaviour relevant to the aim of the investigation
- an explanation of the decisions made when creating the data collection tool
- a discussion of factors that could potentially affect your data collection and findings.

The IB DP Psychology Guide contains the requirement that the data collection tool should contain a minimum of five items (e.g., for a questionnaire there should be at least five questions).

In the way IA requirements are formulated, the line between a research method and a data collection tool is not very clear. Let us clarify what could be used as a data collection tool, using our hypothetical examples.

Example of an investigation	Research method used in the investigation	Examples of data collection tools that could be used
Effects of kindness curriculum on discrimination among school students in the short term	Experiment: one group of students received the experimental curriculum and the other group received the standard curriculum	Questionnaire: a specially designed self-report measure of discriminatory behaviour administered to all participants Implicit association test (IAT): a procedure measuring automatic, implicit prejudice Observation: a structured covert observation in which students are observed during school break with a specially designed observation checklist
Perceived effectiveness of online social support groups for refugees	Survey/questionnaire: a group of refugees fill out an online survey with questions related to participation in online social support groups: frequency of participation, perceived effectiveness, perceived barriers, self-efficacy, self-disclosure, etc.	The survey/questionnaire itself. You can use either a survey in which each question is a measurement in its own right, or a questionnaire in which questions are combined into scales and total scores are calculated per scale
Protective factors for the elderly to cope with loneliness	Semi-structured interview: a purposive sample of older participants from a specific community are interviewed individually	An interview schedule: a list of questions to be asked and some rules regarding follow-up questions

▲ Table 7.2 Examples of data collection tools

The data collection tool in an observational study, an interview, and a survey/questionnaire is quite straightforward: it can be the observation checklist, the interview schedule, and the survey itself, respectively. However, with the experiment you can use various different tools to measure the dependent variable, and these tools may actually include surveys/questionnaires, interviews, and observations, among others.

Exam tip

A note on using a survey/questionnaire as the research method in the Internal Assessment

A survey/questionnaire may be either a research method in its own right or a data collection tool used within some other research method (see Chapter 1: Research methods and data interpretation).

We can use a survey/questionnaire as a data collection tool within an experiment (the research method). At the same time, a survey is also one of the research methods for the IA, and a survey is a data collection tool within that method!

To clarify this, consider the following:

- A research method serves the purpose of testing the hypothesis or answering the research question. If the goal of research is to investigate if one variable is the cause of another, we use an experiment. If the goal is to uncover any other relationship between variables, we use correlational studies. If the goal is to provide an in-depth description of the behaviour of a group of people, we might use observation. If we want to study someone's subjective experiences, we might prefer an interview.
- In contrast, a data collection tool serves the purpose of measuring behaviour, but it does not directly answer the research question. In short, a data collection tool helps us make a measurement, while a research method helps us make an inference. This is why, for example, an observation may be used as a data collection tool to measure participants' behaviour within an experiment. The experiment itself is a research method designed to test if a specific independent variable has an influence on this behaviour.

A survey/questionnaire can be used as a research method in its own right. In this case we refer to it as the quantitative descriptive method. An example would be a study of prevalence rates of depression or a study of patterns of social media use in a group of teenagers. Typically, results of such a study would be reported as percentages: for example, "25% of the respondents said that they spend more than three hours a day on social media", "15% of the study's population reported having moderate-to-severe symptoms of depression", and so on.

When we use the quantitative descriptive method, unlike experiments and correlational studies, we do not test any hypotheses about relationships between variables (either causal or not). We just describe a population of interest. However, unlike qualitative studies (e.g., observations and interviews), this description is constructed using pre-defined parameters and comes in the form of numbers.

Most frequently, surveys/questionnaires are used as a measurement (data collection) tool within another research method. For example, you might give participants two questionnaires and correlate their results, in which case it would be appropriate to categorize it as a correlational study. Or, you might conduct an experiment and use a questionnaire to measure the dependent variable.

The second component required for this section of the written research proposal is an explanation of the decisions made when creating the data collection tool.

There must be a reason why your data collection tool is the way it is. Typically, the reasons are either methodological (you construct the measurement tool in a certain way to avoid potential biases or confounding variables) or ethical (you construct the data collection tool in a certain way to avoid doing any sort of harm to participants). An example of the former is the decision to have an equal number of directly and inversely scored items in a questionnaire in order to minimize the influence of acquiescence bias. An example of the latter is including questions about the participant's friends in an interview in order to avoid directly asking the participant personal and sensitive questions, making them uncomfortable.

Finally, you are also required to include a discussion of factors that could potentially affect your data collection and findings.

No matter how well you control your variables and how clearly you anticipate possible problems, it is impossible to protect against everything. You cannot design a perfect study. However, you can be aware of the possible weaknesses and the exact limitations of your study, and how they influence the decisions that you can and cannot make based on your result.

Here are a few examples from the IB DP Psychology Guide of possible factors that could affect the results of a study:

- Participant variability
- Practice effects (fatigue/boredom)
- Order effects
- Researcher bias
- Response bias
- Validity of the data collection tool
- Controlling variables
- Demand characteristics/social desirability bias.

As you can see, these are very diverse suggestions, so the list is not exhaustive and you can bring up many other considerations as long as they are relevant to your investigation.

7.5 Criterion D: Discussion

There are three key requirements to this section of the written research proposal:

1. A discussion of the potential findings of the investigation and the implications for policy/practice.
2. A discussion of how researcher bias may have affected the investigation.
3. A discussion of one additional method for investigating the same topic.

The first point is a discussion of how the potential findings of your proposed research study will have wider implications for solving the practical problem that you identified at the start of the investigation. In other words, this is about how the findings of the study can be practically applied. Be specific rather than generic, but at the same time avoid suggesting practical applications that do not directly follow from your research.

To clarify how to find the appropriate balance, let us turn back to our examples and consider what would be a relevant practical implication versus one that is too vague or irrelevant.

Example of research proposal	Example of relevant implication for practice	Example of implication for practice that is too generic or less effective
Perceived effectiveness of online social support groups for refugees	Could establish an NGO that works to support the mental health of refugees in an evidence-based way by curating online peer support groups and ensuring that these groups have all the features that have been identified by respondents as impactful.	"Results of this investigation could be used to help refugees better cope with their difficult life situations."
Protective factors for the elderly to cope with loneliness	Could negotiate with policy-makers to spend some budget on creating opportunities for older people. For example, if it is discovered that community events are perceived as important protective factors, we could pilot a number of new community events organized by the municipal authority.	"Knowing protective factors that support older people in overcoming their loneliness may increase their quality of life because if they feel less helpless, they will engage in activities more and be healthier."
Effects of kindness curriculum on discrimination among school students in the short term	If shown to be effective, kindness curriculum can be gradually rolled out across year groups and schools. This should be done step-by-step, continuing to register both short-term and long-term effects on behaviour, because it is possible that the short-term effects may be negated or even reversed in the long term.	"If the kindness curriculum works, we should introduce it across all schools and year groups because this will help us reduce discrimination on a large scale. Reduced discrimination in school may positively affect many social practices, making people more tolerant to each other."

▲ Table 7.3 Relevant implications for practice

The second point is a discussion of how researcher bias may have affected the investigation. The purpose of this part of your written proposal is to assess how well you can reflect on your own role as a researcher in the research process, and on the potential biases and inaccuracies that you may bring (which is unavoidable). This is the important process of a researcher's reflexivity. It is especially prominent in qualitative research where the researcher serves as the main tool of measurement (through their interpretations of participants' behaviour or responses), but also essential in quantitative research.

As stated in the assessment rubric, "one or more examples" are required for this part of the report, so it is not about the quantity but how relevant your discussion is to your investigation.

Some examples of the types of researcher bias that you could discuss include, but are not limited to:

- How has your personal history influenced the choice of topic?
- How might your personal values or beliefs influence your interpretation of the data or your conclusions?
- How could the fact that you were aware of your experimental hypothesis have influenced the behaviour of your participants?

The third point is a discussion of one additional method for investigating the same topic. This has been added to recognize the fact that no research method is perfect and there is always something to be gained from method triangulation. Different research methods allow us to use different angles on the same research topic, thus reaching a deeper understanding of it. The expectation is that you propose one additional method and explain how exactly it is expected to contribute to the deepening of insights that can be derived from the investigation.

7.6 Ethical considerations in the research proposal

Ethical considerations in the IA are somewhat relaxed due to the fact that it is a hypothetical research study that you are not actually conducting.

Having said that, the proposal must still adhere to key ethical guidelines. You cannot propose a study that would not normally be accepted by an ethics committee in real life. Studies that breach ethical norms of conducting psychological research without justification will not be accepted.

For these reasons, there are some specific topics and types of investigation that are not allowed in the IB DP Psychology Internal Assessment. These include, but are not limited to:

- Studies in which participants are required to ingest something as part of the experimental procedure (although investigating self-reported dietary choices is allowed)
- The use of non-human animals.

In addition, we need to be mindful of teachers and examiners who will read the report and might find it difficult to read about certain sensitive topics. Therefore, there are additional topics that are not allowed. These include, but are not limited to:

- abuse (physical or emotional)
- self-harm
- serial killers or acts of torture.

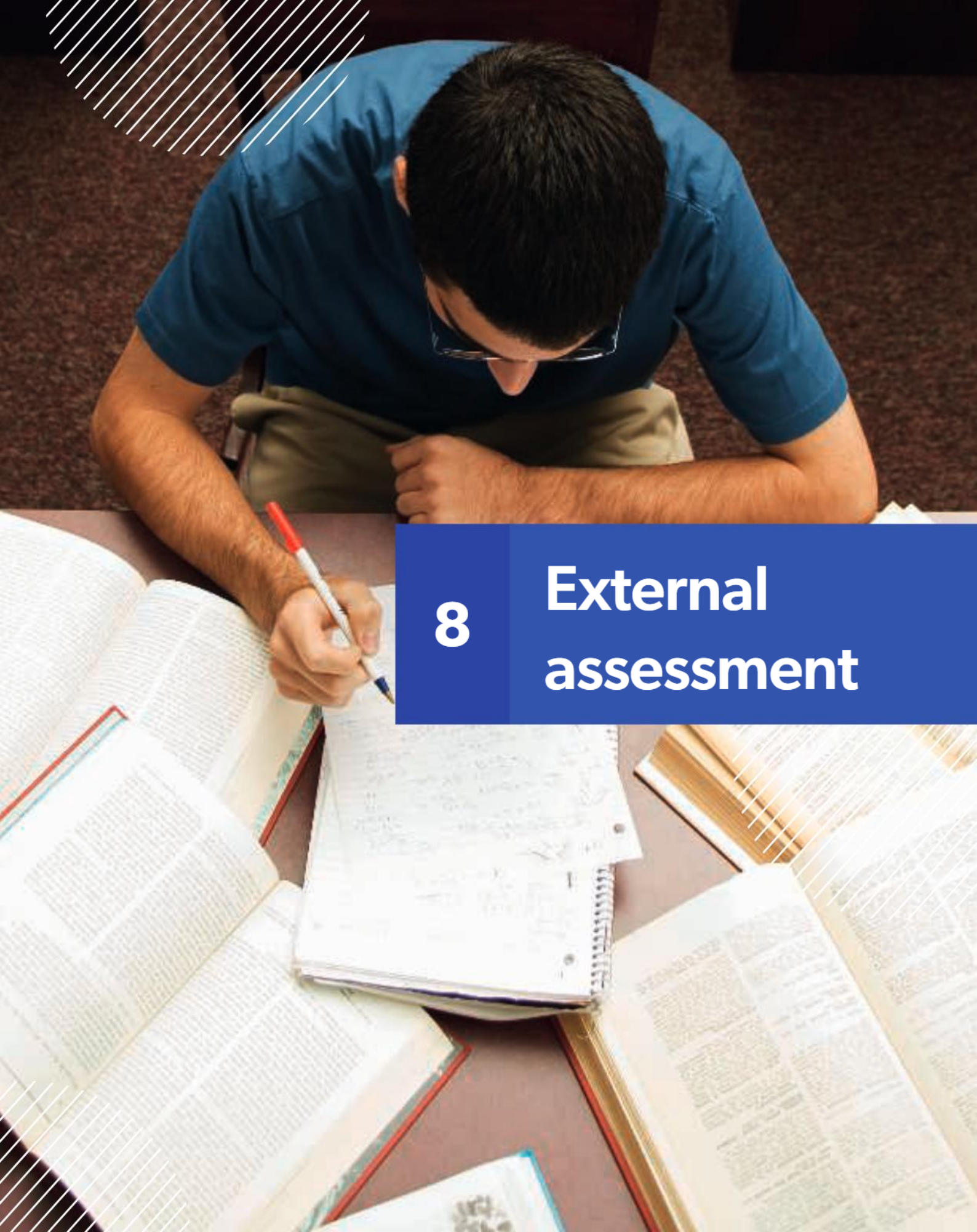
Remember: always get approval from your teacher before proceeding with the research proposal.

7.7 The presentation of the research proposal

Your report must contain references and appendices.

- References are not assessed, but must be included to meet the requirements of academic integrity. If you have references to research studies that are not written in the same language as your research proposal, you are required to translate the title of this research in the references to assist examiners.
- Appendices are not assessed either, but they are important for the examiners to fully understand your investigation. The crucial element that must be included in your appendices is the data collection tool. You are not required to provide completed material beyond the data collection tool itself. Other elements to be included in the appendices are tables, graphs, or pictures that are relevant to the report.

Your written report must not exceed 2,200 words. Typically, examiners are instructed to stop reading after this limit is reached, and to give you marks based on what they have read. Importantly, references and appendices are not included in the word count.



8 External assessment

8.1 Structure of exam papers

Overview

External assessment in IB Psychology includes timed assessment papers that are taken under exam conditions and marked by external examiners.

There are three papers—Paper 1, Paper 2, and Paper 3. The first two papers are subdivided into sections, but the individual sections are not timed. The time limit is for the whole paper and students decide themselves how they allocate time to sections within this. Paper 3 is only for higher level (HL) students. The weight denotes the importance of the paper in calculating the final score for IB Psychology. Students also have the Internal Assessment (IA) which is responsible for 20% of the marks at HL and 30% of the marks at SL. The IA was considered in Chapter 7; in this chapter we focus on the external assessment components.

Paper	Duration	Weight SL	Weight HL
Paper 1 Section A	1.5 hours	35%	25%
Paper 1 Section B			
Paper 1 Section C			
Paper 2 Section A	1.5 hours	35%	25%
Paper 2 Section B			
Paper 3 (HL only)	1.75 hours	N/A	30%
Internal Assessment	*20 hours*	*30%*	*20%*

▲ Table 8.1 External and internal assessments in IB Psychology

Paper 1 Section A

Paper 1 Section A will include two compulsory short-answer questions designed to assess knowledge of theories and content from two of the three content areas (biological, cognitive, sociocultural). The IB DP Psychology Guide contains tables that list all the content that can be asked in exam questions in Paper 1A. For convenience, they are combined in Table 8.2.

Biological approach	Cognitive approach	Sociocultural approach
Brain imaging techniques	Classical conditioning	Models of acculturation
Chemical messengers	Operant conditioning	Cognitive dissonance
Neurotransmission	Schema theory	Compliance techniques
Neuroplasticity	Cognitive models	Conformity
Localization of function	Dual-processing theory	Cultural dimensions
Genetic inheritance	Cognitive load theory	Enculturation
Diathesis–stress model	Anchoring bias	Social identity theory
Animal research/animal models	Confirmation bias	Social learning theory
Biological reductionism	–	Emic approach
–	–	Etic approach

▲ Table 8.2 List of content for Paper 1 Sections A and B

Questions in Paper 1 Section A will ask students to explain or describe a psychological idea or a theory and provide an example. Examples may be a relevant, real or hypothetical research study or situation.

It is important to note that your example may be a research study, but it does not have to be. Knowledge of details of research studies will not be assessed, so do not worry, for example, if you do not remember how exactly the researcher manipulated the independent variable or how exactly they recruited participants in their sample. The focus of assessment will be on your understanding of the content listed in Table 8.2, as evidenced by your explanation and relevance of the supporting example.

Each question in Paper 1A is worth 4 marks, for a total of 8 marks for the two questions. Looking at the total duration of Paper 1 and the maximum marks possible, this translates into approximately 10 minutes per question—and so a short, focused response.

Paper 1 Section B

Paper 1 Section B will include two compulsory short-answer questions. You will be asked to apply your knowledge of psychological content (biological, cognitive, and sociocultural, as listed in Table 8.2) to new situations or unseen scenarios. The situation or scenario will be provided as part of the question. The situation will be embedded in one of the four contexts (Learning and cognition, Human development, Health and well-being, Human relationships).

What is being assessed is your ability to apply your knowledge to new situations, the so-called transfer skills. It is through transfer that understanding is truly demonstrated. Responses to Paper 1A could be prepared in advance and memorized (although this is not a recommended strategy). Responses to Paper 1B cannot be memorized because you cannot predict which situation will be given. However, if you understand the psychological idea or theory, you will have no trouble explaining how it applies to a new situation.

Each of the two questions in this section is worth 6 marks, for a total of 12 marks. A simple calculation based on the duration of the entire Paper 1 suggests that you should spend approximately 15 minutes per question.

Paper 1 Section C

This paper will include two concept-based extended response questions, each from a different context (Learning and cognition, Human development, Health and well-being, Human relationships). You will need to choose one of these two questions.

The content for this paper is embedded in the four contexts. In fact, each content point is a unit in the relevant chapter in this book (Chapters 2–5). For convenience, Table 8.3 outlines the full list of content for Paper 1C.

Context	Group of content		
Learning and cognition	**Thinking and learning**	**Cognitive processes**	
	Cognitive biases	Biological factors in cognitive processes	
	Schema theory	Cognitive models	
	Conditioning (classical and operant)	Cultural factors in cognitive processes	
	Dual processing model	Environmental influences on cognitive processes	
	Social learning theory	Potential for improving a cognitive process	
Human development	**Models of development**	**Development of self**	
	Brain development	Attachment	
	Sociocultural factors in development	Enculturation of social norms	
	Stage theories and continuous models	Peer influence	
	Theory of mind	Role of childhood experiences	
Health and well-being	**Mental health disorders**	**Health problems**	**Prevention and treatment**
	Biological explanations	Stress and health	Biological treatment for one disorder
	Cognitive models	Social learning and health	Psychological treatment for one disorder
	Cultural differences	Prevalence of health problems	Prevention and/or treatment for one health problem
	Environmental factors	–	–
Human relationships	**Group behaviour**	**Interpersonal relationships**	
	Acculturation	Chemical messengers	
	Conformity	Cognitive explanations	
	Compliance techniques	Communication/language	
	Cultural dimensions	Strategies for improving relationships	
	Social identity theory	–	
	Social learning	–	

▲ Table 8.3 List of content for Paper 1C

Command terms in this paper will be at the so-called AO3 level (Assessment Objective 3: high-order thinking):

- Evaluate
- Discuss
- To what extent.

Each exam question will be formulated as a combination of three elements: a key concept, a content unit, and a context. Each question will identify the context area, include one content point, and link to one of the six concepts (perspective, causality, bias, measurement, change, and responsibility). To understand how exam questions will be formulated, consider the examples in Table 8.4.

Question	Explanation
The dual processing model can provide insight into how people learn and process information. Discuss the role of measurement in this model.	The concept is measurement, the content is the dual processing model, the context is Learning and cognition.
It has been claimed that biological treatment of mental disorders addresses symptoms but not the cause of the problem. Evaluate this claim with reference to one or more mental disorders.	The concept is causality, the content is biological treatment for one disorder, the context is Health and well-being.

▲ Table 8.4 Example questions for Paper 1C

This section is worth 15 marks and is assessed against a holistic rubric. Based on the number of marks and the total duration of Paper 1, it may be suggested that students should spend around 40 minutes on writing this response.

Paper 2 Section A

This paper is based on the four class practicals that you are expected to complete as you are studying the four contexts. The IB DP Psychology Guide sets out the expectation as follows:

- Learning and cognition: Experiment (true, quasi-, or natural)
- Human development: Observation (naturalistic or controlled, overt or covert, participant or non-participant)
- Health and well-being: Interview (structured, semi-structured, unstructured, or focus group)
- Human relationships: Survey/questionnaire.

The paper is built upon a set of four questions that are "semi-static" in nature. This means that the questions will be formulated similarly by following a similar template, but the exact research method and the key concept referenced in the question may change.

This is best understood by looking at the four questions one by one.

1. The first question will ask you to describe how a research method (experiment, interview, observation, or survey) was applied in one of your class practicals.
2. The second question will ask you to explain one of the key concepts (perspective, causality, bias, measurement, change, or responsibility) in relation to one of your class practicals.
3. The third question will ask you to compare and contrast the use of the research method in one of your class practicals with an alternative research method (both methods will be given).
4. The fourth question will ask you to design a research study given an alternative research method (that will be given) to investigate the same topic as you investigated in a given class practical.

If this still sounds abstract, have a look at the full sample paper in Unit 8.3. Your teacher should also have access to some specimen papers provided by the IB.

Paper 2A is worth 20 marks in total: four marks each for the first two questions and six marks each for the last two.

Paper 2 Section B

The requirement in Paper 2 Section B will be to discuss a research study with regard to two or more concepts (the question will specify a larger list of concepts from which you can select).

For example: Discuss the following study with reference to two or more of the following concepts: bias, causality, measurement, and/or responsibility.

The study itself will be provided as part of the stimulus material. The study may be experimental or non-experimental.

The question is worth 15 marks.

