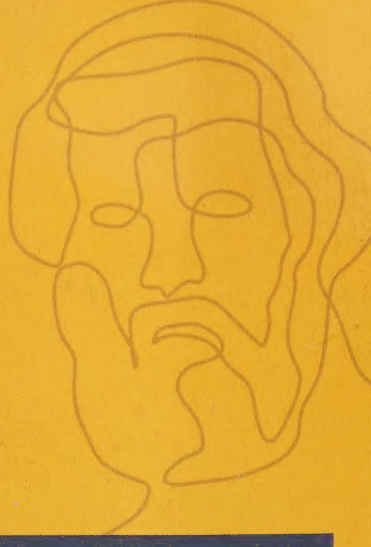


How to get an **A** in

A-LEVEL PHILOSOPHY

Deconstructing the **AQA syllabus** to achieve maximum marks with minimum effort



CHRISTIAN B. FEEST


Digitized by the Internet Archive
in 2022 with funding from
Kahle/Austin Foundation

https://archive.org/details/howtogetaninleve0000chri

How to get an **A** in

A-LEVEL PHILOSOPHY

Deconstructing the **AQA syllabus** to achieve maximum marks with minimum effort

CHRISTIAN B. FEEST

Copyright © 2021 by Christian B. Feest

All rights reserved. No part of this book may be reproduced or used in any manner without written permission of the copyright owner except for the brief use of quotations.

First paperback edition April 2021

ISBN: 9798719524658

CONTENTS

Introduction.....19

Basic philosophical reasoning.....21

Deductive reasoning.....21

Valid argument.....21

Invalid argument.....22

Inductive reasoning.....22

Epistemology.....25

What is knowledge?.....26

Overview.....26

The different kinds of knowledge.....26

Necessary and sufficient conditions.....27

The tripartite definition of knowledge.....27

Responses to the tripartite definition.....28

Gettier case 1.....28

Gettier case 2.....29

The JTB + no false lemmas definition of knowledge.....30

Responses to JTB+N.....31

Fake barn county.....31

Reliabilist definitions of knowledge.....33

Responses to reliabilism.....34

Fake barn county (again).....34

Virtue epistemology definitions of knowledge.....34

Linda Zagzebski: *What is Knowledge?*.....34

Formula for creating Gettier-style cases.....34

Zagzebski's definition of knowledge.....35

Sosa's virtue epistemology.....35

Responses to virtue epistemology.....36

Children and animals....36
The infallibilist definition of knowledge....36
Responses to infallibilism....37
Too strict....37
Perception as as a source of knowledge....38
Overview....38
Direct realism....39
Responses to direct realism....40
Time lag....40
Perceptual variation....41
Illusions....42
Hallucination....43
Indirect realism....44
Sense data....44
Locke's primary and secondary qualities....45
Responses to indirect realism....45
Mind-independent objects cannot be like sense data....45
Scepticism and the veil of perception....46
Indirect realist responses to the sceptical problem....47
Russell: External world is the best hypothesis....47
Locke: The involuntary nature of perception....48
Locke/Trotter Cockburn: Coherence of different senses....48
Idealism....49
George Berkeley: *Three Dialogues Between Hylas and Philonous*....50
Attack on Locke's primary/secondary quality distinction....50
The master argument....51
God as the cause of perceptions....52
Responses to idealism....53
Problems with the role of God in idealism....53
Solipsism....53
Illusion (again)....54
Hallucination (again)....54
Reason as a source of knowledge....56

Overview....56
A priori / a posteriori....56
Analytic / synthetic....57
Innatism vs. empiricism....57
Innatist arguments....58
Plato: Meno's slave....58
Leibniz: Necessary truths....59
Empiricist arguments....60
John Locke: *Essay Concerning Human Understanding*....60
Locke's arguments against innate knowledge....60
Innate knowledge would be universal....60
Argument against innate concepts....60
Tabula rasa....61
Simple, complex, and abstract general ideas....62
Rationalism (intuition and deduction) vs. empiricism....63
Rationalist arguments....64
Rene Descartes: *Meditations*....64
Three waves of doubt....64
Cogito argument for the existence of the self....65
Clear and distinct ideas....66
Trademark argument for the existence of God....66
Argument for the existence of the external world....67
Empiricist arguments....68
David Hume: *An Enquiry Concerning Human Understanding*....68
Hume's fork....68
Responses to rationalism....69
Response to Descartes' cogito argument....69
Responses to Descartes' arguments for the existence of God....69
Causation requires experience....69
The concept of God comes from experience....70
Responses to Descartes' argument for the existence of the external world....70
Relies on Descartes' argument that God exists....70
Causation (again)....70

The limits of knowledge **72**
Overview 72
Philosophical doubt vs. ordinary doubt 72
Scepticism 73
Descartes' evil demon 73
Responses to scepticism 75
Descartes' own response to scepticism 75
Russell's response to scepticism 75
Locke's response to scepticism 76
Berkeley's response to scepticism 76
Reliabilist responses to scepticism 77
Epistemology – module summary **79**

Moral philosophy **87**
Normative ethical theories **88**
Overview 88
Utilitarianism 89
Act utilitarianism 89
The felicific calculus 89
Responses to act utilitarianism 90
Difficult to calculate 90
The moral status of particular relationships 90
Tyranny of the majority 91
Ignores preferences (Nozick's experience machine) 92
Other utilitarian philosophies 92
John Stuart Mill: *Utilitarianism* 93
Mill's "proof" of utilitarianism 93
Higher and lower pleasures (qualitative hedonism) 93
Rule utilitarianism 94
Preference utilitarianism 95
Kant's deontological ethics 96
The good will 96
Duty 96

The categorical imperative....96
Test 1: Contradiction in conception....97
Test 2: Contradiction in will....97
The humanity formula....98
Responses to Kant's deontological ethics....99
Tweaking the maxim to justify any course of action....99
Ignores consequences....99
Ignores other valuable motivations....100
Conflicts between duties....101
Foot: *Morality as a system of hypothetical imperatives*....102
Aristotle's virtue ethics....103
Eudaimonia....104
Ergon, arête, and virtue....104
The doctrine of the mean....105
The skill analogy....106
Phronesis....107
Voluntary and involuntary action....107
Responses to Aristotle's virtue ethics....108
No clear guidance....108
Circularity....108
Competing virtues....109
Difference between moral good and eudaimonia....109
Metaethics....111
Overview....111
Moral realism....112
Moral naturalism....112
Utilitarian moral naturalism....113
Virtue ethics moral naturalism....113
Responses to moral naturalism....114
Moral non-naturalism....114
G.E. Moore: *Principia Ethica*....114
The naturalistic fallacy....114
The open question argument....115

Intuitionism....115
Responses to moral non-naturalism....116
Responses to moral realism....116
Moral anti-realism....116
Error theory....117
J.L. Mackie: *Ethics: Inventing Right and Wrong*....117
Arguments for cognitivism....117
Argument from relativity....117
Arguments from queerness....118
Responses to error theory....119
Non-cognitivism....120
Emotivism....120
A.J. Ayer: Verification principle....121
Hume: *Treatise of Human Nature*....122
Hume's fork (again)....122
Moral judgements motivate action....123
The is/ought problem....124
Prescriptivism....125
R.M. Hare: *The Language of Morals*....125
Responses to non-cognitivism....126
Moral argument and reasoning....126
Mackie's arguments against non-cognitivism....127
Responses to moral anti-realism....127
Moral nihilism....127
Moral progress....128
Applied ethics....129
Overview....129
Stealing....129
Utilitarianism and stealing....129
Kant's deontological ethics and stealing....130
Aristotle's virtue ethics and stealing....130
Metaethics and stealing....131
Simulated killing....131

Utilitarianism and simulated killing....132
Kant's deontological ethics and simulated killing....132
Aristotle's virtue ethics and simulated killing....132
Metaethics and simulated killing....133
Eating animals....133
Utilitarianism and eating animals....133
Kant's deontological ethics and eating animals....134
Aristotle's virtue ethics and eating animals....134
Metaethics and eating animals....135
Telling lies....135
Utilitarianism and telling lies....135
Kant's deontological ethics and telling lies....135
Aristotle's virtue ethics and telling lies....136
Metaethics and telling lies....137
Moral philosophy – module summary....138

Metaphysics of God....147
The concept of God....148
Overview....148
The divine attributes....148
Omnipotence....148
Omniscience....149
Omnibenevolence....149
Eternal/everlasting....149
Responses to the concept of God....151
The problem of the stone....151
The euthyphro dilemma....151
Omniscience vs. free will....153
Arguments concerning God's existence....155
Overview....155
Ontological arguments....156
Anselm's ontological argument....156
Descartes' ontological argument....156