Paper 3 (HL only)

Paper 3 is focused on data analysis and interpretation, but also the HL extensions (culture, motivation, technology). It will include a research booklet with several "sources" that will present findings of real or hypothetical studies. All the questions will refer to these sources. You will be asked to analyse data and findings from these sources.

The sources will include both quantitative and qualitative research. They may be experimental or non-experimental. They may simply state the findings of the study or present more detailed data using tables or graphs. All of the sources will be aligned to one of the HL extensions with one of the contexts—for example, "motivation" in Human development, or "culture" in Learning and cognition.

There are four questions in Paper 3 and they are semi-static in nature. This means that the questions will follow the same template, but there could be slight variations. This is best understood when considering the four questions one by one:

1. The first question will ask you to interpret results of a study presented in graphical form. For example, you could be asked to interpret a graph provided in one of the sources in the resource booklet. Or you may be asked to explain a problem with the graph that influences your ability to interpret it. The purpose of this question is to assess specific skills related to reading and understanding results of research. [3 marks]
2. The second question will ask you to analyse the findings from one of the sources and state a conclusion. The purpose of this question is to assess your understanding of the connection between research data and theoretical conclusions that can be derived from it. [6 marks]
3. The third question will be based on qualitative research (i.e., one of the sources in the resource booklet that describes the findings of a qualitative study). You will be asked a question related to research considerations in this study, such as its credibility, bias, or transferability. For example, you may be asked how the researcher could improve the credibility of the findings, how the researcher could avoid bias, or to what extent the findings can be seen as generalizable to other populations or contexts. [6 marks]
4. Finally, the fourth one will be a synthesis question—you will be required to use at least three of the given sources, but also your background knowledge, to discuss the validity of a claim. The claim will be provided as part of the question. [15 marks]

To get further insight into what to expect from the exams, see the full sample paper provided in Unit 8.3.

8.2 General exam tips

Paper 1 Section A

Where to find content

In this IB Psychology Course Book, the content required for Papers 1A and 1B is embedded within the four contexts (Learning and cognition, Human development, Health and well-being, Human relationships). For convenience we have highlighted this content in a special SAQ content icon. These features are scattered throughout Chapters 1–5 and appear where relevant in the context of a larger discussion. You can also find the full list of such features in Table 1 on page viii.

This content is not to be perceived as separate from the rest of the course. There is a reason why it is not presented as a separate unit. The icons highlight the material that will be sufficient for you to answer questions in Paper 1A effectively. However, you are invited to support this with any example. You could use examples provided in this book (both within the in-text features and outside) or you could use your own examples.

What to consider an example

The IB states that examples may be both real and hypothetical, and that they could be a research study, but they do not have to be. What matters here is that your example is relevant and actually helpful to explain the psychological idea or theory.

In this book, we have focused on providing research studies as examples to support the psychological ideas and theories. If you do use a research study, remember that specific details of research studies are not assessed. You do not need to be concerned about memorizing the details of the study. Just make sure that you understand both the study and how it is related to the psychological content (theory or idea), and that you remember the idea behind it. Assessment in this paper is not about how well you have memorized the content, but rather how well you have conveyed your understanding.

Paper 1 Section B

Predicting the question

Unlike Paper 1A, Paper 1B is fairly unpredictable. This is because you will be given an unseen situation or scenario and asked to explain how your knowledge of a specific psychological idea applies to this situation.

Practising

Since it is impossible to predict the question, the best strategy to prepare for this exam is to understand the content and practise its application to all sorts of scenarios as you are going through the course. Always try to find links between what you already know and any new material you are learning. These skills are transferable. For example, if you really understand what brain imaging techniques are and how they are used, and if you can apply this understanding to a few different situations, then you can probably do the same for any situation at all.

Staying focused

Remember: Sections 1A and 1B are designed to assess distinctly different skills: the ability to retrieve relevant knowledge from your memory (1A) and the ability to transfer this knowledge to an unseen situation (1B). Retrieval from memory has already been assessed in Section 1A, so the focus of 1B is on transfer. Start applying to the situation as soon as possible, preferably straight away. You can explain the relevant aspects of the theory as and when it becomes relevant. If a certain aspect of the theory is not relevant to this particular situation, it makes sense not to mention it. It may be compelling to explain a psychological idea in detail before you apply it to the situation, but you should avoid this approach.

In short, let your response be driven by the situation rather than your pre-existing knowledge, and let it be improvised rather than rehearsed.

Paper 1 Section C

Links to other papers

In some sense, preparation for this paper may be seen as central to the entire course. This is because the skills and knowledge needed to successfully complete many other exam components are acquired as you are studying the content that is explicitly assessed in Paper 1 Section C. For example, the content necessary for 1A and 1B is embedded in the material relevant to the contexts (Learning and cognition, Human development, Health and well-being, Human relationships). As you study for 1C, you also "automatically" study for 1A and 1B. Furthermore, you study the contexts through a conceptual lens. This is reflected in the concept-based nature of questions in Paper 1 Section C. Understanding the concepts and practising their application to a range of specific research studies across different contexts prepares you for analysing how such concepts apply to any unseen study—the skill that is targeted in Paper 2 Section B.

Knowing what can and cannot be asked

It is important to remember that exam questions cannot include any key words other than the ones listed in the prescribed content. For example, you may be asked about "bias" in social identity theory, but you cannot be asked specifically about "publication bias" or "researcher bias". Similarly, you can be asked about social identity theory, but you cannot be asked specifically about in-group favouritism or out-group discrimination.

Exhausting all combinations

As you know, questions in Paper 1C will link one of the six concepts (perspective, causality, measurement, bias, change, responsibility) to one of the 38 areas of study (see Table 8.3 on page 507). This may suggest that you need to prepare for $6 \times 38 = 228$ possible questions. However, thinking about it this way is counterproductive and not advised. Here's why:

- First, even within the combination the exact way the question will be worded cannot be predicted. For example, one may think of several possible ways to ask about "bias in social identity theory".

Exam tip

A note about the possibility of changes in IB exams

Sometimes the IB slightly modifies the rules of formulating exam questions after the course has already been launched. This may mean that revisions to the IB DP Psychology Guide are published, or it may take the form of a separate clarification document that serves as an addendum to the guide. This does not happen very often and the changes are usually insignificant, but sometimes they can still be consequential.

For this reason, always use your teacher as the source of the most up-to-date information about exams and assessments. If the IB has made some small changes to how exam questions can be asked, your teacher will know. Keep an eye also on the digital subscription resources that supplement this book—any updates will be immediately shared there.

- Second, not all combinations are reasonable. For example, there is a lot to explore when we think about causality in relation to biological factors in mental disorders, or the concept of perspective in relation to cultural dimensions. However, it is perhaps less meaningful to speak about problems of change in relation to cognitive models, or problems of responsibility in relation to schema theory. You need to focus on learning what is important and fundamental for understanding psychology and avoid focusing too much on less important aspects just because they represent a plausible combination of concepts and content. We can operate on the assumption that exam setters want to ask about things that are crucial and fundamental.
- Third, conceptual understandings are transferable. If you understand the role of causality in a few key topics, you probably understand the role of causality in most of the remaining topics also. The same conceptual understandings are reiterated throughout the course in different shapes or forms. What is being assessed is your ability to recognize that the new situation or example is similar to something that you have already encountered. On exam day, it is very likely that you will have to bring up arguments and make connections that you have never practised, and this is expected. The ability to "improvise" a meaningful response that is relevant to the question is valued more highly than the ability to reproduce something you have heard in class or practised beforehand.

Paper 2 Section A

The first question on this paper asks you to describe how a specific research method was applied in one of your class practicals. This is a simple response that can be planned in advance. The purpose of this question is to give examiners some background information to assist them with marking your responses to the other three questions. Although the question is only worth 4 marks, the way you answer it may affect the examiner's understanding of your other responses. This is the only information they have about what you have done for your class practicals. Stay brief and relevant to the question, and try to provide all information that is crucial for understanding your study. This typically includes the aim (what you were investigating), the method you used, the sample (who participated), the procedure (what exactly you did, what the participants were required to do in your study). There is no need to include any findings. Neither analysis nor evaluation are necessary—you just need to describe.

The second question asks you to explain how one of the concepts is related to your class practical. There are many angles that can be explored and there is no prescribed way to approach this. Links will be accepted as long as they are relevant and justified. It is a good opportunity for you to practise conceptual thinking. After each class practical, it is recommended that you engage in an activity where you explicitly explain how it links to each of the concepts (either in writing or in a conversation with a partner). The question is only worth 4 marks, so stay brief and to the point.

The third question invites you to compare and contrast the research method used in your practical with one alternative research method. Remember: both methods will be given. For exam purposes in Paper 2A, when we say "research method",

we mean one of the following six broad methods: experiment, correlational study, interview (including focus group), observation, survey/questionnaire, case study. The type within that (e.g., field experiment, structured observation) will not be specified and it is up to you to choose. Also, requirements of the command term "compare and contrast" are to speak about both similarities and differences.

The fourth question on this paper asks you to design a research study using a given alternative method. For example, if you investigated a topic using an observation, the question may ask you to design a study that would investigate the same topic using an experiment. Note: the reference is to the "topic" being investigated, not to a research question or, for example, hypothesis. This is because not every method is appropriate for investigating every research question. Suppose you used an interview to study how people experience the loss of a job after being employed for over 20 years. You cannot use an experiment to address the same research question: experiments cannot give us an insight into people's experiences. However, if you think about "job loss" as a broader topic, you could design an experiment to investigate a related, but different research question within that. You could investigate how participants' judgements of a fictional employee are affected by their belief about how long that employee had been working at the company before they lost their job.

Paper 2 Section B

In this paper, you will be given an unseen research study and asked to discuss it through the lens of two or more concepts. You will not be able to choose from all six concepts—you will be given a shorter list of concepts to choose from instead.

Concepts link to content in many ways, and there are always multiple relevant angles you could explore. The very nature of concepts is that they are very broad, and that they manifest in various specific instances. For example, a study of the effect of chemical messengers on behaviour could be discussed through the lens of the triad "biological–cognitive–sociocultural". You could bring up the idea of biological reductionism and challenges associated with building a holistic picture of reality through research. All this would be relevant to the concept of perspective. However, you could also not mention any of that at all. If there exist alternative explanations of the results of the study, you could discuss these alternative explanations. That would be linked to the concept of perspective as well, because different theoretical orientations dictate how we interpret research findings.

The best way to prepare for this paper is practice! Look at the given list of concepts and choose the ones that are most relevant to the study. Spend some time thinking about all the possible ways in which these concepts link to the given study, choose a few relevant ones with the view of staying focused, but also effectively demonstrating your knowledge, and do not be afraid to improvise. Remember: critical thinking is the major part of the exam rubric in this assessment component, so let your response be argumentative (driven by analysis and evaluation) rather than descriptive.

Paper 3 (HL only)

The first question in this paper is very straightforward. It is an assessment of your ability to understand charts and graphs. The response is expected to be short and relevant. Stay focused on the question and be specific in your response.

The main idea behind the second question is to test how well you understand the transition from research findings (data) to theoretical conclusions. It is important to be clear as to what conclusions can and cannot be made from the given data, and avoid conclusions that are not fully justified by evidence. Spend some time studying the data that is provided, and think about various limitations. Read the given conclusion very carefully, questioning each word of it. In case you believe that the given conclusion is not fully supported by the given data, try to re-formulate it in a way that is more in line with the evidence (for example, a common suggestion is to remove the word "influences" because cause–effect inferences are not fully justified, and replace it with the more neutral phrase "plays a role in").

The third question is focused on qualitative research. You will be asked about credibility, bias, or transferability in one of the studies from the resource booklet. The main advice here would be to remember that qualitative research methodology is different from quantitative methodology: it uses different terms to denote similar ideas and it has different standards for judging the quality of a study. Avoid applying the principles, standards, and terminology of quantitative research when discussing a qualitative study. Here are some specific examples:

- It would be wrong to use the term "validity" when analysing qualitative research. This is a term from quantitative research methodologies. An equivalent in qualitative research is credibility (and transferability).
- The idea that qualitative research is limited in its generalizability to larger populations is not wrong. However, by itself it is a superficial argument that misses the point, that qualitative research does not pursue this kind of generalizability in the first place.
- The researcher in a qualitative study takes a role that is much less impartial as compared to, for example, an experiment. In some sense, the researcher is the measurement tool. This implies certain requirements to researcher reflexivity and triangulation.

See Chapter 1 for a more detailed discussion of quantitative and qualitative research methodology.

The fourth question is "holistic". It asks you to evaluate a given claim based on multiple sources of evidence: at least three of the sources provided in the resource booklet as well as your own knowledge on the topic. This is where your study of HL extensions comes in.

There is no content in HL extensions that needs to be memorized. Indeed, it is impossible to know what the given claim would be. When you study HL extensions, focus on the following:

- Understand the challenges and limitations that seem to be common in this entire area of research.

- Practise evaluating various claims. As with many other parts of the IB Psychology course, the skills assessed in this component are transferable. If you have been successful at evaluating somewhat similar claims in somewhat similar contexts, you will do well.

As in the second question, remember to read the claim very carefully and analyse every word of it. The following are just some of the possibilities of what could be concluded based on your analysis:

- Yes, the claim is fully supported by the evidence that I have from these sources and from my own background knowledge.
- Yes, the claim is fully supported, but we need to be careful about the following limitations: …
- The claim in its current form is not justified. It would be justified if we changed the wording by making it less categorical, like this:…
- Some parts of the claim are supported, but other parts are not.
- The claim is supported, but only within the following constraints: …
- The claim is not supported.

See Chapter 6 for further examples of analysis of psychological claims. We have broken this chapter down into parts, to cover all possible combinations of the three HL extensions (culture, motivation, technology) and the four contexts. We have also provided a variety of claims and diverse approaches to analysing them.

8.3 Sample paper

This section contains one full sample set of exam papers in IB Psychology.

Psychology
Higher level and standard level
Paper 1

1 hour 30 minutes

Instructions to candidates:

- Section A: answer both questions
- Section B: answer both questions
- Section C: answer one question
- The maximum mark for this examination paper is **[35 marks]**

Section A

Answer **both** questions.

Cognitive approach to understanding human behaviour

1. Explain cognitive load theory with reference to one example. (4 marks)

Sociocultural approach to understanding human behaviour

2. Describe the role of cultural dimensions in one human behaviour. (4 marks)

Section B

Answer **both** questions.

Question 1 (Context: Health and well-being)

The mayor of a small town is calling a meeting to discuss the problem of childhood obesity and unhealthy eating habits. It was reported by one of the schools that children are less and less likely to have a substantial healthy breakfast in the morning. Schools in the town have a long lunch break in the afternoon, and children after a certain age are normally allowed to leave school premises for lunch. There are many food outlets available in close proximity to the schools. Perhaps related to the issue is the fact that many such food outlets broadcast popular TV shows on screens.

As the problem is becoming serious, the mayor seeks to understand why students' eating behaviours are becoming less healthy, and what can be done to remedy the situation. You have been invited as an expert in psychology. Please provide an explanation based on operant conditioning. (6 marks)

Question 2 (Context: Human development)

Anthropologists have been investigating a small ethnic community where children are encouraged to think very carefully before they make any kind of decision, and trial-and-error learning is discouraged. You are the researcher that is responsible for organizing an expedition to that ethnic community and investigating the process of cognitive development in their children. You are supposed to use the emic approach in your investigation. Explain what an emic approach would mean in this context. (6 marks)

Section C

Answer **one** of the following questions:

1. Discuss the role of **measurement** in **psychological treatment** for one or more **mental disorders**. (15 marks)
2. Evaluate one or more possible sources of **bias** in constructing a **cognitive explanation** of **human relationships**. (15 marks)

Psychology
Higher level and standard level
Paper 2

1 hour 30 minutes

Instructions to candidates:

- Section A: answer all the questions
- Section B: answer the question
- The maximum mark for this examination paper is **[35 marks]**

Section A

The following questions refer to your participation in the class practicals. Answer all the questions.

1. Describe how you used an observation in your class practical, including the aim and procedure. (4 marks)
2. Explain the concept of responsibility in relation to your class practical that used an observation. (4 marks)
3. Compare and contrast the use of observation in your class practical with a survey/questionnaire. (6 marks)
4. Design an experiment to investigate the same topic as you investigated in your class practical that used an observation. (6 marks)

Section B

Answer the question:

Discuss the following study with reference to two or more of these concepts: bias, perspective, measurement, and/or change. (15 marks)

Zak, Stanton, and Ahmadi (2007) studied the effects of oxytocin on altruism and generosity. Altruism is the act of helping another person at a cost to oneself, and generosity is offering more than that other person needs or expects. For example, giving a homeless person $1 is altruism, but giving $100 is altruism combined with generosity.

Participants (68 male subjects) were split randomly into two groups. Depending on the group, they were infused by a nasal inhaler with either oxytocin or normal saline. Oxytocin is a hormone that is known to be involved in feelings of attachment, bonding, and trust. Saline is a harmless salt solution with no physiological effect.

Participants in this study played two decision-making games—at the start Player 1 is given $10 and Player 2 is given nothing. After this:

- In the Dictator Game, Player 1 can then give some of this money to Player 2, and Player 2 has to accept. This game was used as a measure of altruism.
- In the Ultimatum Game, Player 1 is asked to offer some of this money to Player 2. Player 2 has a choice: they can accept the offer (in which case both partners get the money) or they can reject the offer (in which case neither of the players gets anything). Player 2 can reject the offer if they consider it unfair. This game was used as a measure of generosity.

Firstly, all participants had to make choices for both players. For example, in the Ultimatum Game they were asked to indicate the minimum offer that they would accept as Player 2 as well as the offer that they would make as Player 1. After this, pairs of participants were randomly formed by a computer, pay-offs determined based on their prior responses, and participants paid.

Results showed that the minimum acceptable offer indicated by Player 2 in the Ultimatum Game was unaffected by oxytocin—it was on average $3 in either group. However, the amount of money offered by Player 1 in this game was significantly higher in the oxytocin condition (mean $4.86) than in the placebo condition (mean $4.03). By contrast, the offer given by Player 1 in the Dictator Game was not affected by the hormone. As predicted, it was lower than in the Ultimatum Game (oxytocin group mean $3.77, placebo group mean $3.58), but the difference was not significant.

Researchers concluded that oxytocin does not influence altruism but does influence generosity.

Psychology
Higher level
Paper 3

1 hour 45 minutes

Instructions to candidates:

- Answer all the questions
- The accompanying **Paper 3–resource booklet** is required for this examination paper
- The maximum mark for this examination paper is **[30 marks]**

Answer all the questions.

The sources in this examination have been collated to assess the claim that culture determines the way we cognitively process information.

1. Refer to **source 1** in the accompanying resource booklet.
   Explain **one** issue that limits the interpretation of the data in **source 1**. (3 marks)
2. Refer to **source 2** in the accompanying resource booklet.
   Analyse the findings from **source 2** and state a conclusion linked to the claim that culture determines the way we cognitively process information. (6 marks)
3. Refer to **source 3** in the accompanying resource booklet.
   Discuss how the researcher could avoid bias in the study described in **source 3**. (6 marks)
4. Using at least three of the sources in the resource booklet (**sources 1–5**), and your own knowledge, answer the following question:
   To what extent can we conclude that culture determines the way we cognitively process information? (15 marks)

Psychology
Higher level
Paper 3–resource booklet

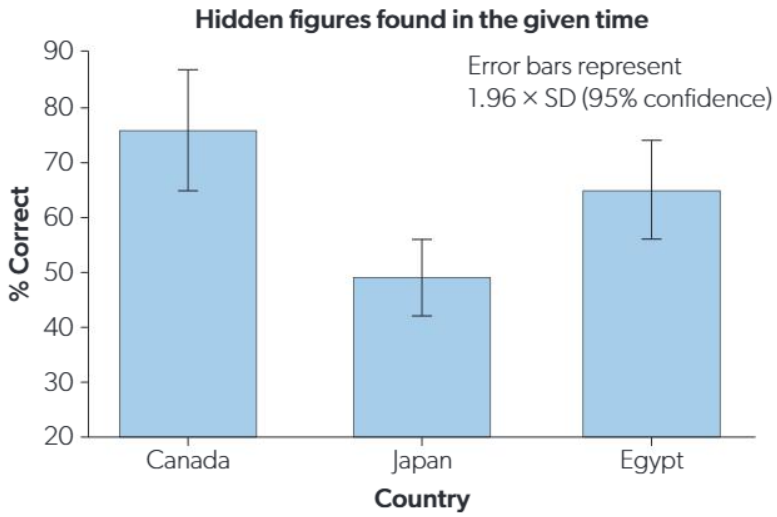
Instructions to candidates

- Do not open this booklet until instructed to do so.
- This booklet accompanies Paper 3.

The sources in this examination have been collated to assess the claim that culture determines the way we cognitively process information.

Source 1

Graph showing results of a study in which participants from different countries searched for hidden figures in an image within a limited time.