Responses to (standard) ontological arguments....157
Gaunilo's island....157
Hume: "God does not exist" is not a contradiction....157
Kant: Existence is not a predicate....158
Malcolm's ontological argument (*necessary* existence)....159
Responses to Malcolm's ontological argument....160
The concept of God *is* self-contradictory....160
Fallacy of equivocation....160
Teleological arguments....160
Teleological arguments from spatial order....161
Paley's design argument....161
Hume's design argument....161
Responses to teleological arguments from spatial order....162
Hume's objections to teleological arguments....162
Problems with the analogy....162
Infinite time, finite matter....163
Causation....163
Spatial *dis*order....164
Darwin: Evolution by natural selection....164
Swinburne's teleological argument from *temporal* order....165
Responses to Swinburne's teleological argument....166
Multiple universes....166
Is the designer God?....167
Cosmological arguments....167
Cosmological arguments from *causation*....168
The Kalam argument....168
Aquinas' first way....168
Aquinas' second way....169
Descartes' cosmological argument....169
Responses to cosmological arguments from causation....170
Is a first cause necessary?....170
Hume: "Everything has a cause" is not an analytic truth....171
Is the first cause God?....171

Cosmological arguments from *contingency*....171
Aquinas' third way....171
Leibniz: Sufficient reason....172
Responses to cosmological arguments from contingency....173
Russell: Fallacy of composition....173
Hume: "God does not exist" is not a contradiction (again)....174
The problem of evil....175
The *logical* problem of evil....175
J.L. Mackie: The inconsistent triad....175
Responses to the logical problem of evil (theodicies)....176
God couldn't exist without evil....176
We need *some* evil to maximise goodness....177
We need evil for free will....178
Alvin Plantinga: *God, Freedom and Evil*....179
The free will defence....179
Natural evil as a form of moral evil....180
The *evidential* problem of evil....180
Responses to the evidential problem of evil (theodicies)....181
Free will (again)....181
John Hick: Soul making....181
Religious language....183
Overview....183
Cognitivism and non-cognitivism....183
Arguments that religious language is meaningless....184
A.J. Ayer: Verification principle (again)....184
Falsifiability....185
Anthony Flew: Invisible gardener....186
Arguments that religious language is meaningful....187
Hick: Eschatological verification....187
Basil Mitchell: Resistance fighter....188
R.M. Hare: Bliks....189
Metaphysics of God – module summary....191

Metaphysics of mind.....197

Dualist theories.....198

Overview.....198

Substance dualism.....199

Descartes' arguments for substance dualism.....199

The conceivability argument.....199

Responses to the conceivability argument.....200

Mind without body is not conceivable.....200

What is conceivable may not be possible.....201

What is possible tells us nothing about the actual world.....202

The divisibility argument.....203

Responses to the divisibility argument.....203

The mind is divisible.....203

Not everything that is physical is divisible.....204

Responses to substance dualism (in general).....205

The problem of other minds.....205

Causal interaction.....206

Version 1: The *conceptual* interaction problem.....206

Version 2: The *empirical* interaction problem.....207

Property dualism.....208

Qualia.....208

Supervenience.....208

Interactionism and epiphenomenalism.....209

David Chalmers: The zombie argument for property dualism.....209

Responses to the zombie argument.....210

Zombies are not conceivable.....210

Zombies are metaphysically impossible.....211

What is possible tells us nothing about the actual world.....213

Frank Jackson: The knowledge argument for property dualism.....213

Responses to the knowledge argument.....214

Mary would already know.....214

The ability response.....215

The acquaintance response.....216

New knowledge, old fact....217
Responses to property dualism (in general)....218
Qualia causes introspective self-knowledge....218
Evolution....219
The problem of other minds (again)....221
Physicalist theories....222
Overview....222
Behaviourism....223
Behavioural dispositions....223
Hard and soft behaviourism....224
Gilbert Ryle: *The Concept of Mind*....224
General arguments against dualism....224
Category mistake....225
Responses to behaviourism....226
Asymmetry between self- and other- knowledge....226
Super Spartans....227
Zombies (again)....228
The conceivability argument (again)....229
Multiple realisability....229
Circularity....231
Type identity theory....232
J.J.C. Smart: *Sensations and Brain Processes*....232
Ontological reduction / contingent identity....232
Ockham's razor (against dualism)....233
Responses to type identity theory....234
Location Problem....234
Zombies (again)....235
The conceivability argument (again)....235
The divisibility argument (again)....236
Multiple realisability....236
Functionalism....237
Hilary Putnam: *The Nature of Mental States*....237
Arguments against type identity theory....238

Arguments against behaviourism....238
Responses to functionalism....239
Inverted spectrum....239
Zombies (again)....240
The knowledge argument (again)....240
The China brain....241
Eliminative materialism....241
Paul Churchland: *Eliminative Materialism and the Propositional Attitudes*....242
Folk psychology....242
Folk psychology as a scientific theory....242
Folk psychology vs. neuroscience....244
Responses to eliminative materialism....245
Folk psychology has good predictive power....245
The direct certainty of folk psychology....245
Belief that eliminative materialism is true is self-refuting....246
Eliminative materialism can't explain intentionality....247
Metaphysics of mind – module summary....248