Source 2

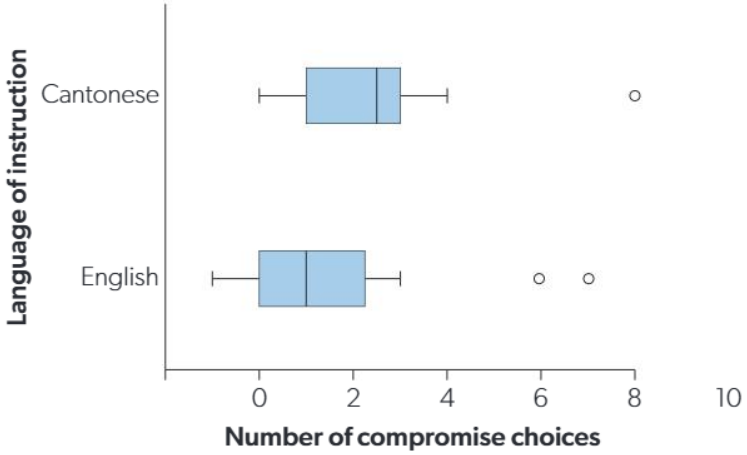
A sample of students in Hong Kong were given a set of decision-making tasks. The set included 10 hypothetical scenarios, each having three options to choose from. One of the options was always superior to the others in parameter A but inferior in parameter B, another option was the opposite, and a third option was always a compromise. An example of a scenario is:

You are shopping for a smartphone and there are three options: (1) a smartphone with a long-lasting battery, but a low-quality camera, (2) a smartphone with a high-quality camera, but a short-lived battery, (3) a smartphone with a medium-duration battery and a medium-quality camera. Which one would you choose?

All students were bilingual—equally fluent in English and Cantonese. They were randomly split into two groups. The instructions and procedure were identical, but the study was conducted in different languages in the two groups.

The main measure was the number of scenarios (out of 10) where the participant chose the compromise option. Results of the study are presented in the following table and plot.

	English	Cantonese
Mean	3.5	4.45
Median	3.0	4.5
Standard deviation	2.13	1.80
Min	1	2
Max	9	10



Source 3

Researchers were interested in investigating a group of older Mexicans living in the USA who had started to experience cognitive decline due to progressing Parkinson's disease. They were specifically interested in how participants experienced what they were going through in the context of keeping ties with their native culture, but also being acculturated to the new context.

They conducted semi-structured interviews and transcribed the responses. The transcripts were then subjected to thematic analysis by two independent

researchers. They arrived at the categories (themes) independently, compared them, then coded the responses again together. They also counted the number of times each theme was encountered in the responses.

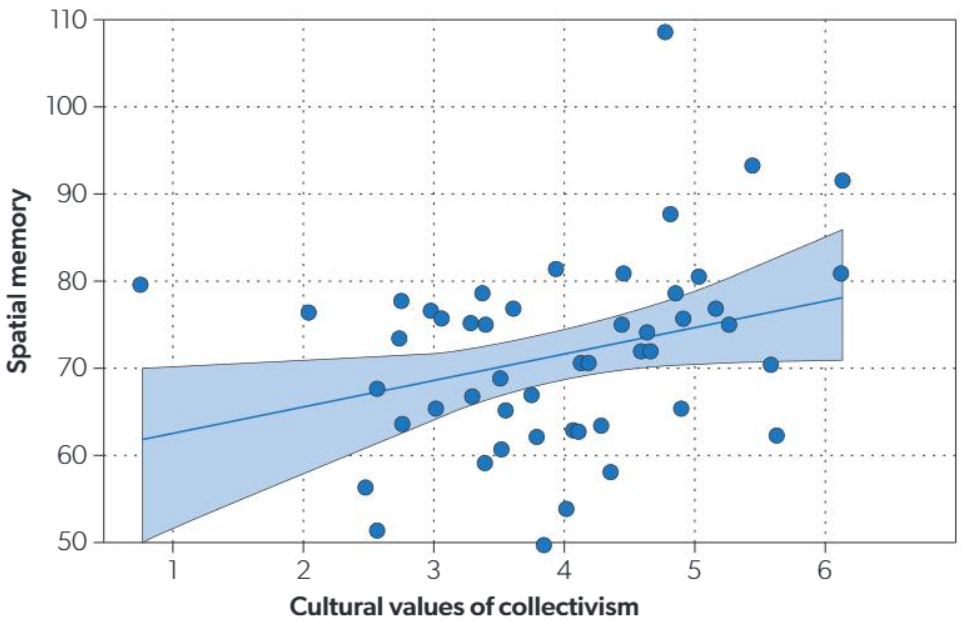
The three main themes that emerged from the analysis were: (1) Feelings of loneliness and disconnect from their home culture, (2) Helplessness associated with the sense that they have no control over their future, (3) Support from relatives and other close people.

Additionally, some participants mentioned that English language proficiency is an important part of their experiences. Researchers recruited an additional sample of participants—whose English proficiency was low—and repeated the procedure with them. Analysing these transcripts led to the emergence of an additional theme: (4) Feeling of belonging to the host culture.

Source 4

Researchers investigated the correlation between spatial memory and cultural values of collectivism in a group of international students from a major university. The sample included 50 participants. They obtained the following results:

$$r(48) = 0.30,\ p = 0.034$$



Source 5

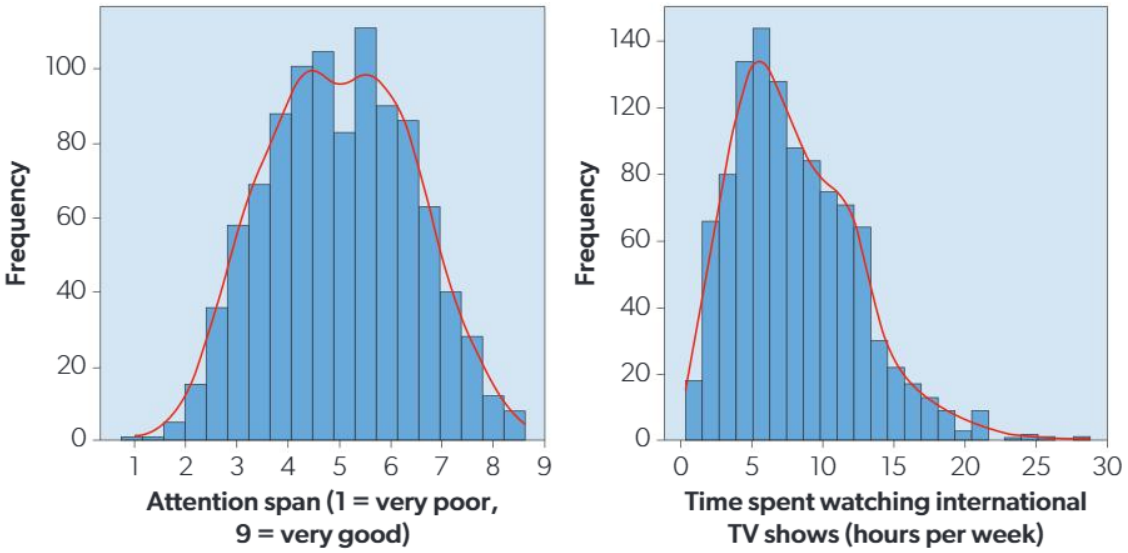
Researchers conducted an online survey among university students in Lesotho (a country in southern Africa). Two of the questions on the survey:

(1) How would you rate your attention span (from 1 = very poor to 9 = very good)?

(2) How many hours a week do you typically spend watching international TV shows?

Participants had to use a slider to answer both questions, which is why fractional values are possible.

Results of the survey are presented below:



Bibliography

Chapter 1: Research methods and data interpretation

American Psychological Association (2012). *Guidelines for ethical conduct in the care and use of nonhuman animals in research*. Washington DC. Retrieved from https://www.apa.org/science/leadership/care/guidelines.aspx

Bailey, J. (2008). "An assessment of the role of chimpanzees in AIDS research." *Alternatives to Laboratory Animals, 36(4)*, 381–428.

British Psychological Society (2012). *Guidelines for psychologists working with animals*. Leicester. Retrieved from https://www.bps.org.uk/news-and-policy/bps-guidelines-psychologists-working-animals-20212

Campbell, D.T. (1957). "Factors relevant to the validity of experiments in social settings." *Psychological Bulletin,* 54(4), 297–312.

Carbone, L. (2021). "Estimating mouse and rat use in American laboratories by extrapolation from Animal Welfare Act-regulated species." *Scientific Reports, 11(1)*, 493. https://doi.org/10.1038/s41598-020-79961-0

Charlton, T., Panting, C., Davie, R., Coles, D., & Whitmarsh, L. (2000). "Children's playground behaviour across five years of broadcast television: A naturalistic study in a remote community." *Emotional & Behavioural Difficulties, 5(4)*, 4–12.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. London, UK. Routledge.

Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago, IL: Rand McNally.

Elo, E. & Kungäs, H. (2008). "The qualitative content analysis process." *Journal of Advanced Nursing, 62(1)*, 107–115.

Firestone, W.A. (1993). "Alternative arguments for generalizing from data as applied to qualitative research." *Educational Researcher, 22*, 16–23.

Goodwin, C.J. (2010). *Research in Psychology: Methods and Design* (Sixth edition). New York: USA. Wiley.

Gorman, K.B., Williams, T.D., Fraser, W.R. (2014). "Ecological sexual dimorphism and environmental variability within a community of Antarctic penguins (genus Pygoscelis)." *PLoS ONE, 9(3)*: e90081. https://doi.org/10.1371/journal.pone.0090081

Klorman, R. (1975). "The phono cartridge fidgetometer: Methodological considerations." *Behavior Research Methods & Instrumentation, 7(6)*, 577–579.

MacLean, P.D. (1990). *The Triune Brain in Evolution: Role in Paleocerebral Functions*. New York, USA. Plenum Press.

Maguire, E.A., Guardian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S.J., & Frith, C.D. (2000). "Navigation-related structural change in the hippocampi of taxi drivers." *Proceedings of the National Academy of Sciences, 97(8)*, 4398–4403.

Merzenich, M.M., Nelson, R.J., Stryker, M.P., Cynader, M.S., Schoppmann, A., & Zook, J.M. (1984). "Somatosensory cortical map changes following digit amputation in adult monkeys." *The Journal of Comparative Neurology, 224*, 591–605.

Piliavin, I.M., Rodin, J., & Piliavin, J.A. (1969). "Good Samaritanism: An underground phenomenon?" *Journal of Personality and Social Psychology, 13*, 289–99.

Polit, D.F., & Beck, C.T. (2010). "Generalization in quantitative and qualitative research: Myths and strategies." *International Journal of Nursing Studies, 47*, 1451–1458.

Popper, K.R. (1959). *The logic of scientific discovery*. Basic Books.

Rosenthal, R., & Fode, K. (1963) "The effects of experimenter bias on the performance of the albino rat." *Behavioral Science, 8*, 183–9.

Sharot, T., Martorella, E.A., Delgado, M.R., & Phelps, E.A. (2007). "How personal experience modulates the neural circuitry of memories of September 11." *Proceedings of the National Academy of Sciences, 104(4)*, 389–394.

Watson, J.B., & Rayner, R. (1920). "Conditioned emotional reaction." *Journal of Experimental Psychology, 2*, 1–14.

Chapter 2: Learning and cognition

Abbott, S.M., & Wong, G.K. (2008). "The conditioning and memory retention of planaria (Dugesia tigrina) for directional preferences." *BIOS, 79(4)*, 160–170. https://doi.org/10.1893/0005-3155-79.4.160

Ariely, D. (2009). *Predictably irrational*. HarperCollins.

Atkinson, R.C., & Shiffrin, R.M. (1968). "Human memory: A proposed system and its control processes." *Psychology of Learning and Motivation, 2*, 89–195.

Bach-y-Rita, P., Collins, C.C., Saunders, F.A., White, B., & Scadden, L. (1969). "Vision substitution by tactile image projection." *Nature, 221*, 963–964.

Baddeley, A. D., & Hitch, G. J. (1974). "Working Memory." In G. A. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A. D., & Hitch, G. J. (2000). "Development of working memory: Should the Pascual-Leone and the Baddeley and Hitch models be merged?" *Journal of Experimental Child Psychology, 77(2)*, 128–137.

Baddeley, A.D., Lewis, V., & Vallar, G. (1984). "Exploring the articulatory loop." *Quarterly Journal of Experimental Psychology, 36A(2)*, 233–252.

Baddeley, A.D., Thomson, N., & Buchanan, M. (1975). "Word length and the structure of short-term memory." *Journal of Verbal Learning and Verbal Behavior, 14(6)*, 575–589.

Bailey, J. (2008). "An assessment of the role of chimpanzees in AIDS vaccine research." *Alternatives to Laboratory Animals, 36(4)*, 381–428.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. Ross, D., & Ross, S.A. (1961). "Transmission of aggression through the imitation of aggressive models." *Journal of Abnormal and Social Psychology, 63*, 575–582

Barajas, G.R., Philipsen, N., & Brooks-Gunn, J. (2007). "Cognitive and emotional outcomes for children in poverty." In Crane, D.R. & Heaton, T. (Eds), *Handbook of Families and Poverty* (pp. 311–333). New York, USA. Sage.

Baron, J. (2008). *Thinking and deciding* (4th ed.). Cambridge University Press.

Berry, J.W. (1989). "Imposed etics–emics-derived etics: The operationalization of a compelling idea." *International Journal of Psychology*, *24*, 721–735.

Bower, G.H., Black, J.B., & Turner, T.J. (1979). "Scripts in memory for text." *Cognitive Psychology, 11*, 177–220.

Brewer, W.F. & Treyens, J.C. (1981). "Role of schemata in memory for places." *Cognitive Psychology, 13(2)*. 207–230

Cialdini, R.B. (2007). *Influence: The Psychology of Persuasion*. New York: HarperCollins.

Cole, M., & Scribner, S. (1974). *Culture & thought: A psychological introduction*. John Wiley & Sons.

Columb, C., & Plant, E.A. (2011). "Revisiting the Obama Effect: Exposure to Obama reduces implicit prejudice." *Journal of Experimental Social Psychology, 47(2)*, 499–501.

Conrad, R. & Hull, A.J. (1964). "Information, acoustic confusion and memory span." *British Journal of Psychology, 55(4)*, 429–432

Dickerson, A. & Popli, G.K. (2016). "Persistent poverty and children's cognitive development: Evidence from the UK Millennium Cohort Study." *Journal of the Royal Statistical Society, Series A, 179*, 535–558.

Englich, B., Mussweiler, T., & Strack, F. (2006). "Playing Dice with Criminal Sentences: The Influence of Irrelevant Anchors on Experts' Judicial Decision Making." *Personality and Social Psychology Bulletin, 32(2)*, 188–200. https://doi.org/10.1177/0146167205282152

Festinger, L., Riecken, H.W., & Schachter, S. (1956). *When Prophecy Fails: A Social and Psychological Study of a Modern Group that Predicted the Destruction of the World*. University of Minnesota Press.

Fox, M.W. (1974). *Concepts in Ethology: Animal and Human Behavior*. Minneapolis: University of Minnesota Press.

Galotti, K.M. (2008). *Cognitive Psychology: Perception, Attention and Memory*. New Delhi, India. Cengage Learning Ltd.

Giaimo, C. (2015). *Nervous system hookup leads to telepathic hand-holding*. Atlas Obscura, 10 June, 2015.

Glanzer, M., & Cunitz, A.R. (1966). "Two storage mechanisms in free recall." *Journal of Verbal Learning and Verbal Behavior, 5(4)*, 351–360.

Heider, F., & Simmel, M. (1944). "An experimental study of apparent behavior." *The American Journal of Psychology, 57(2)*, 243–259.

James, W. (1890). *The Principles of Psychology, in two volumes*. New York: Henry Holt and Company.

Joyce, N., & Faye, C. (2010). "Skinner Air Crib." *Association for Psychological Science*, 1 September 2010. https://www.psychologicalscience.org/observer/skinner-air-crib

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Lorenz, K., Tinbergen, N. (1970). "Taxis and instinct in egg-rolling response of the greylag goose." *Studies in Animal and Human Behavior, 1*, 328–342.

Ludolph, R., & Schulz, P.J. (2018). "Debiasing health-related judgments and decision making: A systematic review." *Medical Decision Making, 38(1)*, 3–13.

Martins, E., Mendes, F., Fernandes, R., & Sargento, J. (2020). "Cognitive impact of breakfast consumption on school age children and youth." In M. Jaworski (Ed.), *Health & Health Psychology*–icH&Hpsy 2020, vol 1. *European Proceedings of Health* (pp. 51–59). European Publisher. https://doi.org/10.15405/eph.20101.9

Mohebian, Z., Dehghan, S.F., & Dehghan, H. (2018) "Evaluation of the combined effects of heat and lighting on the level of attention and reaction time: climate chamber experiments in Iran." *The Scientific World Journal, 5171582*. https://doi.org/10.1155/2018/5171582

Piaget, J. & Inhelder, B. (1969). *The Psychology of the Child*. New York: Basic Books.

Pike, K.L. (1967). *Language in relation to a unified theory of the structure of human behaviour*. The Hague: Mouton

Scoville, W.B., & Milner, B. (1957). "Loss of recent memory after bilateral hippocampal lesions." *Journal of Neurology, Neurosurgery, and Psychiatry, 20(1)*, 11–21.

Sharot, T., Martorella, E.A., Delgado, M.R., & Phelps, E.A. (2007). "How personal experience modulates the neural circuitry of memories of September 11." *PNAS Proceedings of the National Academy of Sciences of the United States of America, 104(1)*, 389–394. https://doi.org/10.1073/pnas.0609230103

Skinner, B.F. (1938). *The behavior of organisms: an experimental analysis*. Appleton-Century.

Skinner, B.F. (1948a). "'Superstition' in the pigeon." *Journal of Experimental Psychology, 38(2)*, 168–172.

Skinner, B.F. (1948b). *Walden Two*. Macmillan.

Stanley, G.B., Li, F.F., & Dan, Y. (1999). "Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus." *The Journal of Neuroscience, 19(18)*, 8036–8042.

Strack, F., & Mussweiler, T. (1997). "Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility." *Journal of Personality and Social Psychology, 73(3)*, 437–446. https://doi.org/10.1037/0022-3514.73.3.437

Sweller, J. (1988), "Cognitive Load During Problem Solving: Effects on Learning." *Cognitive Science, 12*, 257–285. https://doi.org/10.1207/s15516709cog1202_4

Taylor, L., Watkins, S.L., Marshall, H., Dascombe, B.J., & Foster, J. (2016). "The impact of different environmental conditions on cognitive function: a focused review." *Frontiers of Physiology, 6: 372*. https://doi.org/10.3389/fphys.2015.00372

Tolman, E.C. (1948). "Cognitive maps in rats and men." *Psychological Review, 55(4)*, 189–208.

Turnbull, C.M. (1961). *The Forest People. A Study of the Pygmies of the Congo*. Simon & Schuster.

Tversky, A., & Kahneman, D. (1981). "The framing of decisions and the psychology of choice." *Science, 211(4481)*, 453–458. https://doi.org/10.1126/science.7455683

Twiley, N. (2017). "Seeing with your tongue. Sensory-substitution devices help blind and deaf people, but that's just the beginning." *The New Yorker*. 8 May 2017.

Velliste, M., Perel, S., Spalding, M.C., Whitford, A.S., & Schwartz, A.B. (2008). "Cortical control of a prosthetic arm for self-feeding." *Nature, 453(7198)*, 1098–1101.

Wang, C., Zhang, F., Wang, J., Doyle, J.K., Hancock, P.A., Mak, C.M., & Liu, S. (2021). "How indoor environmental quality affects occupants' cognitive functions: A systematic review." *Building and Environment, 193*, 107647. https://doi.org/10.1016/j.buildenv.2021.107647

Wason, P.C. (1968). "Reasoning about a rule." *Quarterly Journal of Experimental Psychology, 20(3)*, 273–281

Watson, J.B., & Rayner, R. (1920). "Conditioned emotional reactions." *Journal of experimental psychology, 3(1)*, 1.

Willett, F.R., Avansino, D.T., Hochberg, L.R., Henderson, J.M., & Shenoy, K.V. (2021). "High-performance brain-to-text communication via handwriting." *Nature, 593(7858)*, 249–254.

Chapter 3: Human development

Ainsworth, M., Blehar, M., Waters, E., & Wall, S. (1978). *Patterns of Attachment: A Psychological Study of the Strange Situation*. Hillsdale, New Jersey, USA. Lawrence Erlbaum.

Baron-Cohen, S., Leslie, A.M., & Frith, U. (1985). "Does the autistic child have a 'theory of mind'?" *Cognition, 21(1)*, 37–46.

Blakemore, C., & Cooper, G. (1970). "Development of the Brain depends on the Visual Environment." *Nature, 228*, 477–478. https://doi.org/10.1038/228477a0

Borghol, N., Suderman, M., McArdle, W., Racine, A., Hallett, M., Pembrey, M., Hertzman, C., Power, C., & Szyf, M. (2012). "Associations with early-life socio-economic position in adult DNA methylation." *International journal of epidemiology, 41(1)*, 62–74. https://doi.org/10.1093/ije/dyr147

Borke, H. (1975). "Piaget's mountains revisited: Changes in the egocentric landscape." *Developmental Psychology, 11(2)*, 240–243.

Bowlby, J. (1958). "The nature of the child's tie to his mother." *International Journal of Psycho-Analysis, Vol. XXXIX*, 1–23.

Bowlby, J., (1959). "Separation anxiety." *International Journal of Psycho-Analysis, Vol. XLI,* 1–25.

Brainerd, C.J. (1978). "The stage question in cognitive-developmental theory." *The Behavioral and Brain Sciences, 2*, 173–213.

Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). "Enculturated chimpanzees imitate rationality. *Developmental Sciences." Vol. 10*. F31–F38.

Byrne, D. (1961). "Interpersonal attraction and attitude similarity." *The Journal of Abnormal and Social Psychology, 62(3)*, 713–715. https://doi.org/10.1037/h0044721

Call, J., & Tomasello, M. (2008). "Does the chimpanzee have a theory of mind? 30 years later." *Trends in Cognitive Sciences, 12(5)*, 187–192.