How to get an A/A* in the exam....253

Overview....254
AO1: 3, 5, and 12 mark questions....255
3 mark questions....255
Example 3/3 responses....255
5 mark questions....256
Example 5/5 responses....256
12 mark questions....257
Example 12/12 response....258
AO2: 25 mark questions....260
The 25 mark template....260
Introduction....261
A consistent argument....261
A balanced argument....262
Conclusion....263

General writing tips and examples....263
Level of detail....263
Precise use of philosophical language....264
Example 25 mark answer plans....266
Epistemology answer plans....267
The definition of knowledge....267
Direct realism....268
Indirect realism....269
Idealism....270
Innatism vs. empiricism....271
Rationalism vs. empiricism....272
The limits of knowledge....273
Moral philosophy answer plans....274
Utilitarianism....274
Kant's deontological ethics....275
Aristotle's virtue ethics....276
Metaethics....277
Metaphysics of God answer plans....278
The concept of God....278
Ontological arguments....279
Teleological arguments....280
Cosmological arguments....281
The problem of evil....282
Religious language....283
Metaphysics of mind answer plans....284
Substance dualism....284
Property dualism....285
Behaviourism....286
Type identity theory....287
Functionalism....288
Eliminative materialism....289
Exam technique summary....290

Practice questions....291

Epistemology....292

3 mark questions....292

5 mark questions....292

12 mark questions....293

25 mark questions....294

Moral philosophy....295

3 mark questions....295

5 mark questions....295

12 mark questions....296

25 mark questions....297

Metaphysics of God....298

3 mark questions....298

5 mark questions....298

12 mark questions....299

25 mark questions....300

Metaphysics of Mind....301

3 mark questions....301

5 mark questions....301

12 mark questions....302

25 mark questions....303

Revision exercises....305

Quizzes....306

Epistemology....306

Moral philosophy....310

Metaphysics of God....314

Metaphysics of mind....318

Tables....322

Epistemology....322

Moral philosophy....323

Metaphysics of God....324

Metaphysics of Mind....325

Last minute revision checklist....326

Syllabus....327

Epistemology....328
Moral philosophy....331
Metaphysics of God....334
Metaphysics of mind....337

Bibliography/reading list....341

Epistemology....342
Moral philosophy....342
Metaphysics of God....343
Metaphysics of Mind....344

Glossary....345



INTRODUCTION

I love philosophy – I really do. It's something worth doing for its own sake.

Not only that, it's genuinely useful. In a world where everyone outsources their thinking to group consensus and authority figures, being able to think for yourself is basically a superpower. Done right, philosophy teaches sound reasoning from first principles – and this can help you identify undervalued investments and get rich, discover new scientific truths outside of accepted paradigms, or even uncover the secrets of the universe itself. Philosophy can also help you live a more meaningful life.

But there is a difference between learning philosophy and *learning to pass the philosophy A-level exam*. The two things overlap quite a bit, but your final grade ultimately comes down to two 3-hour exam papers at the end of the course. These exam papers have a set formula and can only contain topics listed in the AQA syllabus. So, like the subtitle says, this book is about deconstructing this syllabus to achieve maximum marks. It's about getting you the best grade in your philosophy A-level exam – and that's it.

Some of the syllabus topics can be complicated, but I have tried to explain them as simply as possible and in a way that is practically useful for the exam. This usually means breaking down each topic or theory into three things:

- What the theory says
- The arguments *for* that theory (and responses)
- The arguments *against* that theory (and responses)

Some topics don't fit perfectly into the binary of arguments for vs. arguments against, but most do.

Knowing these arguments and responses is particularly important for the 25 mark questions. And, in my experience, it's these essay-style 25 mark questions that students struggle with most. Whereas the other question types (3, 5, and 12 marks) are basically just memory test questions, the 25 mark questions require you to construct an *argument* for a particular viewpoint.

To help you understand how to construct an argument that will get high marks, I have included a bunch of example 25 mark answer plans starting from page 266. I highly recommend that you come up with and *memorise* your own answer plans to these questions because doing so is much easier than trying to construct an argument on the fly in the limited time you have in the exam! Luckily, there are only around 20 or so 25 mark questions you can be asked (although they might be worded slightly differently), so it's not like you have to memorise thousands of essay plans. In my view, planning and memorising your answers to the 25 mark questions is the single most effective revision task.

Anyway, I really hope you enjoy this book, that you find the course content interesting, and that it makes you think. But also, I hope this book gets you that A or A* grade in the philosophy A-level!

BASIC PHILOSOPHICAL REASONING

We make arguments and give reasons for things all the time. For example, you might say something like "it can't have been him because he was in Spain that night" without even realising the underlying logic of that argument. Instinctively, we recognise the logic of people's claims and arguments. We can often tell good arguments from bad arguments – even if we don't always consciously realise why.

This is basically what philosophy is: it's about using logic and reason to argue that a certain claim or viewpoint is true or false. Often, this will involve bringing the subconscious understanding of logic before the conscious mind. The examples below illustrate some basic logical concepts and reasoning methods. As you go through the arguments in each module, be mindful of the underlying logic and reasoning. If possible, try to consciously identify *why* the argument works or doesn't.

DEDUCTIVE REASONING

A deductive argument is one where the premises are supposed to *logically guarantee* the conclusion. Another word for this is a logically *valid* argument.

Valid argument

A valid argument is one where if the premises are true the conclusion must be true as well. For example:

1. If Rex is a dog, then Rex is a mammal
2. Rex is a dog
3. Therefore, Rex is a mammal

Notice that the validity of this argument means it's impossible for the conclusion (3) to be false if the premises above (1 and 2) are true. In other words, it would be *illogical* to agree with 1 and 2 but disagree with 3.

However, just because an argument is valid, this doesn't guarantee its conclusion is true. Validity is just about the *structure* of the argument - not the *truth* of the premises. It's possible to have a valid argument with false premises and a false conclusion. For example:

1. If Rex is a dog, then the moon is made of green cheese
2. Rex is a dog
3. Therefore, the moon is made of green cheese

This argument is still valid – the conclusion *does* follow logically from the premises – but the premises and conclusion are clearly false. So, the argument is *valid*, but not *sound*. A sound argument is one that is logically valid *and* where the premises are true.

Invalid argument

An invalid argument is an argument where it's possible for the premises to be true but the conclusion be false. For example:

1. If Rex is a dog, then Rex is a mammal
2. Paris is the capital of France
3. Therefore, Rex is a mammal

This argument is obviously invalid – the conclusion (3) does not follow from the premises (1 and 2). Even if all the premises and the conclusion are true, the *structure* of this argument is not valid.

Sometimes invalid arguments aren't so obvious. For example, the following argument is also logically invalid:

1. If Rex is a dog, then Rex is a mammal
2. Rex is a mammal
3. Therefore, Rex is a dog

Again, even if the premises and the conclusion of this argument are all true, this argument is not valid. The reason it isn't valid is because it's *possible* for the premises to be true and the conclusion to be false. For example, it's possible that Rex is the name of my *cat* – in which case premises 1 and 2 are still true – but the conclusion "Rex is a dog" would be false.

INDUCTIVE REASONING

Whereas a deductive argument is supposed to logically guarantee its conclusion, an inductive argument is one where the premises only provide *support* for the conclusion. For example:

1. Most dogs have 4 legs
2. Rex is a dog
3. Therefore, Rex probably has 4 legs

Notice that the premises of this argument (1 and 2) provide support for the conclusion (3) but do not logically guarantee it. For example, Rex might have lost a leg somehow and

only have 3 legs. So, a good inductive argument shows that the conclusion is *probably* true, even though it doesn't show that the conclusion is *certainly* true.



EPISTEMOLOGY

Epistemology means theory of knowledge. The epistemology module covers what the definition of knowledge is, as well as how much knowledge comes from our perception of the world and how much of our knowledge comes from reason and thinking. It also covers the idea of scepticism: whether it is possible to know anything at all.

WHAT IS KNOWLEDGE?

OVERVIEW

"What is knowledge?" is one of the oldest questions of philosophy. The definition of knowledge as "justified true belief" can be traced all the way back to Plato and stood for thousands of years – until a 1963 philosophy paper by Edmund Gettier challenged this definition.

Gettier described two scenarios – now known as *Gettier cases* – where an individual has a justified true belief that is *not* knowledge. Gettier cases thus show that justified true belief is not a sufficient definition of knowledge.

Since Gettier's argument against the justified true belief definition, philosophers have proposed various alternative accounts of knowledge. The goal of these definitions is to define "knowledge" in a way that rules out Gettier cases whilst still capturing all instances of what we consider to be knowledge.

The A-level philosophy syllabus looks at five definitions of knowledge:

1. **Justified true belief (the tripartite definition)**
2. **JTB + No false lemmas**
3. **Reliabilism**
4. **Virtue epistemology**
5. **Infallibilism**

The different kinds of knowledge

First, we need to distinguish the *kind* of knowledge we're defining. Broadly, there are three kinds of knowledge: ability, acquaintance, and *propositional*.