Cohen, G.L., & Prinstein, M.J. (2006). "Peer contagion of aggression and health risk behavior among adolescent males: An experimental investigation of effects on public conduct and private attitudes." *Child Development, 77(4)*, 967–983. https://doi.org/10.1111/j.1467-8624.2006.00913.x

Curtiss, S., Fromkin, V.A., Krashen, S.D., Rigler, D., & Rigler, M. (1974). "The development of language in Genie: A case of language acquisition beyond the 'critical period'." *Brain and Language, 1(1)*, 81–107.

deVries, M.W. (1984). "Temperament and infant mortality among the Masai of East Africa." *American Journal of Psychiatry, 141*, 1189–1194.

Doidge, N. (2007). *The brain that changes itself: Stories of personal triumph from frontiers of brain science.* New York, USA: Penguin Books.

Fischer, K.W., & Silvern, L. (1985). "Stages and individual differences in cognitive development." *Annual Review of Psychology, 36*, 613–648. https://doi.org/10.1146/annurev.ps.36.020185.003145

Giletta, M., Choukas-Bradley, S., Maes, M., Linthicum, K. P., Card, N.A., & Prinstein, M.J. (2021). "A meta-analysis of longitudinal peer influence effects in childhood and adolescence." *Psychological Bulletin, 147(7)*, 719–747. https://doi.org/10.1037/bul0000329

Göckeritz, S., Schmidt, M.F.H., & Tomasello, M. (2014). "Young children's creation and transmission of social norms." *Cognitive Development, 30*, 81–95.

Greenough, W.T., Black, J.E., & Wallace, C.S. (1987). "Experience and brain development." *Child Development, 58(3)*, 539–559.

Harlow, H.F. (1958). "The nature of love." *American Psychologist, 13*, 673–685.

Haun, D.B., & Tomasello, M. (2011). "Conformity to peer pressure in preschool children." *Child Development, 82(6)*, 1759–1767. https://doi.org/10.1111/j.1467-8624.2011.01666.x

Hollos, M., & Cowan, P.A. (1973). "Social isolation and cognitive development: Logical operations and role-taking abilities in three Norwegian social settings." *Child Development, 44*, 630–641.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., & Rizzolatti, G. (2005). "Grasping the intentions of others with one's own mirror neuron system." *PLoS Biology, 3(3)*, e79.

Kolb, B., & Fantie, B. (1989). "Development of the child's brain and behavior." In C.R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of Clinical Child Neuropsychology* (pp. 17–39). New York: Plenum Press.

Kolb, B., & Fantie, B. (2009). "Development of the child's brain and behavior." In C.R. Reynolds, & E. Fletcher-Janzen (Eds.), *Handbook of Clinical Child Neuropsychology. Third Edition* (pp. 19–46). Springer Science+Business Media.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). "Great apes anticipate that other individuals will act according to false beliefs." *Science, 354*, 110–114.

Lecce, S., Bianco, F., & Ronchi, L. (2020). "Executive function in the school context: The role of peer relationships." *Infant and Child Development, 29(1)*, Article e2151. https://doi.org/10.1002/icd.2151

Leontyev, A.N. (1931). *Razvitie pamyati: Eksperimentalnoe issledovanie vysshih psihicheskikh funktsii*. ("Development of memory: Experimental investigation of higher-order cognitive functions"). Moscow. Uchpedgiz. (In Russian)

Luo, H., Hu, X., Liu, X., Ma, X., Guo, W., Qiu, C., & Li, T. (2012). "Hair cortisol level as a biomarker for altered hypothalamic–pituitary–adrenal activity in female adolescents with posttraumatic stress disorder after the 2008 Wenchuan earthquake." *Biological Psychiatry, 72*, 65–69.

McFarlane, A.C. (1987). "Posttraumatic phenomena in a longitudinal study of children following a natural disaster." *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 764–769.

McGowan, P.O., Sasaki, A., D'Alessio, A.C., Dymov, S., Labonte, B., Szyf, M., & Meaney, M.J. (2009). "Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse." *Nature Neuroscience, 12*, 342–348.

Meltzoff, A. (1995). "Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children." *Developmental Psychology, 31(5)*, 838–850.

Nedospasova, V.A. (1985). "The role of perspective in the development of logical thinking." In Poddyakov, N.N. and Govorkova, A.F. (Eds.), *Razvitiye Mishleniya i umstvennoe vospitaniye shkol'nikov (Development of thinking and intellectual upbringing of pre-school children).* Moscow. Pedagogika. (In Russian.)

Odden, H., & Rochat, P. (2004). "Observational learning and enculturation." *Educational and Child Psychology, 21(2)*, 39–50.

Pettit, G.S., & Harrist, A.W. (1993). "Children's aggressive and socially unskilled playground behavior with peers: Origins in early family relations." In C.H. Hart (Ed.), *Children on playgrounds: Research perspectives and applications* (pp. 240–270). State University of New York Press.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge & Kegan Paul.

Povinelli, D.J., & Eddy, T.J. (1996). "What young chimpanzees know about seeing." *Monograph of the Society for Research in Child Development. Vol 61*, pp. 1–152.

Premack, D., & Woodruff, G. (1978). "Does the chimpanzee have a theory of mind?" *Behavioral and Brain Sciences, 1(4)*, 515–526.

Rakoczy, H., & Schmidt, M.F.H. (2013). "The early ontogeny of social norms." *Child Development Perspectives, 7(1)*, 17–21.

Rieder, H., & Elbert, T. (2013). "Rwanda—lasting imprints of a genocide: trauma, mental health and psychosocial conditions in survivors, former prisoners and their children." *Conflict and Health, 7(1)*, 6. https://doi.org/10.1186/1752-1505-7-6

Roisman, G.I., & Fraley, R.C. (2013). "Developmental mechanisms underlying the legacy of childhood experiences." *Child Development Perspectives, 7(3)*, 149–154. https://doi.org/10.1111/cdep.12030

Ruffman, T., Garnham, W., & Rideout, P. (2001). "Social understanding in autism: Eye gaze as a measure of core insights." *Journal of Child Psychology and Psychiatry, 42(8)*, 1083–1094.

Saxe, R., & Kanwisher, N. (2003). "People thinking about thinking people: The role of the temporo-parietal junction in 'theory of mind'." *NeuroImage, 19(4)*, 1835–1842.

Shaver, P.R., & Hazan, C. (1988). "A biased overview of the study of love." *Journal of Social and Personality Relationships, 5*, 473–501.

Szyf, M., & Bick, J. (2013). "DNA methylation: a mechanism for embedding early life experiences in the genome." *Child development, 84(1)*, 49–57. https://doi.org/10.1111/j.1467-8624.2012.01793.x

Tronick, E.Z., Morelliz, G.A., & Winn, S. (1987). "Multiple caretaking of Efe (Pygmy) infants." *American Anthropologist, 89(1)*, 96–106.

Van Ijzendoorn, M.H., & Kroonenberg, P.M. (1988). "Cross-cultural patterns of attachment: A meta-analysis of the Strange Situation." *Child Development, 59*, 147–156.

Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Translated by Cole M., John-Steiner, V., Scribner, S. and Souberman, E. (Eds). Cambridge, Massachusetts, USA. Harvard University Press.

Weaver, I.C., Cervoni, N., Champagne, F.A., D'Alessio, A.C., Sharma, S., Seckl, J.R., & Meaney, M.J. (2004). "Epigenetic programming by maternal behavior." *Nature Neuroscience, 7*, 847–854.

Werker, J.F., Gilbert, J.H., Humphrey, K., & Tees, R.C. (1981). "Developmental aspects of cross-language speech perception." *Child Development, 52(1)*, 349–355. https://doi.org/10.2307/1129249

Chapter 4: Health and well-being

Alghowinem, S., Goecke, R., Epps, J., Wagner, M., & Cohn, J. (2016). "Cross-Cultural Depression Recognition from Vocal Biomarkers." In N. Morgan (Ed.), *Proceedings of Interspeech 2016* (Vol. 08-12-September–2016, pp. 1943–1947). (Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH). *International Speech Communication Association (ISCA)*. https://doi.org/10.21437/Interspeech.2016-1339

Alloy, L.B., Abramson, L.Y., & Francis, E.L. (1999). "Do Negative Cognitive Styles Confer Vulnerability to Depression?" *Current Directions in Psychological Science, 8(4)*, 128–132. http://www.jstor.org/stable/20182581

Andreassen, C.S., Pallesen, S., & Griffiths, M.D. (2017). "The relationship between addictive use of social media, narcissism, and self-esteem: Findings from a large national survey." *Addictive Behaviors, 64*, 287–293. 10.1016/j.addbeh.2016.03.006

Andreassen, C.S., Torsheim, T., Brunborg, G.S., & Pallesen, S. (2012). "Development of a Facebook addiction scale." *Psychological Reports, 110(2)*, 501–517. https://doi.org/10.2466/02.09.18.PR0.110.2.501–517

Andrews, G., Cuijpers, P., Craske, M.G., McEvoy, P., & Titov, N. (2010). "Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis." *PLoS One, 5(10)*: e13196. https://doi.org/10.1371/journal.pone.0013196

Bandura, A. (1989). "Human agency in social cognitive theory." *American Psychologist, 44,*1175–1184.

Baumeister, R.F. , Bratslavsky, E. , Muraven, M., & Tice, D.M. (1998). "Ego depletion: Is the active self a limited resource?" *Journal of Personality and Social Psychology, 74(5)*, 1252–1265. https://doi.org/10.1037/0022-3514.74.5.1252

Beck, A.T. (1967). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beck, A.T. (2008). "The evolution of the cognitive model of depression and its neurobiological correlates." *The American Journal of Psychiatry, 165(8)*, 969–977. https://doi.org/10.1176/appi.ajp.2008.08050721

Beck, A.T., & Bredemeier, K. (2016). "A unified model of depression: Integrating clinical, cognitive, biological, and evolutionary perspectives." *Clinical Psychological Science*, *4*(4), 596–619. https://doi.org/10.1177/2167702616628523

Beck, A.T., & Steer, R.A. (1984). "Internal consistencies of the original and revised Beck Depression Inventory." *Journal of Clinical Psychology*, *40(6)*, 1365–1367. https://doi.org/10.1016/j.jpsychires.2021.09.018

Black, M.M., Hager, E.R., Le K, A.J., Arteaga, S.S., Diclemente, C., Gittelsohn, J., Magder, L., Papas, M., Snitker, S., Treuth, M.S. & Wang, Y. (2010). "Challenge! Health promotion/obesity prevention mentorship model among urban, black adolescents." *Pediatrics, 126(2), 280–8*. https://doi.org/10.1542/peds.2009-1832. Epub 2010 Jul 26. PMID: 20660556; PMCID: PMC4124131.

Blair, J.E. (1993). "Social learning theory. Strategies for health promotion." *AAOHN Journal, 41(5)*, 245–249.

Brevers, D., Foucart, J., Turel, O., Bertrand, A., Alaerts, M., Verbanck, P., Kornreich, C., & Bechara, A. (2018). "The impact of self-control cues on subsequent monetary risk-taking." *Journal of Behavioral Addictions, 7(4),* 1044–1055. https://doi.org/10.1556/2006.7.2018.97

Brevers, D., & Turel, O. (2019). "Strategies for self-controlling social media use: Classification and role in preventing social media addiction symptoms." *Journal of Behavioral Addictions, 8(3)*, 554–563. https://doi.org/10.1556/2006.8.2019.49

Brooks-Gunn, L., & Duncan, G.J. (1997). "The effects of poverty on children." *The Future of Children: Children and Poverty, 17(2)*, 55–71.

Brown, G.W., & Harris, T.O. (1978). *Social Origins of Depression: A Study of Psychiatric Disorder in Women*. London, UK: Tavistock Publications.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A., & Poulton, R. (2003). "Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene." *Science, 301(5631)*, 386–389.

Chang, H., Meng, X., Li, Y., Liu, J., Yuan, W., Ni, J., & Li, C. (2023). "The effect of mindfulness on social media addiction among Chinese college students: A serial mediation model." *Frontiers of Psychiatry, 14*:1087909. https://doi.org/10.3389/fpsyt.2023.1087909

Chee, H.P., Hazizi, A.S., Barakatun Nisak, M.Y., & Mohd, Nasir, M.T. (2014). "A randomised controlled trial of a Facebook-based physical activity intervention for government employees with metabolic syndrome." *Malaysian Journal of Nutrition, 20(2)*, 165–181.

Cheng, C., Lau, Y.C., Chan, L., & Luk, J.W. (2021). "Prevalence of social media addiction across 32 nations: Meta-analysis with subgroup analysis of classification schemes and cultural values". *Addictive behaviors, 117*, 106845. https://doi.org/10.1016/j.addbeh.2021.106845

Chiao, J.Y., & Blizinsky, K.D. (2010). "Culture–gene coevolution of individualism–collectivism and the serotonin transporter gene. Proceedings." *Biological sciences*, *277(1681)*, 529–537. https://doi.org/10.1098/rspb.2009.1650

Fincher, C.L., Thornhill, R., Murray, D.R., Schaller, M. (2008). "Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism." *Proc. R. Soc. B 275*, 1279–1285.

Furnham, A. & Malik, R. (1994). "Cross-cultural beliefs about depression." *The International Journal of Social Psychiatry, 40(2)*, 106–123.

Hamilton, M. (1960). "A rating scale for depression." *Journal of Neurology, Neurosurgery and Psychiatry, 23*, 56–62.

Holland, J. (2013). "Fixing a broken drug development process". *Journal of Commercial Biotechnology, 19*. https://doi.org/10.5912/jcb588.

Inkster, B., Sarda, S., & Subramanian, V. (2018). "An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study." *JMIR mHealth and uHealth, 6(11)*, e12106. https://doi.org/10.2196/12106

Kendler, K.S., Gatz, M., Gardner, C.O., & Pedersen, N.L. (2006). "A Swedish national twin study of lifetime major depression." *American Journal of Psychiatry, 163(1)*, 109–114.

Kleinman, A. (1977). "Depression, somatisation and the new 'cross-cultural psychiatry'." *Social Science and Medicine, 11*, 3–10.

Lazarus, R.S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer publishing company.

Lewis, S.C., & Warlow, C.P. (2004). "How to spot bias and other potential problems in randomised controlled trials." *Journal of Neurology, Neurosurgery & Psychiatry, 75*, 181–187.

Liu, H., Peng, H., Song, X., Xu, C., & Zhang, M. (2022). "Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness." *Internet interventions, 27*, 100495. https://doi.org/10.1016/j.invent.2022.100495

Loh, Y.L., Yaw, Q.P., & Lau, Y. (2023). "Social media-based interventions for adults with obesity and overweight: a meta-analysis and meta-regression." *International Journal of Obesity, 47*, 606–621. https://doi.org/10.1038/s41366-023-01304-6

Lumsden, C.J., & Wilson, E.O. (1981). *Genes, mind and culture: the coevolutionary process* Cambridge, MA: Harvard University Press

Manwong, M., Lohsoonthorn, V., Booranasuksakul, T., & Chaikoolvatana, A. (2018). "Effects of a group

activity-based motivational enhancement therapy program on social media addictive behaviors among junior high school students in Thailand: a cluster randomized trial." *Psychology research and behavior management, 11*, 329–339. https://doi.org/10.2147/PRBM.S168869

Oei, T.P., & Kwon, S.M. (2007). "Evaluation of the integrated cognitive model of depression and its specificity in a migrant population." *Depression and anxiety, 24(2)*, 112–123. https://doi.org/10.1002/da.20225

Patten, S.B. (1991). "Are Brown and Harris 'vulnerability factors' risk factors for depression?" *Journal of Psychiatric Neuroscience, 16(5)*, 267–271.

Prochaska, J.O., & Norcross, J.C. (2018). *Systems of Psychotherapy: A Transtheoretical Analysis*. Oxford University Press.

Resenquist, J.N., Fowler, J.H., & Christakis, N.A. (2011). "Social network determinants of depression." *Molecular Psychiatry, 16*, 273–281.

Schramm, S., Sørensen, T.I.A., Davidsen, M., & Tolstrup, J.S. (2023). "Changes in adult obesity prevalence in Denmark, 1987–2021: age–period–cohort analysis of nationally representative data." *European Journal of Public Health, 33(3)*, 463–467. https://doi.org/10.1093/eurpub/ckad024

Schulz, K.F., Altman, D.G., Moher, D. et al. (2010). "CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials." *BMC Medicine, 8*, 18. https://doi.org/10.1186/1741-7015-8-18

Selye, H. (1976). "Stress without Distress." In G. Serban (Eds.), *Psychopathology of Human Adaptation*. Springer, Boston, MA. pp. 137–146. https://doi.org/10.1007/978-1-4684-2238-2_9

Stephens, J.D., Yager, A.M., & Allen, J. (2017). "Smartphone technology and text messaging for weight loss in young adults: a randomized controlled trial." *Journal of Cardiovascular Nursing, 32(1)*, 39–46.

Sullivan, P.F., Neale, M.C., & Kendler, K.S. (2000). "Genetic epidemiology of major depression: Review and meta-analysis." *American Journal of Psychiatry, 157(10)*, 1552–1562.

The Treatment for Adolescents with Depression Study (TADS): Long-term Effectiveness and Safety Outcomes. (2007). *Archives of General Psychiatry, 64(10)*, 1132–1143. https://doi.org/10.1001/archpsyc.64.10.1132

Turel, O., Brevers, D., & Bechara, A. (2018). "Time distortion when users at-risk for social media addiction engage in non-social media tasks." *Journal of Psychiatric Research, 97*, 84–88. https://doi.org/10.1016/j.jpsychires.2017.11.014. Epub 2017 Dec 2. PMID: 29220826.

Wankerl, M., Miller, R., Kirschbaum, C. et al. (2014) "Effects of genetic and early environmental risk factors for depression on serotonin transporter expression and methylation profiles." *Translational Psychiatry, 4, e402*. https://doi.org/10.1038/tp.2014.37

Witters, D. (2023). "U.S. depression rates reach new heights." *Gallup* (17 May 2023)

World Health Organization. (17 June 2022). "Mental Health" Retrieved from https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response

Zhao, N., & Zhou, G. (2021). "COVID-19 stress and addictive social media use (SMU): Mediating role of active use and social media flow." *Frontiers in Psychiatry, 12*, 85. https://doi.org/10.3389/fpsyt.2021.635546

Chapter 5: Human relationships

Alegría, M., Canino, G., Shrout, P.E., Woo, M., Duan, N., Vila, D., & Meng, X-L. (2008). "Prevalence of mental illness in immigrant and non-immigrant U.S. Latino groups." *American Journal of Psychiatry, 165*, 359–369.

Allen, M.L., Elliott, M.N., Fugligni, A.J., Morales, L.S., Hambarsoomian, K., & Schuster, M.A. (2008). "The relationship between Spanish language use and substance use behaviors among Latino youth: A social network approach." *Journal of Adolescent Health, 43*, 372–379.

Arnow, B.A., Desmond, J.E., Banner, L.L., Glover, G.H., Solomon, A., Polan, M.L., Lue, T.F., & Atlas, S.W. (2002). "Brain activation and sexual arousal in healthy, heterosexual males." *Brain, 125*, 1014–1023.

Aron, A., Fisher, H.E., Mashek, D.J., Strong, G., Li, H.F., & Brown, L.L. (2005). "Reward, motivation and emotion systems associated with early-stage intense romantic love: an fMRI study." *Journal of Neurophysiology, 94*, 327–337.

Asch, S.E. (1951). "Effects of group pressure upon the modification and distortion of judgments." In H. Guetzkow, (Ed.), *Groups, Leadership and Men*, Carnegie Press, Pittsburgh, pp. 177–190.

Asch, S.E. (1955). "Opinions and social pressure." *Scientific American, 193(5)*, 31–35.

Berry, J.W. (1992). Acculturation and adaptation in a new society. *International Migration, 30(1)*, 69–85.

Brady, W.J., McLoughlin, K., Doan, T.N., & Crockett, M.J. (2021). "How social learning amplifies moral outrage expression in online social networks." *Science advances, 7(33)*, eabe5641. https://doi.org/10.1126/sciadv.abe5641

Brewer, M.B., & Yuki, M. (2007). "Culture and social identity." In S. Kitayama & D. Cohen (Eds.), *Handbook of Cultural Psychology* (pp. 307–322). The Guilford Press.

Brown, R. (2000). "Social identity theory: Past achievements, current problems and future challenges." *Europe's Journal of Social Psychology, 30(6)*, 745–778. https://doi.org/10.1002/1099-0992(200011/12)30:6<745::AID-EJSP24>3.0.CO;2-O

Bryan, J.H. (1975). "Children's cooperation and helping behavior." In E.M. Hetherington (Ed.), *Review of child development research (Vol. 5)*. Chicago: University of Chicago Press.

Buchan, N.R., Croson, R., & Dawes, R.M. (2002). "Swift neighbors and persistent strangers: A crosscultural investigation of trust and reciprocity in social exchange." *American Journal of Sociology, 108*, 168–206.

Cabassa, L.J. (2003). "Measuring acculturation: Where we are and where we need to go." *Hispanic Journal of Behavioral Sciences*, *25*(2), 127–146. https://doi.org/10.1177/0739986303253626

Cialdini, R. (1984). *Influence: The Psychology of Persuasion*. HarperCollins, New York.

Cialdini, R.B., Cacioppo, J.T., Bassett, R., & Miller, J.A. (1978). "Low-ball procedure for producing compliance: Commitment then cost." *Journal of Personality and Social Psychology*, *36*(5), 463–476.