Ability	Acquaintance	Propositional
Knowledge *how*	Knowledge *of*	Knowledge *that*
E.g. "I know how to ride a bike"	E.g. "I know Fred well"	E.g. "I know that London is the capital of England"

When we talk about the definition of knowledge in this chapter, we are talking about the definition of *propositional* knowledge – i.e. knowledge of facts – specifically.

Necessary and sufficient conditions

The name of the game in defining "knowledge" is to give a definition where each individual condition is *necessary* and where all the conditions taken together are jointly *sufficient*.

The classic example used to illustrate this sort of definition is the concept of a bachelor. The definition "unmarried man" provides both necessary and sufficient conditions that define "bachelor" because:

- "Unmarried" and a "man" are *necessary* conditions of what it means to be a bachelor. You can't be a bachelor if you fail to meet either of these conditions. If you are married, you cannot be a bachelor. If you are not a man, you cannot be a bachelor.
- Having just one of these conditions in isolation is not *sufficient*, though. For example, being unmarried doesn't guarantee a person is a bachelor – because that person may be an unmarried woman. However, being *both* "unmarried" *and* "man" is sufficient to be a bachelor because everything that is an unmarried man is a bachelor.

By providing necessary and sufficient conditions, "unmarried man" is a perfect definition of bachelor because everything that fits that definition will be a bachelor and everything that is a bachelor will fit that definition. So, the correct definition of "knowledge" would provide necessary and sufficient conditions in the same way that "unmarried man" provides necessary and sufficient conditions for "bachelor".

THE TRIPARTITE DEFINITION OF KNOWLEDGE

In *Theaetetus*, Plato gives a definition of knowledge as "true belief accompanied by a rational account". This seems to be the origin of what is known as the "tripartite" definition of knowledge, which got simplified to:

- Justified
- True
- Belief

On the face of it, the definition "justified true belief" seems to provide necessary and sufficient conditions for "knowledge" in the same way "unmarried man" provides necessary and sufficient conditions for "bachelor".

"True" is necessary because you can't know something if it is false. If someone said, "I know that the moon is made of green cheese" you wouldn't consider that knowledge because it isn't true.

"Belief" is also a necessary condition of knowledge because knowledge is taken as a relationship between a person and the world. It may be a *fact* that "Paris is the capital of France", but I can't *know* that fact unless I also believe it. It just wouldn't make sense, for example, to say "I know Paris is the capital of France but I don't believe that Paris is the capital of France."

And finally, justification. Suppose someone asks you if you know how many moons Pluto has. You have no interest in astronomy but just have a strong feeling about the number five because it's your lucky number or whatever. You'd be right – Pluto does indeed have five moons – but it seems a bit of a stretch to say you *knew* Pluto has five moons. Your true belief "Pluto has five moons" is not properly justified and so would not count as knowledge.

Together, these necessary conditions (justified, true, and belief) are said to be jointly sufficient. This means they capture every instance of knowledge whilst not including anything that isn't knowledge. However, this latter part is challenged by "Gettier cases".

Responses to the tripartite definition

Gettier case 1

In 1963, philosopher Edmund Gettier published a paper titled *Is Justified True Belief Knowledge?* This paper describes two scenarios where an individual has a justified true belief that does *not* count as knowledge. The first Gettier case is as follows:

- Smith and Jones are interviewing for the same job
- Smith hears the interviewer say "I'm going to give Jones the job"
- Smith also sees Jones count 10 coins from his pocket
- So Smith forms the belief that "*the man who will get the job has 10 coins in his pocket*"
- But Smith gets the job, not Jones
- And, by coincidence, Smith also has 10 coins in his pocket

Smith's belief that "the man who will get the job has 10 coins in his pocket" meets all three conditions of the tripartite definition of knowledge:

- Justified: Smith hears the interviewer say Jones will get the job and he sees that Jones has 10 coins in his pocket

- Belief: The justifications above lead Smith to form the belief that "the man who will get the job has 10 coins in his pocket"
- True: The man who gets the job (Smith) does indeed have 10 coins in his pocket

However – and this is the key point – despite being a justified true belief, it seems wrong to say that Smith's belief counts as knowledge because it's just luck that led to him being correct. Smith thought *Jones* would get the job, but that was wrong. It's sheer coincidence that Smith also happens to have 10 coins in his pocket. This shows that the tripartite definition of knowledge is not *sufficient*: you can have a justified true belief that is *not* knowledge.