Ferguson, Y.L., Ferguson, K.T., & Ferguson, G.M. (2015). "I am AmeriBritSouthAfrican–Zambian: Multidimensional remote acculturation and well-being among urban Zambian adolescents." *International Journal of Psychology: Journal international de Psychologie, 52(1)*, 67–76. https://doi.org/10.1002/ijop.12191

Fisher, H., Aron, A., & Brown, L.L. (2005). "Romantic love: an fMRI study of a neural mechanism for mate choice." *Journal of Comparative Neurology, 493*, 58–62.

Fisher, H.E., Aron, A., & Brown, L.L. (2006). "Romantic love: a mammalian brain system for mate choice." *Philosophical Transactions of the Royal Society B*, 3612173–2186. https://doi.org/10.1098/rstb.2006.1938

Freedman, J.L., & Fraser, S.C. (1966). "Compliance without pressure: The foot-in-the-door technique." *Journal of Personality and Social Psychology, 4(2)*, 195–202.

Friedrich-Cofer, L.K., Huston-Stein, A., Kipnis, D.M., Susman, E.J., & Clewett, A.S. (1979). "Environmental enhancement of prosocial television content: Effects on interpersonal behavior, imaginative play, and self-regulation in a natural setting." *Developmental Psychology, 15*, 637–646.

Friend, R., Rafferty, Y., & Bramel, D. (1990). "A puzzling misinterpretation of the Asch 'conformity' study." *European Journal of Social Psychology, 20(1)*, 29–44. https://doi.org/10.1002/ejsp.2420200104

Gerlach, P., & Eriksson, K. (2021). "Measuring Cultural Dimensions: External Validity and Internal Consistency of Hofstede's VSM 2013 Scales." *Frontiers of Psychology*, *12*:662604. https://doi.org/10.3389/fpsyg.2021.662604

Gottman, J.M. (1994). *What predicts divorce? The relationship between marital processes and marital outcomes*. Lawrence Erlbaum Associates, Inc.

Gottman, J.M., & Krokoff, L.J. (1989). "Marital interaction and satisfaction: A longitudinal view." *Journal of Consulting and Clinical Psychology, 57(1)*, 47–52.

Griggs, R.A. (2015). "The disappearance of independence in textbook coverage of Asch's social pressure experiments." *Teaching of Psychology, 42(2)*, 137–142.

Guerrero, L.K., Anderson, P.A., & Afifi, W.A. (2021). *Close encounters : communication in relationships*. Thousand Oaks, Sage.

Hatfield, E. & Traupmann, J. (1980). "Intimate relationships: A perspective from equity theory." In S. Duck, R. Gilmour (Eds.), *Personal relationships I: Studying personal relationships.* Academic Press. pp. 165–178.

Hofstede, G. (1980). *Culture's Consequences: International Differences in Work Related Values*. Beverly Hills, CA: Sage.

Hofstede, G., Hofstede, G.J., & Minkov, M. (2010). *Culture and organizations: Software of the mind*. New York: McGraw-Hill

Kühn, S., Kugler, D., Schmalen, K., Weichenberger, M., Witt, C., & Gallinat, J. (2019). "Does playing violent video games cause aggression? A longitudinal intervention study." *Molecular Psychiatry, 24*, 1220–1234. https://doi.org/10.1038/s41380-018-0031-7

Morgan, T.J.H., & Laland, K.N. (2012). "The biological bases of conformity." *Frontiers in Neuroscience, 6(87)*, 1–7. https://doi.org/10.3389/fnins.2012.00087

Pike, T.W., and Laland, K.N. (2010). "Conformist learning in nine-spined sticklebacks' foraging decisions." *Biology Letters, 6*, 466–468.

Rajaei, A., Daneshpour, M., & Robertson, J. (2019). "The effectiveness of couples therapy based on the Gottman method among Iranian couples with conflicts: A quasi-experimental study." *Journal of Couple & Relationship Therapy, 18(3)*, 223–240. https://doi.org/10.1080/15332691.2019.1567174

Rusbult, C. (1980). "Commitment and satisfaction in romantic associations: A test of the investment model." *Journal of Experimental Social Psychology, 16*, 172–186. https://doi.org/10.1016/0022-1031(80)90007-4

Rusbult, C. (1983). "A longitudinal test of the investment model: The development (and deterioration) of satisfaction and commitment in heterosexual involvements." *Journal of Personality and Social Psychology, 45*, 172–186. https://doi.org/10.1037//0022-3514.45.1.101

Salomons, N., Sebo, S.S., Qin, M., & Scassellati, B. (2021). "A minority of one against the majority of robots: Robots cause normative and informational conformity." *ACM Transactions on Human–Robot Interaction (THRI), 10(2)*, 1–22.

Schwartz, S.J., Unger, J.B., Zamboanga, B.L., & Szapocznik, J. (2010). "Rethinking the concept of acculturation: implications for theory and research." *The American Psychologist*, *65*(*4*), 237–251. https://doi.org/10.1037/a0019330

Schwartz, S.J. & Zamboanga, B.L. (2008). "Testing Berry's model of acculturation: A confirmatory latent class approach." *Cultural Diversity and Ethnic Minority Psychology, 14*, 275–285

Shah, S.M., Loney, T., Dhaheri, S.A. et al. "Association between acculturation, obesity and cardiovascular risk factors among male South Asian migrants in the United Arab Emirates—a cross-sectional study." *BMC Public Health, 15*, 204 (2015). https://doi.org/10.1186/s12889-015-1568-x

Sherif, M. (1956). "Experiments in group conflict." *Scientific American, 195(5)*, 54–59

Stein, A.H., & Friedrich, L.K. (1972). "Television content and youth children's behavior." In J.P. Murray, E.A. Rubinstein, & G.A. Comstock (Eds.), *Television and social behaviour (Vol. 2). Television and social learning*. Washington: U.S. Government Printing Office.

Tajfel, H., & Turner, J.C. (1979). "An integrative theory of intergroup conflict." In W.G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations*. Monterey, California, USA. Books/Cole.

Tajfel, H., Flament, C., Billig, M.G., & Bundy, R.F. (1971). "Social categorization and intergroup behaviour." *European Journal of Social Psychology, 1(2)*, 149–177.

Thibaut, J.W., & Kelley, H.H. (1959). *The social psychology of groups*. Routledge.

Yamagishi, T., Foddy, M., Makimura, Y., Matsuda, M., & Platow, M. (2003). *Contextualized and decontextualized use of social categories: Comparisons of Australians and Japanese on group-based trust and cooperation*. Sapporo, Japan: Hokkaido University, Department of Behavioral Science.

Chapter 6: HL extensions

Arabiat, D., Al Jabery, M., Robinson, S., Whitehead, L., & Mörelius, E. (2022). "Interactive technology use and child development: A systematic review." *Child: Care, Health and Development*, *49(4)*, 679–715.

Arndt, J., Greenberg, J., Schimel, J., Pyszczynski, T., & Solomon, S. (2002). "To belong or not to belong, that is the question: Terror management and identification with gender and ethnicity." *Journal of Personality and Social Psychology, 83(1)*, 26–43.

Axford, C., Joosten, A.V., & Harris, C. (2018). "iPad applications that required a range of motor skills promoted motor coordination in children commencing primary school." *Australian Occupational Therapy Journal, 65(2)*, 146–155.

Bello, R.S., Brandau-Brown, F.E., Zhang, S., & Ragsdale, D. (2010). "Verbal and nonverbal methods for expressing appreciation in friendships and romantic relationships: A cross-cultural comparison." *International Journal of Intercultural Relations, 34*, 294–302.

Braghieri, L., Levy, R., & Makarin, A. (2022). "Social media and mental health." *American Economic Review, 112(11)*, 3660–93.

Briley, D.A., Morris, M.W., & Simonson, I. (2005). "Cultural Chameleons: Biculturals, Conformity Motives, and Decision Making." *Journal of Consumer Psychology, 15(4)*, 351–362.

Carrier, L.M., Spradlin, A., Bunce, J.P., & Rosen, L.R. (2015). "Virtual empathy: Positive and negative impacts of going online upon empathy in young adults." *Computers in Human Behavior, 52*, 39–48.

Chandler, C.L., & Connell, J.P. (1987). "Children's intrinsic, extrinsic and internalized motivation: A developmental study of children's reasons for liked and disliked behaviours." *British Journal of Developmental Psychology, 5*, 357–365.

Cihak, D.F., Moore, E.J., Wright, R.E., McMahon, D.D., Gibbons, M.M., & Smith, C. (2016). "Evaluating augmented reality to complete a chain task for elementary students with autism." *Journal of Special Education Technology, 31*, 99–108.

Das, E., Bushman, B.J., Bezemer, M.D., Kerkhof, P., & Vermeulen, I.E. (2009). "How terrorism news reports increase prejudice against outgroups: a terror management account." *Journal of Experimental Social Psychology, 45(3)*, 453–459. https://doi.org/10.1016/j.jesp.2008.12.001

Dasen, P.R. (1984). "The cross-cultural study of intelligence: Piaget and the Baoulé." *International Journal of Psychology, 19(1–4)*, 407–434.

Deci, E.L., & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Springer Science+Business Media.

Duan, H., Fernández, G., van Dongen, E., Kohn, S. (2020). "The effect of intrinsic and extrinsic motivation on memory formation: insight from behavioral and imaging study." *Brain Structure and Function, 225*, 1561–1574. https://doi.org/10.1007/s00429-020-02074-x

Freud, S. (1916). "Introductory Lectures on Psycho-analysis (part III)." *The Standard Edition of the Complete Psychological Works of Sigmund Freud. Volume 16 (1963)*. Hogarth Press: The Institute of Psycho-Analysis.

Greenberg, J., Pyszczynski, T., & Solomon, S. (1986). "The causes and consequences of a need for self-esteem: A terror management theory." In R.F. Baumeister (Ed.), *Public and Private Self* (pp. 189–212). New York, NY: Springer-Verlag.

Haidt, J., & Allen, N. (2020). "Scrutinizing the effects of digital technology on mental health." *Nature, 578(7794)*, 226–227. https://doi.org/10.1038/d41586-020-00296-x

Hazan, C., & Shaver, P. (1987). "Romantic love conceptualized as an attachment process." *Journal of Personality and Social Psychology, 52(3)*, 511–524. https://doi.org/10.1037/0022-3514.52.3.511

Hogg, M.A., Hohman, Z.P., & Rivera, J.E. (2008). "Why do people join groups? Three motivational accounts from social psychology." *Social and Personality Psychology Compass, 2(3)*, 1269–1280.

Howard-Jones, P. (2011). "The impact of digital technologies on human wellbeing. Evidence from the sciences of mind and brain." www.nominettrust.org.uk

Hull, C.L. (1943). *Principles of behavior: An introduction of behavior theory*. New York: Appleton-Century-Crofts.

Hunt, M.G., Marx, R., Lipson, C., & Young, J. (2018). "No more FOMO: Limiting social media decreases loneliness and depression." *Journal of Social and Clinical Psychology, 37(10)*, 751–768. https://doi.org/10.1521/jscp.2018.37.10.751

Hutchins, E. (1995). *Cognition in the wild*. The MIT Press.

Ji, L.J, & Yap. S. (2016). "Culture and cognition." *Current Opinion in Psychology, 8*, 105–111.

Kanfer. R., & Ackerman, P.L. (1989). "Motivation and cognitive abilities: An integrative/aptitude–treatment interaction approach to skill acquisition." *Journal of Applied Psychology, 74(4)*, 657–690

Karandashev, V., Zarubko, E., Artemeva, V., Evans, M., Morgan, K., Neto, F., Feybesse, C., Surmanidze, L., & Purvis, J. (2020). "Cross-Cultural Comparison of Sensory Preferences in Romantic Attraction." *Sexuality & Culture, 24,* 23–53. https://doi.org/10.1007/s12119-019-09628-0

Keeler, L.A., Skidmore, B., Leenstra, T., MacDonald, J.R., & Stewart, D. (2021). "Treating University Students' Depression using Physical Activity with Peers: Two Field-Based Quasi-Experiments Grounded in the Self-Determination Theory." *Journal of College Student Psychotherapy, 35(3)*, 205–223. https://doi.org/10.1080/87568225.2019.1660293

Kinzie, J.D., Leung, P., Boehnlein, J.K., & Fleck, J. (1987). "Antidepressant blood levels in Southeast Asians: clinical and cultural implications." *Journal of Nervous and Mental Disorders, 175*, 480–485.

Kleemans, M., Daalmans, S., Carbaat, I., & Anschütz, D. (2018). "Picture perfect: The direct effect of manipulated Instagram photos on body image in adolescent girls." *Media Psychology, 21(1)*, 93–110. https://doi.org/10.1080/15213269.2016.1257392

Lin, E.H.B., Carter, W.B. & Kleinman, A.M. (1985). "An exploration of somatization among Asian refugees and immigrants in primary care." *American Journal of Public Health, 75(9)*, 1080–1084.

Masuda, T., & Nisbett, R.E. (2001). "Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans." *Journal of Personality and Social Psychology, 81(5)*, 922–934.

Mathur, P. (2018). "Meet the real-life Mowgli-Dina Sanichar who was raised by a pack of wolves." *India Times*, 10 December 2018.

Mayer, C., Wallner, S., Budde-Spengler, N., Braunert, S., Arndt, P. A., & Kiefer, M. (2020). "Literacy training of kindergarten children with pencil, keyboard or tablet stylus: The influence of the writing tool on reading and writing performance at the letter and word level." *Frontiers in Psychology, 10*, 3054. https://doi.org/10.3389/fpsyg.2019.03054

Michaelian, K., & Sutton, J. (2013). "Distributed Cognition and Memory Research: History and Current Directions." *Review of Philosophy and Psychology, 4*, 1–24. https://doi.org/10.1007/s13164-013-0131-x

Morey, J.N., Gentzler, A.L., Creasy, B., Oberhauser, A.M., & Westerman, D. (2013). "Young adults' use of communication technology within their romantic relationships and associations with attachment style." *Computers in Human Behavior, 29(4)*, 1771–1778. https://doi.org/10.1016/j.chb.2013.02.019.

Nelson, E.C., Sools, A.M., Vollenbroek-Hutten, M.M.R., Verhagen, T., & Noordzij, M.L. (2020). "Embodiment of Wearable Technology: Qualitative Longitudinal Study." *JMIR mHealth and uHealth, 8(11)*, e16973. https://doi.org/10.2196/16973

Nisbett, R.E. (2003). *The Geography of Thought: How Asians and Westerners Think Differently ... and Why*. New York, NY: Free Press.

Okdie, B.M. & Ewoldsen, D.R. (2018). "To boldly go where no relationship has gone before: Commentary on interpersonal relationships in the digital age." *The Journal of Social Psychology, 158(4)*, 508–513. https://doi.org/10.1080/00224545.2018.1473835

Payne, J.S. (2012). "Influence of race and symptom expression on clinicians' depressive disorder identification in African American men." *Journal of the Society for Social Work and Research, 3(3)*, 162–177.

Pew Research Center. (2015, August). "Teens, technology, and friendships." Retrieved from http://www.pewinternet.org/2015/08/06/teens-technology-and-friendships/

Quiroz, S.I., & Mickelson, K.D. (2021). "Are online behaviors damaging our in-person connections? Passive versus active social media use on romantic relationships." *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 15(1)*. https://doi.org/10.5817/CP2021-1-1

Robinson, L.J., Stevens, L.H., Threapleton, C.J.D., Vainiute, J., McAllister-Williams, R.H., & Gallagher, P. (2012). "Effects of intrinsic and extrinsic motivation on attention and memory." *Acta Psychologica, 141*, 243–249.

Sanchez, C.A. (2012). "Enhancing visuospatial performance through video game training to increase learning in visuospatial science domains." *Psychonomic Bulletin Review, 19*, 58–65.

Scelza, B.A., Prall, S.P., Blumenfield, T., Crittenden, A. N., Gurven, M., Kline, M., Koster, J., Kushnick, G., Mattison, S.M., Pillsworth, E., Shenk, M.K., Starkweather, K., Stieglitz, J., Sum, C.Y., Yamaguchi, K., & McElreath, R. (2020). "Patterns of paternal investment predict cross-cultural variation in jealous response." *Nature human behaviour, 4(1)*, 20–26. https://doi.org/10.1038/s41562-019-0654-y

Senzaki, S., Masuda, T., & Nand, K. (2014). "Holistic versus analytic expressions in artworks: cross-cultural differences in and similarities in drawings and collages by Canadian and Japanese school-age children." *Journal of Cross-Cultural Psychology, 45*, 1297–1316.

Sparrow, B., Liu, J., & Wegner, D. M. (2011). "Google effects on memory: Cognitive consequences of having information at our fingertips." *Science, 333(6043)*, 476–478. https://doi.org/10.1126/science.1207745

Uskul, A.K., & Kitayama, S., & Nisbett, R.E.: (2008). "Ecocultural basis of cognition: farmers and fishermen are more holistic than herders." *Proceedings of the National Academy of Sciences, 105*, 8552–8556.

Vaterlaus, J.M., Tulane, S., Porter, B.D., & Beckert, T.E. (2018). "The Perceived Influence of Media and Technology on Adolescent Romantic Relationships." *Journal of Adolescent Research, 33(6)*, 651–671. https://doi.org/10.1177/0743558417712611

Vinden, P.G. (1996). "Junín Quechua children's understanding of mind." *Child Development, 67*, 1707–1716.

Wampold, B.E. (2001). *The great psychotherapy debate: Models, methods, findings.* Hillsdale, NJ: Erlbaum.

Wilhelmsen, M., Lillevoll, K., Risør, M.B., Høifødt, R., Johansen, M-L., Waterloo, K., Eisemann, M., & Kolstrup, N. (2013). "Motivation to persist with internet-based cognitive behavioural treatment using blended care: a qualitative study." *BMC Psychiatry, 13*, 296. https://doi.org/10.1186/1471-244X-13-296

Witkin, H.A., & Berry, J.W. (1975). "Psychological differentiation in cross-cultural perspective." *Journal of Cross-Cultural Psychology, 6*, 4–87.

Zuckerman, M., Porac, J., Lathin, D., & Deci, E.L. (1978). "On the Importance of Self-Determination for Intrinsically-Motivated Behavior." *Personality and Social Psychology Bulletin, 4(3)*, 443–446. https://doi.org/10.1177/014616727800400317

Zuroff, D.C., Koestner, R., Moskowitz, D.S., McBride, C., Marshall, M., & Bagby, M.R. (2007). "Autonomous motivation for therapy: A new common factor in brief treatments for depression." *Psychotherapy Research, 17(2)*, 137–147. https://doi.org/10.1080/10503300600919380

Chapter 8: External assessment

Zak, P.J., Stanton, A.A., & Ahmadi, S. (2007). "Oxytocin Increases Generosity in Humans." *PLoS ONE 2(11)*: e1128. https://doi.org/10.1371/journal.pone.0001128

Index

Page numbers in *italic* refer to figures; page numbers in **bold** refer to tables.