Venn diagram to illustrate that not everything that is JTB is knowledge

Gettier case 2

Gettier's second example is similar to the first Gettier case in that it describes a justified belief that turns out to be true purely as a result of luck.

This example relies on the logical principle of *disjunction introduction*. Disjunction introduction is a logical law that says that if you have a true statement and add "or some other statement" then the full statement (i.e. "true statement *or some other statement*") is also true.

For example:

- "London is the capital of England" is true.
- And so, using disjunction introduction, the statement "either London is the capital of England *or the moon is made of green cheese*" is also true, because London is the capital of England.

- Even though the second part ("the moon is made of green cheese") is false, the overall statement is true because the "or" means only one part has to be true (in this case "London is the capital of England")

In short, disjunction introduction means that if "proposition 1" is true, then "proposition 1 *or proposition 2"* must also be true.

So, Gettier's second example is as follows:

- Smith has a justified belief that "Jones owns a Ford"
- Using the principle of disjunctive introduction above, Smith forms a further justified belief that "*either Jones owns a Ford or Brown is in Barcelona"*
- But it turns out that Jones *does not* own a Ford – he sold it last week
- However, by sheer coincidence, Brown is in Barcelona

So, Smith's belief that "either Jones owns a Ford or Brown is in Barcelona" is both true and justified.

It's *justified* because the original belief "Jones owns a Ford" is justified (let's say Smith and Jones are colleagues and Smith sees Jones pull up to work in a Ford every day). Disjunction introduction is a logically valid inference and so the second belief "either Jones owns a Ford or Brown is in Barcelona" is also justified.

It's also *true* because the entire statement "either Jones owns a Ford or Brown is in Barcelona" does turn out to be correct. However, Smith thought this statement was true because of the *first* condition (Jones owns a Ford) whereas it turns out it's actually true because of the *second* condition (Brown is in Barcelona).

So, despite being a justified true belief, it is wrong to say that Smith's belief counts as *knowledge*, because it was just luck that led to him being correct. This again shows that the tripartite definition of knowledge is not sufficient because it shows that you can have a justified true belief that is not knowledge.

Both Gettier cases are generally considered to be knockout blows for the tripartite definition of knowledge. And so, in response, philosophers have tried to come up with new definitions of knowledge that avoid Gettier cases. Generally, these new definitions seek to refine the justification condition of the tripartite definition.

THE JTB + NO FALSE LEMMAS DEFINITION OF KNOWLEDGE

The no false lemmas definition of knowledge (JTB+N) is a direct response to Gettier cases. It adds an extra condition to the tripartite definition that rules out Gettier cases.

The JTB+N definition of knowledge says that James has knowledge of *P* if:

- *P* is true
- James believes that *P*
- James's belief is justified
- James did not infer that *P* from a false lemma

A false lemma is basically a false belief. So, JTB+N just adds an extra condition to the tripartite definition and says that knowledge is a justified true belief *that is not inferred from anything false* (a false lemma).

Adding this extra condition avoids Gettier cases because:

- In Gettier case 1, Smith's belief "the man who will get the job has 10 coins in his pocket" is inferred from the *false lemma* that "Jones will get the job"
- In Gettier case 2, Smith's belief "either Jones owns a Ford or Brown is in Barcelona" is inferred from the *false lemma* that "Jones owns a Ford"

Again, Gettier cases are a problem for the tripartite definition of knowledge because they are examples of justified true beliefs that are not knowledge.

In other words, the tripartite definition *incorrectly* says Smith's belief *does* count as knowledge. But by adding the no false lemmas condition, JTB+N *correctly* say Smith's belief *does not* count as knowledge.

The story so far:

- The tripartite definition says Smith's belief is knowledge, even though it isn't. This shows the tripartite definition must be wrong.
- The JTB+N definition says Smith's belief is *not* knowledge, which is correct.
- So, in this instance, the JTB+N definition appears to be a more accurate analysis of knowledge than the tripartite view. It can correctly say that Gettier cases *do not* count as knowledge, whereas the ordinary JTB definition (the tripartite view) incorrectly says Gettier cases *do* count as knowledge.

Responses to JTB+N

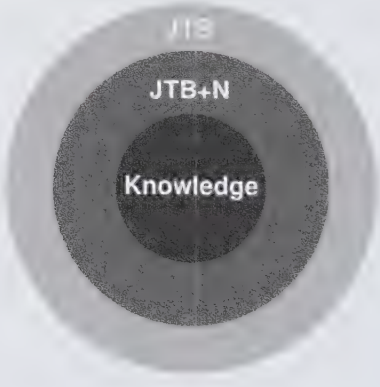
Fake barn county

However, there is another scenario similar to Gettier cases that shows the no false lemmas definition of knowledge is *also* not sufficient: fake barn county.