5-HTT gene (serotonin transporter) 255, 256, 294
5-HTTLPR genotype 285, *286*
9/11 attacks (2001), study 27, 122

accommodation 112, 113
acculturation 207, 352–61
  acculturation contexts 357
  acculturation strategies 353, 356
  conceptual analysis 360–1
  health outcomes 357, 360, 361
  immigrant paradox 357–8
  limitations of Berry's model 355
  models of 353–4
  proxy measures 354–5, 358, 361
  remote acculturation 358–9
  Schwartz's extension of Berry's model 356–7
  two-dimensional model 354, 355, 356–7
  unidimensional model 354, 360
acquiescence bias 39, 52
ACT *see* Air Traffic Controller simulation
action, distortion of 365
active use of social media 311
acute stress 310
addictive behaviour
  and cultural practices 357
  digital technologies 470
  social media addiction 248, 300–3, 305, 311–14, 316, 324–30
additive influence, and interaction of factors 294
adherence to therapy 277, 282
adolescents
  body image of girls 478–9
  depression 260–2
  peer influence 215–16
  remote acculturation 358–9
  romantic relationships and digital technology 472–3
adverse childhood experiences 222–8, 297
  conceptual analysis 227–8
  epigenetic changes 226–7
  individual differences 227
  Maasai people 223–4
  persistent effect of *225*
  and resilience 223–4, 227, 228
  reversibility of effects of 224–5
aetiology 258, 272, 284–8, 289
age effects 303
agency, and cognitive behavioural therapy 276–7, 282
aggressive behaviour, observational learning 146–7, 151, 379–82, 384
AI (artificial intelligence) 8, 279, 280, 281, 406, 469
Ainsworth, Mary 240, 241, 243
air crib, Skinner's 108, 109
Air Traffic Controller (ATC) simulation 438–9
alleles 250, *251*
allocation concealment 263
alternative explanations 114, 115
alternative hypothesis (H_1) 81, 82
American Psychological Association (APA)
  animal research guidelines 62
  code of ethics 279
amnesia 120
amygdala, selective activation of 122
analytic cognitive style 463–4, 466
anchoring bias 139
animal intelligence 5–6
animal models/research 58–63, 189
  advantages/disadvantages of 61–2
  attachment studies 238–9
  brain–computer interfaces 172
  ethical issues 62–3, 236, 239, 244
  regulation of 62–3
  theory of mind studies 230–1
Animal Welfare Act (AWA) 58
anonymity 55, 56
anterograde amnesia 120
antidepressant drugs 254, 258–62, 265, 272, 455
APA *see* American Psychological Association
apps, CBT 279–81
Ariely, Dan 141
articulatory suppression technique 130–1
artificial intelligence (AI) 8, 279, 280, 281, 406, 469
Asch paradigm 364, 365, 366, 368
Asch, S.E. 363–6
ASD *see* autistic spectrum disorder
aspirin 62
assessment
  exam-style practice questions 177, 245, 331, 413
  external assessment 505–22
  HL extensions 415, 509, 514–15
  internal assessment 493–503
assimilation 112, 114, 353
asymmetrical communication 405
attachment 237–44
  in adults 242
  attachment behavioural system 239–40
  attachment styles 240–1, 243
  as casual factor 243
  conceptual analysis 243–4
  contact comfort hypothesis 238
  cultural variations in 241–2, 243
  internal working model 240
  observational studies 244
  secure base hypothesis 239
  theories of 239–40
attention
  and motivation 439–40
  social learning theory 145
  and working conditions 162
attention control 328, *329*
attraction/romantic love
  chemical messengers, role of 388–90
  cultural differences 447–8, 451–2
augmented reality, in schools 483–4
authority, compliance technique 372
autistic spectrum disorder (ASD)
  eye gaze measurement 234
  primary school students 483–4, 486
  theory of mind 233
automatic processing 135
automatic thoughts 267, 276
autonomous motivation, and therapy outcomes 426–31
autonomy 420, 436
AWA *see* Animal Welfare Act
axons 253

Bach-y-Rita, Paul 169–70
Bandura, Albert 146–7, 151, 317, 379–82
Baoulé people 459–60
bar graphs 74–5
bases, DNA 250
BBB *see* blood-brain barrier
BCI *see* brain–computer interfaces
Beck Depression Inventory (BDI) 270
Beck's cognitive theory of depression 267–73
  evaluation of 269–71
  limitations of 272
behaviour 7
  addictive behaviour 300–3, 305, 316, 357, 470
  aggressive behaviour 146–7, 151, 379–82, 384
  biological perspective 117–18
  and causality 108
  charitable behaviour 375–6
  cognitive perspective 117–18
  control-determined behaviours 419
  courtship behaviour 388–9
  and environment 149–50
  externalizing behaviours 216
  health behaviours 316–17
  innate behaviours 99, 100
  internalizing behaviours 216
  law of effect 105
  learned behaviours 99, 100, 101–9
  microanalytic observation of behaviour 407–8
  personal factors 148
  prosocial behaviours 382
  reward-seeking behaviours 390
  and schemas 115
  self-determined behaviours 419
  sociocultural perspective 117–18
  types of *102*
behavioural compliance 374, 375
behavioural exercises 276
behavioural perspective 215
behaviourism 418, 419
  and bias 109
  history of 102–3
  mental representations 111–12
  operant variability 108
  radical behaviourism 107
  teleological behaviourism 112
Bergen Facebook Addiction Scale (BFAS) 301, 302, 313
Bergen Social Media Addiction Scale (BSMAS) 301, 302
Berry, John 155, 159, 353–5
Berscheid, Ellen 404
BFAS (Bergen Facebook Addiction Scale) 301, 302, 313
bias ix, 14–15, 38–40, 235, 257
  acculturation 360
  acquiescence bias 39, 52
  anchoring bias 139
  and behaviourism 109
  biological factors in cognition 123
  clinical trials 264
  cognitive biases 136–42, 268, 273
  cognitive development 197, 205
  cognitive models 133
  cognitive schemas 116
  confirmation bias 39, 137–8, *140*
  confounding variables 17
  in correlational studies 35
  cultural dimensions 351
  cultural studies 159
  debiasing 140, 170–1
  demand characteristics 25
  dominant respondent bias 39
  enculturation of social norms 212
  environment and cognition 166
  experimental mortality 25
  experimenter bias 25–6
  extreme responding bias 52
  framing effect 136–7
  leading questions bias 39
  maturation bias 23
  participant bias **15**, 39
  prevalence rates 306
  publication bias 263–4, 321
  in qualitative research 38–40
  question order bias 39–40
  questionnaires 51–2
  randomized controlled trials 263–4
  reporting bias 40, 287–8, 289
  sampling bias 40, 289, 321
  selection bias 23, 321
  sensitivity bias 39
  social desirability bias 39, 51
  social learning theory 151
  surveys and questionnaires 51–2
  testing effect 23–4
  theory of mind 462
biased reporting 40, 287–8, 289
bibliotherapy 280
biculturalism 353
bidirectional ambiguity 380, 384, 402
bidirectional interactions 147, 148–50
binaural hearing 153
biological basis
  conformity 366–7
  theory of mind 234, *235*
biological maturation 183, 189
biological perspective 117–18
biological reductionism 118–19, 257
black box metaphor 103, 111
blood-brain barrier (BBB) 392
BMI *see* body mass index
Bobo doll experiment 146–7, 379, 380
body image, in adolescent girls 478–9, 480
body mass index (BMI) 303, 304
boundary cases 45
Bowlby, John 239–40
box-and-whisker plots 75–7
BPS *see* British Psychological Society
brain
  amygdala *122*
  blood-brain barrier 392
  cortical remapping of digits, owl monkeys 59–60
  injury to 120–1
  triune brain theory 61
brain–computer interfaces (BCI) 172–4, 175
brain development 179, 181–9
  biological process 181–3
  and cognitive development 188
  conceptual analysis 188–9
  developmental neuroscience 187–8
  experience-dependent neuroplasticity 185–6
  experience-expectant neuroplasticity 184–5
  prenatal brain development *181*
  structure–function relationships 186–7
brain imaging techniques 121–2, 123, 187–8, *234*, 390
breakfast consumption 163–4
British Asians, perception of depression 287–8
British Psychological Society (BPS), animal research guidelines 63
BSMAS (Bergen Social Media Addiction Scale) 301, 302
Burt, Cyril 57

caretaking practices 241–2
case studies 45
case-to-case generalization 42
categorized charts *79*
causality ix–x
  cognitive processes 150, 159, 175
  cognitive schemas 115
  domino causality 225
  perceived locus of causality 419–20
cause–effect hypotheses 60, 217
cause–effect inferences 16–17, 141
CBT *see* cognitive behavioural therapy
CCBT *see* computerized CBT
central tendency 67–8, 446
chair design, sensory substitution 169
change x
charitable behaviours 375–6
chatbots, AI 8, 279, 280, 281, 406, 469

cheating, romantic relationships 397
chemical imbalance theory 262
chemical messengers 386–94
  and attraction/romantic love 388–90
  causal role of 394
  conceptual analysis 393–4
  and partner attachment 391–2
  research limitations 392
  in romantic relationships 388
  and sex drive 391
  types of 387–8
Chicago Word Fluency Test 186–7
childhood, concept of 180
children
  adverse childhood experiences 222–8, 297
  aggressive behaviour 146–7, 151, 379–82, 384
  attachment 240–2
  autistic spectrum disorder 233, 234, 483–4, 486
  breakfast consumption 163–4
  child-rearing practices 180
  cognitive development 164–5, 192–7, 204, 481–6
  cognitive egocentrism 192, 193
  cognitive schemas 113
  concrete operational stage (age 7 to 11 years) 193–4
  diet 163–4
  digital technology use 481–6
  egocentric speech 200
  enculturation of social norms 206–13
  false beliefs 232–3
  feral children 444
  formal operational stage (age 11 to 16 years) 194
  internal working model 240
  internalization of motivation 432–5, 436
  motor skills, development of 482–3
  observational learning 146–7, 151
  peer influence 214–20
  poverty, effect on cognitive functions 164–5
  preoperational stage (age 2 to 7 years) 193
  reading and writing skills 482–3, 485
  resilience 223–4
  sensorimotor stage (birth to 2 years) 192
  separation distress 238, 240, 243
  social learning of aggression 146–7, 151, 379–82, 384
  and socioeconomic status 227
  stages of development 191–2
  theory of mind 231–3, 458–9, 461
  see also adolescents
chimpanzees, theory of mind research 230–1, 236
choice 418–19, 435, 436
chromosomes 250
chronic stress 310
Cialdini, Robert 371–6
classical conditioning 101–4, 105, 108–9
Clever Hans (horse) 5–6
clinical trials 262–3, 264, 265
clinician variables 288, 453
codes of ethics 54, 279
cognitive abilities, and motivation 438–43
cognitive appraisal 267
cognitive behavioural therapy (CBT) 113, 261, 268, 273–4, 275
  adherence to therapy 277
  behavioural exercises 276
  CBT apps 279–81
  codes of ethics 279
  computerized CBT 277–8
  effectiveness of 271, *278*, 282
  ethical issues 282
  how CBT works 276
  and motivation/agency 276–7, 282, 427–8
cognitive biases 136–42, 268, 273
  anchoring bias 139
  conceptual analysis 141–2
  confirmation bias 39, 137–8, *140*
  debiasing 140, 170–1
  framing effect 136–7
  measurement of 142
cognitive constructs 111
cognitive development 192–7, 204
  bias 197, 205
  brain development 188
  conceptual analysis 196–7
  cultural variation 458–62
  and digital technology use 481–6
  individual differences 194–5
  measurement 197
  and peer influence 218–19
  Piaget's theory 192–4, 196, 199
  and poverty 164–5
cognitive dissonance 148–9, 372
cognitive egocentrism 192, 193, 195, 203
cognitive explanations, interpersonal relationships 395–402
cognitive function, restoration of to normal state 169–70
cognitive learning 100
cognitive load theory 131–2, 485
cognitive models 124–33
  conceptual analysis 132–3
  cross-cultural variation 133
  historical context 125
  multi-store memory model 125, 126–9, 132
  scientific models 125
  working memory model 129–32
cognitive perspective 117–18
cognitive processes 99–177
  biological factors 117–23
  brain injury 120–1
  causality 159, 175
  classical and operant conditioning 100, 101–9
  cultural factors 152–60, 445, 446, 463–7
  digital technology use 487–91
  distributed cognition 487, 490
  environmental factors 161–7
  improvement of 168–76
  increasing the efficiency of cognitive function 170–2
  localization of function 120, 122
  restoration of cognitive function 169–70
  schema theory 110–16
  thinking 135–6
cognitive revolution 103, 125
cognitive schemas 110–16
  conceptual analysis 115–16
  development of 113–14
  research challenges 114–15
  usefulness of 114
cognitive structures 461
cognitive styles 269–70, 463–4, 466
cognitive triad 267
Cohen's *d* 83, 86
cohort effects 303
collectivism 284–5, 286, 289, 302, 344, 349
commitment and consistency 372
commitment, interpersonal relationships 397–401
communication
  conceptual analysis 411–12
  conflict resolution 405–8
  four horsemen of apocalypse, Gottmans' 407–10
  improvement strategies 408–10
  instrumental communication 405
  interpersonal communication 404–5, 411–12
  relational communication 405
  in relationships 403–12
  self-presentational communication 405
  symmetrical and asymmetrical communication 405
  see also digital technologies
comparison studies 26, 27, *28*
competence, perceived 420
compliance techniques 149, 370–7
  conceptual analysis 376–7
  conformity comparison 333, 371
  ethical issues 377
  foot-in-the-door 373–4, 375, 376
  low-ball technique 374–6
  principles of compliance 371–3, 376
  situation-dependency 377
compliance with treatment 453, 455, 457
computer analogy 118
computerized CBT (CCBT) 277–8
concepts *iv*, viii–xi, 4
conceptual thinking 200
concordance rate 251–2
concrete operational stage (age 7 to 11 years) 193–4
conditioned response (CR) 103
conditioned stimulus (CS) 103
conditioning 102
  classical conditioning 101–4, 105, 108–9
  operant conditioning 100, 101, 102, 105–9, 111
confidentiality 55
confirmation bias 39, 137–8, *140*
conflict
  realistic group conflict theory 336
  social media addiction 301
conflict resolution 340, 405–8
conformity 362–9
  Asch study of 363–6, 368–9
  biological basis of 366–7
  compliance comparison 333, 371
  conceptual analysis 368–9
  and culture 369
  definition 363
  evolution of 366–7, 368, 369
  robot experiment 367–8
  social identity theory 36
  types of 363
confounding variables 17, 22
conservation, concept of 193, 459–60, 462
CONSORT statement 264
construct validity 14, 21–2, 142, 348, 349
constructs 9, 109, 111
contact comfort hypothesis 238
contempt 407
content analysis 94–5, 431
context sensitivity, social norms 207
continuity 191, 192
control-determined behaviours 419
control groups 23
controlled processing 135
convenience (opportunity) sampling 18–19, 42
conventionality, social norms 207
conversational AI 8, 279, 280, 281, 406, 469
cooperation 349
coping strategy 311
correlation coefficient 32
correlation matrix 346
correlational studies 10, 31–5, 497
  bias 35
  credibility 35
  effect size and statistical significance 32–3
  limitations of 34, 400–1
  peer influence 217, 220
  sampling and generalizability 35
correlations 31–2
  curvilinear relationships 34, 35
  examples of *32*
  multiple variables 165
  negative correlations 32
  pairwise correlations 346
  positive correlations 32
  and scatterplots *77*
  spurious correlations 34, 35
  third variable problem 34, 35
cortical remapping 59–60, 183
cortisol 225, 309
Cotard's delusion 272
counter-stereotypical information 171
counterbalancing 21
couples therapy 410–11
courtship behaviour 388–9
covariates 165
covert observation 43
COVID-19 pandemic, and stress 311–12
CR *see* conditioned response
credibility 22, 36–8
  correlational studies 35
  credibility checks 38
  establishing a rapport 37
  iterative questioning 38
  reflexivity 38
  thick descriptions 38
  triangulation 37
critical periods in development 184
criticism 407
Cronbach's alpha 49
cross-cultural research 217, 350–1, 445–6, 497
CS *see* conditioned stimulus
cultural differences
  mental health 453–7
  romantic relationships 448–9, 451–2
cultural dimensions 343–51, 445
  bias in 351
  conceptual analysis 350–1
  correlation matrix 346
  evaluation of Hofstede's model/surveys 347–9
  measurement of 345–7
  national differences 345, 350
  and social identity theory 349–50
  and within-nation heterogeneity 348
cultural evolution 208
cultural expression of symptoms 453
cultural factors, cognitive processes 152–60
  conceptual analysis 159–60
  cross-cultural comparisons 154–5
  cultural differences in cognition 153
  cultural universals 155–6
  derived etic 155–6
  emic and etic approaches 153–4, 158, 159
  imposed etic 154–5, 158, 159
  memory study, Kpelle society 156–8
cultural identification 356
cultural practices 354, 356
cultural relativism 447
cultural style 445
cultural tools 201–2, 205
cultural values 351, 356
cultural variation 351, 445
  cognitive development 458–62
culture 152, 333, 444–67
  acculturation 352–61
  attachment 241–2, 243
  and cognitive development 458–62
  and cognitive processes 152–60, 445, 446, 463–7
  collectivistic cultures 284–5, 286, 289, 302
  conformity 369
  culture acceptance 355
  definition 344
  developmental psychology 458–62
  emic approach 344
  interpersonal relationships 447–52
  and learning 463–7
  mental disorders 283–90
  and resilience 223–4
culture–gene coevolution 285–7, 290
curiosity 441
curvilinear relationships 34, 35, 476, 481

Darwin, Charles 179–80
DAS *see* Dysfunctional Attitudes Scale
data analysis/interpretation
  assessment of 4
  descriptive statistics 64–70
  inferential statistics 80–93
  plotting and graphing 71–9
  qualitative data 94–7
data collection
  internal assessment 498–500
  surveys and questionnaires 11, 46–53
data fabrication 56, 57
death-related thoughts, terror management theory 421–5
debiasing 140, 170–1
debriefing 55
deception 55, 133, 369
decision-making 135, 140
declarative memory 120
defensiveness 407
delusional thinking 272
demand characteristics 25, 411
dendrites 253
dependent variables (DV) 10, 15, 16–17
  measurement of 260, 265
depression
  in adolescents 260–2
  aetiology of 285–7, 289
  Beck's cognitive theory of depression 267–73
  chemical imbalance theory 262
  cognitive styles 269–70
  cognitive theory of 267–73
  compliance with treatment 455, 457
  concordance rate 251–2
  cultural aetiology of 285–7, 289
  diagnosis 288, 454–5
  drug treatment 254, 258–9, 265, 272, 455
  epigenetic influences 293–4
  genetic heritability 251, 252
  misunderstanding of 247
  molecular genetics study 255–6
  neurotransmitters, role of 253–7
  paralinguistic features 288
  person-to-person spread of depressive symptoms 295–6
  prevalence 248
  reporting bias 287–8
  self-schema 113
  serotonin hypothesis 254–5
  and social media use 479, 480
  sociocultural factors 293–8
  and stressful life events 255–6, 293–4
  symptoms 248
  therapy effectiveness *278*, 426–31
  treatment of adolescents with depression (TADS) study 260–2
  vocal biomarkers 288
  vulnerability factors 295
  see also cognitive behavioural therapy
deprivation 222
depth perception 153
derived etic 155–6
descriptive models 136
descriptive statistics 64–70
  central tendency 67–8
  dispersion measurement 68–70
  measurement levels 64–5
  normal distribution 65–7
determinism 150
developmental neuroscience 183, 187–8
developmental psychology 179–245
  adverse childhood experiences 222–8
  attachment 237–44
  brain development 179, 181–9
  childhood experiences 221–8
  continuous models of human development 190–7
  and culture 458–62
  enculturation of social norms 206–13
  extrinsic motivation 433–4
  intrinsic motivation 432, 435, 436
  and motivation 432–7
  peer influence 214–20
  sociocultural factors 198–205
  stage theories 190–7
  theory of mind 229–36
  see also adolescents; children
developmental unevenness 195
diagnosis
  cultural differences in 454–5
  mental disorders 247–8, 288, 454–5
  obesity 303
  social media addiction 300–2
Diagnostic and Statistical Manual of Mental Disorders (DSM-5) 247, 248, 301
diathesis–stress model 256, 257, 269
diet
  breakfast consumption 163–4
  supplements 162
differentiation 182
digital technologies 468–91
  addiction to 470
  and child development 481–6
  and cognitive development 481–6
  cognitive processes 487–91
  and embodiment 489–90, 491
  and interpersonal relationships 469, 471–5
  mediating variables 471, 475, 476, 482
  and memory 487–9, 491
  and mental health 469, 473–4, 476–80
  negative/positive effects of 469, 470, 475
  research methodology 470
  uses of 470
  wearable technology 489–90, 491
  see also social media
direct learning 144
discontinuity 191, 192
discrimination
  implicit prejudice 171
  out-group discrimination 336, 337–8, 341, 421–5
discrimination paradigm 184
dispersion, measures of 68–70
distinctiveness, social identity 336, 338–9
distortion of action 365
distortion of judgement 365
distortion of perception 365
distress 310
  separation anxiety/distress 238, 240, 243
distributed cognition 487, 490
divorce 410
dizygotic twins 252
DNA (deoxyribonucleic acid) 250–1
  bases 250
  methylation 226, 227, 292
  SERT expression 293, 294
DNA sequences, knockout technique 61
dominant respondent bias 39
domino causality 225
dopamine 387, 389–90, 392
dopaminergic pathway 390
double-blind designs 26
drug development
  ethical issues 265
  randomized controlled trials 259, 262–4,265
drug treatment, for depression 254, 258–62, 265, 272, 455
DSM-5 (Diagnostic and Statistical Manual of Mental Disorders) 247, 248, 301
dual-process model of thinking 135–6
duality of human cognition 135
DV *see* dependent variables
Dysfunctional Attitudes Scale (DAS) 270

e-therapy 469
Eagleman, David 170
ecological validity 14, 22, 166, 217
education 342
Efe people, caretaking practices 241–2
effect size 32–3, 83, 320
EFT *see* Embedded Figures Test
egocentric speech 200
Embedded Figures Test (EFT) 463
embodiment, wearable technology 489–90, 491
emic approach 153–4, 158, 159, 344, 350
empirical evidence 5
enculturation 206–13, 353
  conceptual analysis 212–13
  observational learning 210–12
entrapment 400
environment
  and behaviour 149–50
  and cognitive processes 161–7
  and epigenetic influences 292–3
  in mental disorders 291–8
  and personal factors 149
epigenetic changes 226–7, 293
epigenetics, and environment 292–3
epistemic division 118
epistemological reflexivity 38
epistemology 118
equity theory 396–7
establishing a rapport 37, 39
ethics xi, 54–7
  animal research 62–3, 236, 239, 244
  anonymity 55, 56
  codes of ethics 54, 279
  cognitive behavioural therapy 273–4
  communication studies 412
  compliance techniques 377
  confidentiality 55
  data fabrication 56, 57
  debriefing 55
  deception 55, 133, 369
  drug development 265
  ethics committees 56
  genetic testing 257
  informed consent 55, 213
  internal assessment 497, 502–3
  intervention programmes 330
  Little Albert experiment 56
  protection from harm 55, 369
  randomized controlled trials 262
  reporting of research results 56–7
  research methodology 54–7, 298, 314
  romantic relationships research 412
  sharing research data for verification 56–7
  social implications of reporting scientific result 57
  therapy treatment 282
  withdrawal from participation 55
etic approach 153–4, 159, 350
eustress 309
evolution, of conformity 366–7, 368, 369
exam questions/papers *see* external assessment
excitatory neurotransmitters 387
executive functions 218, 219, 220
expected-utility theory 136
experience-dependent neuroplasticity 183, 185–6
experience-expectant neuroplasticity 183, 184–5
experimental designs 19–21
  counterbalancing 21
  double-blind design 26
  independent measures design 19–20
  matched pairs design 20
  order effects 21
  repeated measures design 21
experimental mortality 25
experimental studies 217
experimenter bias 25–6
experiments 10, 16–30
  bias 23–6
  cause-and-effect inferences 16–17
  confounding variables 17, 22
  control groups 23
  design of 19–21
  ecological validity 14
  field experiments 29
  group allocation 26–7
  internal validity 14–15
  threats to 23–6
  laboratory experiments 29
  natural experiments 29
  non-experiments 27–8
  quasi-experiments 28, 29
  random group allocation 26
  sampling 17–19
  true experiments 26–7
  validity 21–3
explicit attitude 170
external assessment 505–22
  exam tips 510–15
  practice questions 177, 245, 331, 413
  sample exam paper 516–22
  structure of exam papers 505–9
external validity 22
externalizing behaviours 216
externally mediated activities 201
extinction 105
extraneous cognitive load 131
extreme responding bias 52
extrinsic motivation 419–20, 433–4, 439–41
eye-tracking, theory of mind research 233–4

Facebook 479
factor analysis 344
fallacies
  treatment-aetiology fallacy 273
  treatment fallacy 322
false beliefs 232–3, 459
false consensus effect 217
falsifiable/falsifiability 5, 138
family, and peer influence 218
family stress model 164–5
favouritism, in group 336, 337–8, 341, 421–5
fear of missing out (FOMO) 328, *329*
feral children 444
Festinger, Leon 148–9
field dependence 463, 464–7
field experiments 29
field independence 463, 464–7
fight or flight response 308, 309, 310
fish, conformity study 367, 369
fixed action patterns 99
flashbulb memories 122
flatworms, learned behaviour study 99–100
flow 311, 312
fluoxetine (Prozac) 260
fMRI *see* functional magnetic resonance imaging
FOMO (fear of missing out) 328, *329*
foot-in-the-door 373–4, 375, 376
formal operational stage (age 11 to 16 years) 194
four horsemen of apocalypse, Gottmans' 407–10
four-card problem, Wason's 138
framing effect 136–7
free recall 156–7, 158
free will versus determinism 150
frequency distribution charts 72, *73*, *79*
Freud, Sigmund 247, 418

friendships
  cultural differences 448–9, 452
  and digital technologies 471
  formation of 218
functional magnetic resonance imaging (fMRI) 121, 187, *234*, 390

GA-MET programme 326–8
gene–culture coevolution 285–7, 290
gene expression 226, 292, 293
generality, social norms 207
generalizability 13–14, **15**, *22*, 40–1
  case-to-case generalization 14, 42
  in correlational studies 35
  ecological validity 14
  population validity 14
  sample-to-population generalization 14, 42
  theoretical generalization 42
generalization of stimulus 104
generative AI 8, 279, 280, 281, 406, 469
genes 250
genetic epistemology 192
genetic inheritance, mental disorders 250–2
genetic relatedness 251
genetic testing, ethical issues 257
genetics
  epigenetic changes 226–7, 293
  epigenetics and environment 292–3
  gene expression 226, 292, 293
  genetic heritability 250–2
  gene–culture coevolution 285–7, 290
  molecular genetics 252, 255–6
genotypes 251, 293
germane cognitive load 131
GI *see* glycaemic index
globalization 358
glucocorticoid receptor (GR) gene 226
glycaemic index (GI) 163
goodness-of-fit 126
Gottman, John and Julie 407–10, 412
GR *see* glucocorticoid receptor gene
graphs and plotting 71–9
  bar graphs 74–5
  box-and-whisker plots 75–7
  categorized charts *79*
  frequency distribution charts 72, *73*, *79*
  histograms *66*, 71–2
  scatterplots 31–2, 77–8, *79*
Gray, Carol 233
greylag geese, fixed action patterns 99
grounded cognition theory 485
grounded theory 96–7
group allocation, experiments 26
group behaviour, social learning in 378–85
group membership
  and cultural dimensions 349–50
  distinctiveness 338–9
  in-group favouritism 336, 337–8, 341, 421–5
  out-group discrimination 336, 337–8, 341, 421–5
  positive social identity 338–9
  social identity theory 335–42
  terror management theory 421–5
Guerrero, L.K. 404–6

Harlow, Harry 238–9, 243, 244
Hatfield, Elaine 404
HCI *see* human–computer interaction
health
  and acculturation 357, 360, 361
  and digital technology use 476–80
  and motivation 426–31
health behaviours 316–17
health problems
  causality 313
  changes over time 303–4, 306
  conceptual analysis 305–6, 321–2, 329–30
  dynamics of change 322
  harmful/helpful strategies 322
  intervention programmes 318, 322, 329, 330
  and personal responsibility 317
  prevalence rates 300, 305–6
  prevention programmes 316, 318–20, 321
  prevention strategies 323–30
  self-control 324–6, 330
  social learning theory 315–22
  social media addiction 300–3, 329–30
  sociocultural perspective 315–22
  stress 307–14
  see also depression; mental disorders
hearing 153
Heraclitus 179
heuristics 136
higher-order cognitive functions 199–200, 202
higher primates, theory of mind 230–1, 234, 236
histograms *66*, 71–2
Hofstede, Geert 154–5, 344–8
holism 119
holistic cognitive style 463–4, 466
homeostasis 308
hormones 387–8
human agency 144
human–computer interaction (HCI) 490
human development *see* developmental psychology
human psyche 247

IA *see* internal assessment
IAT *see* implicit association test
identification with the model 379
identity 179
identity perspective 215
idiographic approach 12
IEQ *see* indoor environmental quality
immigrant paradox 357–8
immigrants, acculturation 357–8
immigrants, diagnosis of mental disorders 454, 456
immune system 316
implicit association test (IAT) 171–2
implicit prejudice 171
imposed etic 154–5, 158, 159, 212
in-group favouritism 336, 337–8, 341, 421–5
independent measures design 19–20
independent variables (IV) 10, 14–17
indirect learning 145
individual differences, adverse childhood experiences 227
individual and national level of analysis 348
individualism–collectivism 284–5, *286*
individualism versus collectivism 344, 349
indoor environmental quality (IEQ) 162
inductive content analysis 95–6, 431
indulgence versus restraint 345
infants, linguistic abilities study 184
inferential statistics 80–93
  choice of inferential test 84
  effect size 83
  Mann–Whitney *U* test 86–8, 91–2
  null hypothesis (H_0) and alternative hypothesis (H_1) 81, 82
  parametric and non-parametric tests 83
  Pearson's correlation coefficient 88–9, 92
  reporting and interpretation of results 90–3
  Spearman's rank correlation coefficient 89–90, 92–3
  statistical significance 82, 83
  type I and type II error rates 81–2
  unrelated *t*-test 84–6, 90–1
infidelity, interpersonal relationships 449–51, 452
informational influence 363
informed consent 55, 213
inhibitory neurotransmitters 387
innate behaviours 99, 100
instrumental communication 405
integration, acculturation strategies 353
intelligence 57, 67, 155, 458
intentions 231–2
interaction of factors, and additive influence 294
interactive technology 481, 482
internal assessment (IA) 493–503
  background research **495**
  correlational studies 497
  data collection 498–500
  discussion section 501–2
  ethical issues 497, 502–3
  introduction section 494–5
  research methodology section 496–7
  research proposal
  ethical issues 502–3
  presentation of 503
  surveys/questionnaires 499
internal validity 14–15, 22, 166
  threats to 23–6
internalization *201*, 432–5, 436
internalizing behaviours 216
internet 469, 470, 487
interpersonal communication 404–5, 411–12
interpersonal process 200, *201*
interpersonal relationships *see* relationships
interval-level variables 64–5
intervention programmes 318, 322, 329
  effectiveness measures 321, 330
  ethical issues 330
interviews 44
  semi-structured interviews 44
  structured interviews 44
  unstructured interviews 44
intracultural variation 241
intrapersonal process 200, *201*
intrinsic cognitive load 131
intrinsic motivation 419–20, 432, 435, 436, 439–41
Inuit 459–60
investment, interpersonal relationships 397
investment model of commitment 397–401
  empirical support for 398–400
  research challenges/limitations 400–2
  role-playing experiment 398–9
IQ scores/tests 57, 67, 155
Iran, relationship therapy 410–11
irreversibility 193

iterative questioning 38
IV *see* independent variables

jealousy 449–51, 452
judgement, distortion of 365
Junín Quechua community 458–9, 461

Kahneman, Daniel 135–6, 137
Kpelle society, memory study 156–8
kurtosis 73

laboratory experiments 29
laboratory observation 43
lack of conservation 193
language 202
  infants' abilities 184
  mental states, concepts of 458–9
  Word Fluency Test 186–7
large language models (LLMs) 174, 406
law of effect 105
leading questions bias 39
learned behaviours 99, 100, 101–9
learning 99
  classical conditioning 101–4, 105, 108–9
  cognitive learning 100
  and culture 463–7
  direct learning 144
  indirect learning 145
  and maturation 179
  and motivation 438–43
  observational learning 144–7, 151, 210–12, 316, 379–82, 384
  operant conditioning 100, 101, 102, 105–9, 111
  reinforcement learning 383
  schema theory 110–16
  zone of proximal development 201
  see also social learning theory
Leontyev, A.N. 202–3, 205
lifetime prevalence 300
Likert scale **65**
liking, compliance techniques 372
line comparison task, conformity study 363–6, 368
Little Albert experiment 56, 104
LLMs *see* large language models
localization of function 120, 122, 174, 186, 234–5
locked-in syndrome 174
logical thinking 138, 194
long-term effects 470, 476
long-term memory (LTM) 127, 128, 132
longitudinal studies 215, 217, 220, 225, 228, 399–400
Lorenz, Konrad 99
love 388–90, 391
low-ball technique 374–6
low-order cognitive functions 199–200
LTM *see* long-term memory

Maasai people, adverse childhood experiences 223–4
major depressive disorder (MDD) *see* depression
Mann–Whitney *U* test 86–8
  calculations of **87**
  reporting and interpretation of results 91–2
marginalization, acculturation strategies 353, 355
marital satisfaction questionnaires 408, 411
marriage counselling, role-play exercise 409
matched pairs design 20
matching variables 20
maternal stress 293
maturation 179, 189
maturation bias 23
MDD (major depressive disorder) *see* depression
mean 67–8, 69, 83
measurement x
mechanistic motivation 418–19
median 68
mediating factors 145
mediating variables *145*, 471, 475, 476, 482
memory
  declarative memory 120
  and digital technologies 487–9, 491
  flashbulb memories 122
  impairment of 120
  in Kpelle society 156–8
  long-term memory 127, 128, 132
  and motivation 439–41
  multi-store memory model 125, 126–9, 132
  Newcastle Spatial Memory Test 439–40
  procedural memory 120
  semantic memory 157, 200
  sensory memory 126
  short-term memory 127, 128
  working memory model 129–32, 218
mental disorders 247–98
  biological explanations 249–57
  biological treatment 254, 258–65, 272, 455
  causality 256–7
  cognitive appraisal 267
  cognitive models 266–74
  conceptual analysis 272–4,289–90, 296–8

cultural aetiology of 284–8
cultural differences 283–90, 453–7
cultural expression of symptoms 453
diagnosis 247–8, 288, 454–5
environmental factors 291–8
genetic heritability 250–2, 257
prevalence rates 287–8
psychological treatment 275–82
reporting bias 287–8
as a social problem 297–8
social spread of symptoms 295–6
sociocultural factors 291–8
see also depression
mental health
definition 247
and digital technology use 469, 473–4, 476–80
mental processes 7
mental representations 110, 111–12
mental states (beliefs, intentions, knowledge) 230
meta-analyses 215–16, 220, 277–8
methylation 226, 227, 292
microanalytic observation of behaviour 407–8
migrants
acculturation 357–8
diagnosis of mental disorders 454, 456
migration of neurons 182
mind, theory of *see* theory of mind
mindfulness 328–9
minimal group paradigm 336, 337–40, 341, 342
mirror neurons 234–5
MKO *see* more knowledgeable other
mobile phones 471
mode 68
Molaison, Henry 120
molecular genetics 252, 255–6
monothetic classification 301
monozygotic twins 252
mood modification, social media addiction 301
moral outrage 383
more knowledgeable other (MKO) 214
motivation 145, 147, 417–43
autonomous motivation 426–31
cognitive abilities 438–43
cognitive behavioural therapy 276–7, 282
concept of 417
developmental psychology 432–7
etymology of 417
extrinsic motivation 419–20, 433–4, 439–41
internalization of 432–5
interpersonal relationships 421–5
intrinsic motivation 419–20, 432, 435, 436, 439–41
mechanistic motivation versus choice 418–19
memory 439–41
mental health 426–31
self-determination theory 419–20
social media addiction 326–8
theories of, overview 418
motivational debiasing strategy 140
motivational enhancement therapy 326
motor reproduction 145
motor skills, child development 482–3
multi-store memory model (MSMM) 125, 126–9, 132
experimental test of 127–9
multiple variables 165

narrative recall 157–8
natural experiments 29
naturalistic observation 43
nature versus nurture 150, 179
negative correlations 32
negative reciprocity 406
negative reinforcements 100, 105
negative schemas 267, 268
negative thinking 273, 275
nervous system 253–4, 387
neurogenesis 182
neurons
migration of 182
mirror neurons 234–5
pruning of 183
structure of 253
neuroplasticity 169, 174, 183, 184–6, 189, 490
neuroprosthetics 172
neurotransmission 253–4, 390
neurotransmitters 387, 392
imbalance of 253–7
Newcastle Spatial Memory Test (NSMT) 439–40
noise 162
nominal-level variables 64
nomothetic approach 9
non-experiments 27–8
non-parametric tests 83
non-reductionism (holism) 119
see also holistic cognitive style
non-verbal messages 404–5
normal distribution 65–7
normative influence 363
normative models 136, 142
normativity, social norms 207
norms 207, 383, 384
NSMT *see* Newcastle Spatial Memory Test
null hypothesis (H_0) 81, 82
nursery schools, digital technology use 482–3
nutrition
breakfast consumption 163–4
supplements 162

Obama effect 171
obesity 248
changes over time 303–4
diagnosis 303
prevalence rates 304
prevention programmes 318–20, 321
object permanence 192
observation 42–3
covert observation 43
laboratory versus naturalistic observation 43
observational studies 244
outliers 67
overt observation 43
participant/non-participant observation 43
structured versus unstructured observation 43
observational learning 144–7, 316, 379
of aggression 146–7, 151, 379–82, 384
Bobo doll experiment 146–7, 379, 380
in enculturation 210–12
Odden, Harold 210–12
office schema, study of 114–15
ontology 118
operant conditioning 100, 101, 102, 105–9, 111
applications of 107–8
conceptual analysis 108–9
pigeon study 106–7
operant conditioning chamber (Skinner box) 105–6, 109
operant variability 108
operationalizations
of constructs 10, 21–2
social learning 385
opportunity sampling 18–19, 42
ordinal-level variables 64
Osten, Wilhelm von 5–6
out-group discrimination 336, 337–8, 341, 421–5
out-of-reach zone 201
outliers 67, 72, *76*, *77*, *78*
overt observation 43
owl monkeys, cortical remapping research 59–60
oxytocin 388, 391

pair bonding 391–2
pairwise correlations 346
panic attacks, therapy effectiveness *278*
paralinguistic features, of depression 288
parametric tests 83
parental investment model 165
parenting style 218
parsimony 126
participant bias **15**, 39
acquiescence bias 39
dominant respondent bias 39
sensitivity bias 39
social desirability bias 39
participant observation 43
partner attachment 388, 391–2
paternal investment 450
patient variables 288, 453
Pavlov, Ivan 102, 103–4
Pearson's correlation coefficient 88–9
reporting and interpretation of results 92
Pearson's *r* 83
peer influence 214–20
and cognitive development 218–19
conceptual analysis 219–20
experiments versus correlational studies 217
peers and family, relative contributions of 218
research challenges 216–17
selection effects 216–17
self-report measures 217, 220
perceived competence 420
perceived locus of causality 419–20
perceived quality of alternatives 397
perception, distortion of 365
performance accomplishments 316
period effects 303
period prevalence 300
person-to-person spread of depressive symptoms 295–6
personal factors
and behaviour 148
and environment 149
personal reflexivity 38
perspective x
Pfungst, Oskar 6
Phase 0, I, II, and III trials 262, 265
phenotypes 251
phonological loop 129, 130
phonological similarity effect 129, 130
phonological store 129
physical activity, treatment for depression 428–9, 431
Piaget, Jean 113, 192–4, 196, 199, 200, 204, 214–15, 461
pigeons, operant conditioning study 106–7
placebo effect 259
placebo group 259, 261, 263, *274*
plotting *see* graphs and plotting
point prevalence 300
polythetic classification 301
population validity 14, 22, 42
positive correlations 32
positive–negative asymmetry phenomenon 341
positive reinforcements 105
post-traumatic stress disorder (PTSD) 222
poverty 297
and child development 222
and cognitive development 164–5, 166–7
power distance 344, 349
predictive validity 50, 400
prejudice *see* discrimination
prenatal stress 293
preoperational stage (age 2 to 7 years) 193
prevalence rates
bias 306
depression 287–8
health problems 305–6
measurement of 300
obesity 304
social media addiction 302–3, 305
types of 300
prevention programmes 316, 318–20, 321
primary appraisal 310
proactive self-control 325
procedural memory 120
prosocial behaviours 382
protection from harm 55, 369
proxy measures 354–5, 358, 361
Prozac (fluoxetine) 260
pruning 182, 184–5
psyche 247
psychoanalysis 418–19
psychology 3, 5–8
etymology of 7
IB Psychology, focus of 7
psychometrics 49–50
psychotherapy, and autonomous motivation 426–31
PTSD (post-traumatic stress disorder) 222
publication bias 263–4, 321
punishments 105, 106, 145
purposive sampling 41

qualitative research 11–12, **12**, **15**, 36–45
bias 38–40
case studies 45
content analysis 94–5
credibility 36–8
focus groups 44–5
grounded theory 96–7
interviews 44
observation 42–3
sampling and generalizability 14, 40–2
thematic analysis 95–6
transferability 42
trustworthiness 36–8
quantitative research 9–11, **12**, **15**
correlational studies 10, 31–5
experiments 10
quantitative descriptive studies 11, 51
sampling and generalizability 14
quartiles 69, *76*
quasi-experiments 28, 29
question order bias 39–40
questionnaires 11, 46–53
acculturation 354, 361
bias 51–2
construction of 46–8, 52–3
internal assessment 499
internal consistency 49
marital satisfaction 411
metrics 48, 49
predictive validity 50
psychometrics 49–50
survey comparison 50–1
test-retest reliability 49
quiz questions study 488
quota sampling 41

radical behaviourism 107
random sampling 18, 41
randomization 263
randomized controlled trials (RCTs) 259
bias in 263–4
CONSORT statement 264
criticism of 262
rapport 37, 39
Rasbult, C. 398–401
ratio-level variables 65
rational imitation paradigm 231
rats, mental representations study 111–12

RCTs *see* randomized controlled trials
reactive self-control 325
reading skills, child development 482–3, 485
realistic group conflict theory 336
recency effect 128
reciprocal determinism 147–8, 150, 151
reciprocity 371–2, 376
reductionism 118–19, 123, 257
reflexivity 38
refugees, diagnosis of mental disorders 454, 456
see also migrants
rehearsal 127
reinforcement learning 383
reinforcements 105, 106, 145
relapse rate 260
relapse, social media addiction 301
relatedness 420
relational communication 405
relationships 333–4
cheating 397
chemical messengers, role of 386–94
cognitive explanations for 395–402
commitment 397–401
conceptual analysis 411–12
conflict resolution 405–8
couples therapy 410–11
and culture 447–52
and digital technologies 469, 471–5
equity theory 396–7
four horsemen of apocalypse, Gottmans' 407–10
friendships 448–9, 452, 471
improvement strategies 408–10
infidelity 449–51, 452
interpersonal communication 404–5, 411–12
investment model of commitment 397–401
jealousy 449–51, 452
maintenance mechanisms 397
marital satisfaction questionnaires 408
motivation in 421–5
paternal investment 450
perceived quality of alternatives 397
relationship therapy 410–11
romantic relationships 448–9, 451–2, 472–5
social exchange theory 396
and social media 475
Sound Relationship House theory 408–10
reliability
questionnaires 49
surveys 348, 351
test-retest reliability 49
remission rate 260
remote acculturation 358–9
repeated measures design 21
replicatability 5
reporting bias 40, 287–8, 289
reporting of research results 56–7
representative samples 300
research methodology 3–63
analysis of research 13–15
assessment of 3
bias 14–15
central tendency and variability 446
correlational studies 31–5
credibility 14–15, *22*
digital technologies 470
ethics 54–7, 298, 314
experiments 10, 16–30
generalizability 13–14, **15**
informed consent 213
qualitative research 11–12, **12**, 14, **15**, 36–45
quantitative research 9–11, **12**, 14, **15**
questionnaires 46–53
reporting of research results 56–7
sampling 14, **15**
surveys 46–53
trauma 228
Vygotsky's approach 202
see also animal models/research
researcher bias **15**, 39–40, 235, 257
biased reporting 40, 287–8, 289
confirmation bias 39, 137–8, *140*
leading questions bias 39
question order bias 39–40
sampling bias 40, 289, 321
resilience 223–4, 227, 228
respect 407
response rate 260
responsibility xi
restraint 345
retention 145
retrograde amnesia 120
reward-seeking behaviour 390
rhesus monkeys, attachment studies 238–9
robot experiment, conformity study 367–8
Rochat, P. 210–12
romantic love 388–90, 391
romantic relationships 242, 388
attachment styles 472
cognitive explanations for 395–402
conceptual analysis 401–2
cultural differences 448–9, 451–2
and digital technologies 472–5
Ross, D. 146–7, 151
Ross, S.A. 146–7, 151
Rwandan Genocide (1994) 222

salience, social media addiction 301
Sally-Anne task 232, *233*, 234
Samoa, enculturation study 210–12
sample-to-population generalization 42
sampling 14, **15**, 17–19, 41–2
convenience (opportunity) sampling 18–19, 42
in correlational studies 35
purposive sampling 41
quota sampling 41
random sampling 18, 41
representative samples 300
representativeness 17
sample 17
sampling bias 40, 289, 321
self-selected sampling 19
snowball sampling 41
stratified sampling 18
target population 17
theoretical sampling 41
Sanichar, Dina 444
satisfaction, interpersonal relationships 397
scaffolding 195, 201, 203
scarcity, compliance technique 372
scatterplots 31–2, *77–8*, *79*
schema theory 110–16
conceptual analysis 115–16
development of schemas 113–14
mental representations 110, 111–12
negative schemas 267, 268
research challenges 114–15
schemas 112–13
usefulness of schemas 114
schematic processing 114
Schmidt, M.F.H. 208–10
schools, digital technologies in 482–3
Schwartz, S.J. 355, 356–7
scientific models 125
scientific studies 5–6
scientific theories 281
scripts 113
SDT *see* self-determination theory
secondary appraisal 311
secure base hypothesis 239
selection bias 23, 321
selection effects 216–17
selective activation of the amygdala 122
selective serotonin reuptake inhibitors (SSRIs) 254, 258–9
self-control 324–6, 330
self-determination theory (SDT) 419–20
autonomous motivation and therapy outcomes 426–31
intrinsic motivation 435, 436
self-determined behaviours 419
self-efficacy 144, 147, *277*, 316, 317, 431
self-esteem 336, 340–1, 473
self-presentational communication 405
self-report measures 217, 220, 272–3, 314
self-schema 113
self-selected sampling 19
Selye, Hans 308, 309
semantic memory 157, 200
semi-interquartile range 69–70
semi-structured interviews 44
sensitive (critical) periods in development 184
sensitivity bias 39
sensorimotor stage (birth to 2 years) 192
sensory augmentation 170
sensory memory 126
sensory substitution 169
sentiment analysis 95
separation, acculturation strategies 353
separation anxiety/distress 238, 240, 243
serial position effect 128
serotonin 254–5
serotonin transporter gene 285, 292, 293
SERT expression 293, 294
SES *see* socioeconomic status
SET *see* social exchange theory
sex drive 388, 391
sharing research data for verification 56–7
short-term effects 470, 476
short-term memory (STM) 127, 128
signs and symbols 202
SIT *see* social identity theory
skewness *72*, *73*, *77*
Skinner, B.F. 102, 105–8, 109
Skinner box (operant conditioning chamber) 105–6, 109
sleep deprivation, experiment 17, 20
SLT *see* social learning theory
snowball sampling 41
social categorization 336
social comparison 336, 478–9, 480
social desirability bias 39, 51
social environment 292
social exchange theory (SET) 396
social identity theory (SIT) 333, 335–42
bias in 341
cognitive component 341
conceptual analysis 341–2
conflict resolution 340
and conformity 363
cultural dimensions 349–50
distinctiveness 338–9
evaluation of 340–1
in-group favouritism 336, 337–8, 341, 421–5
limitations of 340–1
manipulation of social identity 342
minimal group paradigm 336, 337–40, 341, 342
operational definitions 341
out-group discrimination 336, 337–8, 341, 421–5
positive social identity 338–9
self-esteem 336, 340–1
Tajfel's minimal group experiments 337–40
social implications of reporting scientific results 57
social influence 363, 371
social learning theory (SLT) 143–51, 315–22, 333, 378–85
aggressive behaviour 146–7, 151, 379–82, 384
bias 151
bidirectional interactions 147, 148–50
conceptual analysis 150–1, 384–5
health behaviours 316–17
and health problems 318
obesity prevention programmes 318–20
observational learning 144–6
operationalizations 385
prosocial behaviours 382
public behaviour on social networks 383
reciprocal determinism 147–8, 150, 151
social norms 383
social media 469, 471
body image study 478–9, 480
and interpersonal relationships 475
and mental health 473–4, 476–80
moral outrage in 383
social-media-based interventions 319–20
uses of 470
see also digital technologies; social networks
social media addiction 248, 300–3, 305, 329–30
diagnosis 300–2
measurement of 313–14
and mindfulness 328–9
and motivation 326–8
prevalence rates 302–3
prevention strategies 324–9
and self-control 324–6
social learning 316
stress as causal factor 311–12
and time distortion 312–13
social networks
public behaviour on 383
and spread of mental disorder symptoms 295–6
see also social media
social norms 206–13, 383, 384
conceptual analysis 212–13
early development of 208
transmission of 208–10
social phobia, therapy effectiveness *278*
social proof, compliance techniques 372
social skills 218, 219
social stories 233
socialization 207, 432–3
sociocultural factors, in mental disorders 291–8
sociocultural perspective 117–18
sociocultural theory of human development 199–205
conceptual analysis 204–5
cultural tools 201–2, 205
low-order and higher-order cognitive functions 199–200
social sources of development 200–1
zone of proximal development 201
socioeconomic status (SES) 227
somatization 453, 454, 456
Sound Relationship House theory 408–10
spatial resolution 121
Spearman's rank correlation coefficient 89–90
reporting and interpretation of results 92–3
spurious correlations 34, 35
SSRIs (selective serotonin reuptake inhibitors) 254, 258–9
stages of development 190–7
arguments for and against 191–2
standard deviation 69
statistical power 82
statistical separation of effects 165
statistical significance 33, 82, 83, 320
statistics
descriptive statistics 64–70
inferential statistics 80–93
stereotypes 112, 116, 171, 340
stimulus, generalization of 104

STM *see* short-term memory
stonewalling 407
strange situation paradigm 240, 243
stratified sampling 18
strength-model of self-control 325
stress 307–14
  acute and chronic stress 310
  appraisal 310–11
  biology of 308–9
  cognitive dissonance 148–9
  conceptual analysis 313–14
  and COVID-19 pandemic 311–12
  definitions 308
  and depression 255–6, 293–4
  diathesis–stress model 256, *257*, 269
  fight or flight response 308, 309, 310
  maternal stress 293
  and performance 34
  post-traumatic stress disorder 222
  prenatal stress 293
  resilience 223–4
  and social media addiction 311–12
  stress hormones 308–9
  transactional model of stress and coping 310–11
  types of 309–10
Stroop tasks 488
structural equation modelling 271
structure–function relationships 186–7
structure validity 49
structured interviews 44
structured observation 43
styles of conflict resolution 405–6
suicide, and depression 248
superstition 107
surveys 11, 46–53
  bias 51–2
  construction of 52–3
  internal assessment 499
  questionnaire comparison 50–1
  reliability 348, 351
Sweller, John 131
symmetrical communication 405
symptoms, cultural expression of 453
synapses
  differentiation and pruning 182
  elimination of 182–3
  synaptic density *182*
  synaptic gap 253, 254
  synaptic plasticity 183
System I and System II thinking, Kahneman 135–6

t-tests 84–6
tablet computers 482
tactile television *170*
TADS *see* treatment of adolescents with depression study
Tajfel, Henry 336–42
target population 17
taxi drivers, brain scan study 27
technological debiasing strategy 140
technology *see* digital technologies
teeth-brushing study, autistic spectrum disorder children 483–4, 486
teleological behaviourism 112
telepathic communication 174
television *170*, 382
temperature, working conditions 162
temporal resolution 121
temporo-parietal junction (TPJ) 234
terror management theory (TMT) 421–5
test-retest reliability 49
testing effect 23–4
testosterone 391, 392
thematic analysis (inductive content analysis) 95–6
theoretical generalization 42
theoretical sampling 41
theory of mind 229–36, 458–9, 461
  animal research 230–1
  biological basis of 234, 235
  in children 231–3
  conceptual analysis 235–6
  eye-tracking research 233–4
  false beliefs 232–3
  in higher primates 230–1, 234, 236
  intentions 231–2
  localization of function 234–5
  measurement methods 236
therapy *see* treatment
thick descriptions 38
thinking
  automatic thoughts 267, 276
  conceptual thinking 200
  delusional thinking 272
  dual-process model 135–6
  logical thinking 138, 194
  negative thinking 273, 275
  System I and System II thinking 135–6
third variable problem 34, 35
Thorndike, Edward 102, 105
three mountains task 193, 203
time distortion 312–13
Tinbergen, Niko 99
TMT *see* terror management theory
tolerance, social media addiction 301
Tolman, E.C. 111–12
Tomasello, M. 208–10
TPJ *see* temporo-parietal junction
trade-off between internal and external validity 262
transactional model of stress and coping 310–11
transcription 292
transferability 14, 42
translation 292
trauma 222
  and cortisol levels 225
  post-traumatic stress disorder 222
  research studies 228
  SERT expression 294
treatment 258–65
  adherence to therapy 277, 282
  and autonomous motivation 426–31
  biological treatment for depression 254, 258–62, 265, 272, 455
  clinical trials 262–3, 264, 265
  compliance with 453, 455, 457
  conceptual analysis 264–5, 281–2
  cultural sensitivity of clinician 290
  cultural variation 453
  dependent variable, measurement of 260, 265
  effectiveness measures 260, *278*, 281
  and motivation 426–31
  randomized controlled trials 259, 262–4
  treatment-aetiology fallacy 273
  treatment fallacy 322
  treatment and prevention 323
treatment of adolescents with depression (TADS) study 260–2
triangulation 37, 186, 187
triune brain theory 61
true experiments 26–7
trust and cooperation 349
trustworthiness 36–8
turkeys, fixed action patterns study 99
twin studies 252
two-dimensional model of acculturation 354, 355, 356–7
Type A (avoidant attachment) 240
Type B (secure attachment) 241
Type C (ambivalent/resistant attachment) 241
type I error (false positive) 81–2, 381
type II error (false negative) 82

unblinding 263
uncertainty avoidance index 345
unconditioned response (UR) 103
unconditioned stimulus (US) 103
unidimensional model of acculturation 354, 360
universal patterns 447
unrelated *t*-test 84–6
  reporting and interpretation of results 90–1
unstructured interviews 44
UR *see* unconditioned response
US *see* unconditioned stimulus
USA, Facebook rollout 479

validity 21–3
  construct validity 14, 21–2, 142, 348, 349
  ecological validity 14, 22, 166, 217
  external validity 22
  internal validity 14–15, 22, 23–6, 166
  population validity 14, 22, 42
  predictive validity 50, 400
  questionnaires 49
  structure validity 49
  trade-off between internal and external validity 262
Values Survey Module (VSM) 348–9
variables 9, 64–5
  clinician variables 288, 453
  confounding variables 17, 165
  dependent variables 10, 15, 16–17, 260, 265
  independent variables 10, 14–17
  interval-level variables 64–5
  matching variables 20
  mediating variables *145*, 471, 475, 476, 482
  multiple variables 165
  nominal-level variables 64
  ordinal-level variables 64
  patient variables 288, 453
  ratio-level variables 65
  third variable problem 34, 35
vasopressin 391, 392
verbal compliance 374, 375
verbal messages 404–5
verbal persuasion 317
vested interests 263–4
vicarious experience 316–17
vicarious reinforcement (or punishment) 145
video games, social learning of violence 379–82, 384
visualization exercises 325
visuospatial sketchpad 129
vocal biomarkers 288
von Osten, Wilhelm 5–6
VSM *see* Values Survey Module
vulnerability factors of depression 295
Vygotsky, Lev 199–205, 214

waitlist control 259
Wankerl, M. 293–4, 297
Wason's four-card problem 138
WeChat (Chinese social network) 280
Weihenmayer, Erik 170
withdrawal from participation 55, 213
withdrawal, social media addiction 301
within-nation heterogeneity 348
WMM *see* working memory model
Word Fluency Test 186–7
word length effect 129, 130
working conditions 162
working memory model (WMM) 129–32, 218
writing skills, child development 482–3, 485

Zambia, remote acculturation study 358–9
Zhao, N. 311–12
Zhou, G. 311–12
zone of proximal development (ZPD) 201

Take learning online with Kerboodle

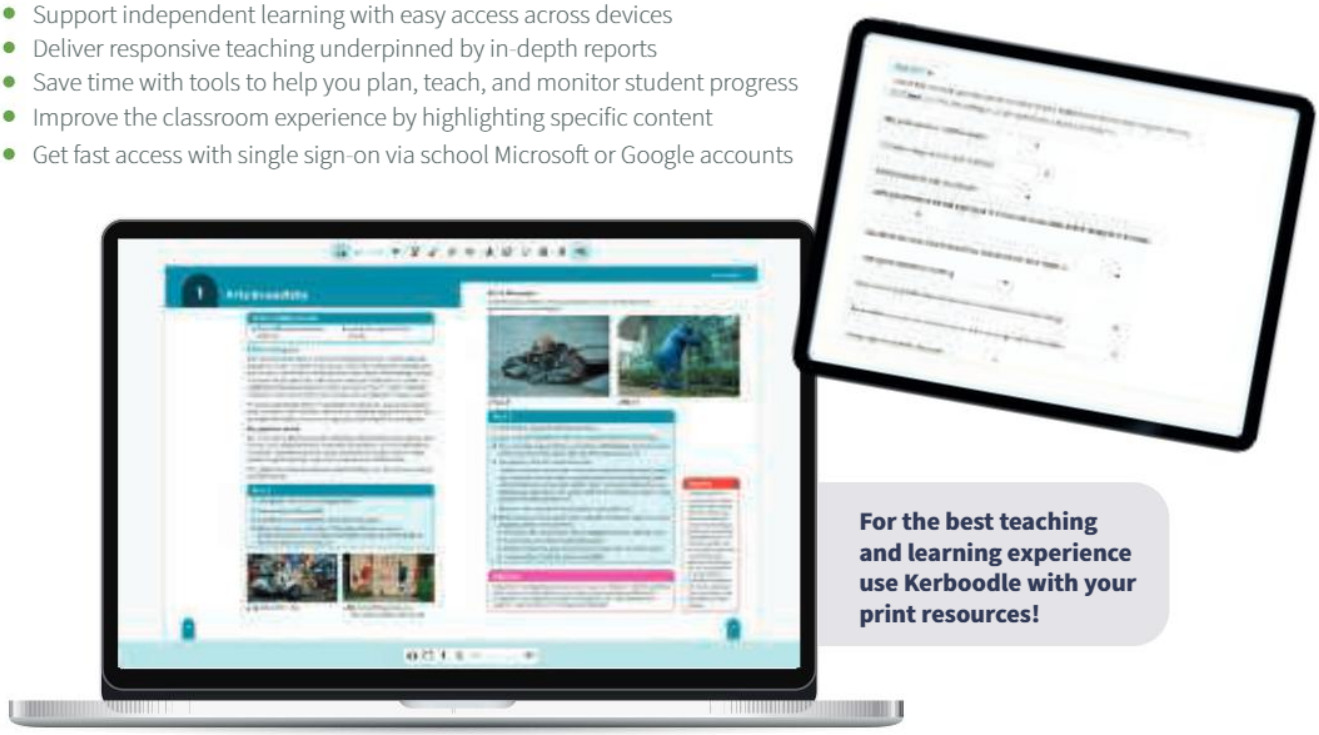
What is Kerboodle?

Kerboodle is a digital learning platform that works alongside your print textbooks to create a supportive learning environment. Available for UK and international curricula, Kerboodle helps you save time and reinforces student learning with a range of supportive resources.

Find out more and sign up for a free trial!

Use Kerboodle to:

- Enable learning anywhere with online and offline access to digital books
- Enhance student engagement with activities and auto-marked quizzes
- Boost performance and exam confidence with assessment materials
- Support independent learning with easy access across devices
- Deliver responsive teaching underpinned by in-depth reports
- Save time with tools to help you plan, teach, and monitor student progress
- Improve the classroom experience by highlighting specific content
- Get fast access with single sign-on via school Microsoft or Google accounts



For the best teaching and learning experience use Kerboodle with your print resources!

For more information, visit:
www.oxfordsecondary.com/kerboodle

Need help?
Contact your local Educational Consultant: www.oxfordsecondary.com/contact-us

Oxford Resources for IB
Diploma Programme

2025 EDITION

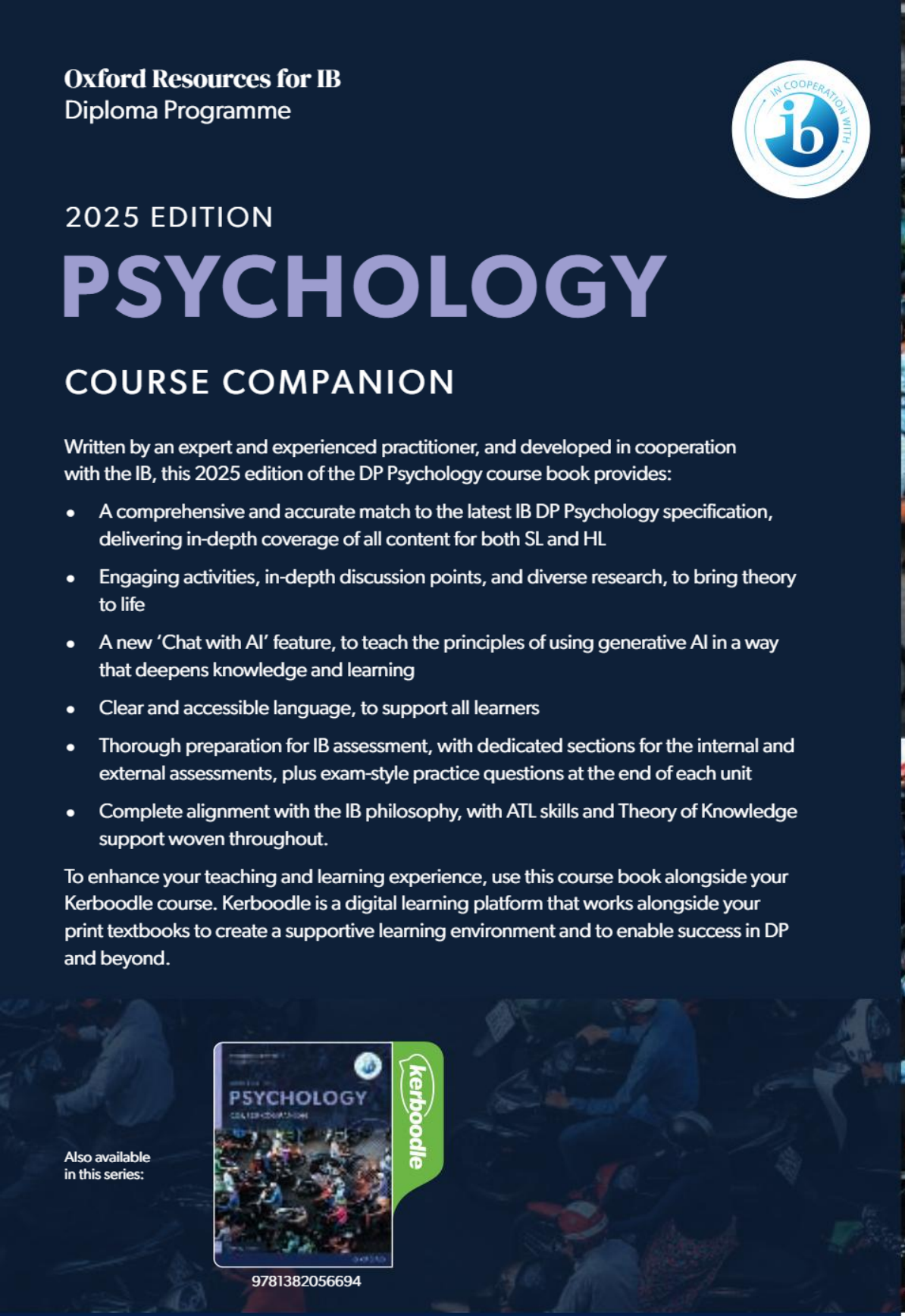
PSYCHOLOGY

COURSE COMPANION

Written by an expert and experienced practitioner, and developed in cooperation with the IB, this 2025 edition of the DP Psychology course book provides:

- A comprehensive and accurate match to the latest IB DP Psychology specification, delivering in-depth coverage of all content for both SL and HL
- Engaging activities, in-depth discussion points, and diverse research, to bring theory to life
- A new 'Chat with AI' feature, to teach the principles of using generative AI in a way that deepens knowledge and learning
- Clear and accessible language, to support all learners
- Thorough preparation for IB assessment, with dedicated sections for the internal and external assessments, plus exam-style practice questions at the end of each unit
- Complete alignment with the IB philosophy, with ATL skills and Theory of Knowledge support woven throughout.

To enhance your teaching and learning experience, use this course book alongside your Kerboodle course. Kerboodle is a digital learning platform that works alongside your print textbooks to create a supportive learning environment and to enable success in DP and beyond.



Also available in this series:

9781382056694

FOR FIRST ASSESSMENT IN 2027

OXFORD UNIVERSITY PRESS

www.oup.com

How to get in touch:
web www.oxfordsecondary.com/ib
email schools.enquiries.uk@oup.com
tel +44 (0)1536 452620

ISBN 978 1 382 05665 3
9 781382 056663