The fake barn county scenario is as follows:

- In fake barn county, the locals create fake barns that look identical to real barns
- Henry is driving through fake barn county, but he doesn't know the locals create these fake barns
- Henry often thinks "there's a barn" when he looks at the fake barns
    - These beliefs are not knowledge, because they are not *true* – the barns are fake
- However, on one occasion, Henry looks at the one *real* barn and thinks "there's a barn"
    - This time the belief is true
    - It's also justified by his visual perception of the barn
    - And it's not inferred from anything false (i.e. it's not inferred from a false lemma)

According to the JTB+N definition of knowledge, when Henry is looking at the only real barn, his belief "there's a barn" counts as knowledge. But this doesn't seem right: Henry's belief is only true as a result of luck – if Henry had been looking at any other barn his belief would have been false and he would have no idea. Henry's belief is not knowledge – he's just lucky in that one instance.



Venn diagram to illustrate that although JTB+N is a better definition of knowledge than JTB, it is still possible to have a JTB+N that is not knowledge

The fake barn county example shows that it is possible to have a JTB+N that is not knowledge. So, like the tripartite definition, the JTB+N definition of knowledge is also not a sufficient definition of knowledge.

RELIABILIST DEFINITIONS OF KNOWLEDGE

The reliabilist definition of knowledge gets rid of the "justification" condition and instead replaces it with the condition that the true belief must be formed via some (generally) reliable method.

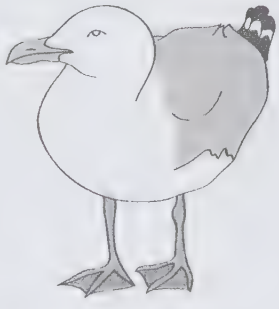
So, reliabilism says that James knows that *P* if:

- *P* is true
- James believes that *P*
- James's belief that *P* is caused by a reliable method

A reliable method is one that produces a high percentage of true beliefs.

So, if you have good eyesight, it's likely that your eyesight would constitute a reliable method of forming true beliefs. Likewise, if you have an accurate memory, it's likely your memory would also be a reliable method for forming true beliefs. If you form a true belief using these reliable methods, then that belief counts as knowledge.

But if you form a belief through an unreliable method – for example by simply guessing – then it would not count as knowledge even if the resultant belief is true. And obviously, if you form a false belief – even if the method used to form the belief is normally reliable – then that belief isn't knowledge either because it isn't true.



An advantage of reliabilism is that this definition allows that young children and animals can have knowledge. And typically, we do attribute knowledge to young children and animals. For example, it seems perfectly sensible to say that a seagull *knows* where to find food or that a baby *knows* when its mother is speaking. However, pretty much all the other definitions of knowledge considered here imply that animals and young children *can't* have knowledge. For example, a seagull or a baby can't *justify* its beliefs and so the tripartite and JTB+N definitions of knowledge seem to exclude them from having knowledge. Similarly, if virtue epistemology (page 34) is the correct definition, it's hard to see how a seagull or a newly born baby could possess intellectual virtues such as caring about forming true beliefs. So, virtue epistemology also seems to exclude animals and babies from possessing knowledge. But it seems obvious that young children and animals *do* possess knowledge, and so these definitions must be wrong.

Unlike the other definitions of knowledge, reliabilism can account for this: animals are capable of forming true beliefs via reliable processes, e.g. their eyesight, and so according to the reliabilist definition they are capable of possessing knowledge.

Responses to reliabilism

Fake barn county (again)

You can argue against reliabilism using the same fake barn county argument above (page 31): Henry's true belief that "there's a barn" is caused by a reliable cognitive process – his visual perception. Reliabilism would thus (incorrectly) say Henry's belief is knowledge, despite the fact that it's just lucky the belief happens to be true.

VIRTUE EPISTEMOLOGY DEFINITIONS OF KNOWLEDGE

There are several forms of virtue epistemology (we will look at two), but common to all of them is a link between the belief and *intellectual virtues*. Intellectual virtues are somewhat analogous to the sort of moral virtues considered in Aristotle's virtue ethics (page 103). However, instead of being concerned with *moral* good, intellectual virtues are about *epistemic* good. For example, an intellectually virtuous person would have traits such as being rational, caring about what's true, and a good memory.

Linda Zagzebski: *What is Knowledge?*

Formula for creating Gettier-style cases

Philosopher Linda Zagzebski argues that definitions of knowledge of the kind we have looked at so far (i.e. "true belief + *some third condition"*) will always fall victim to Gettier-style cases. She provides a formula for constructing such Gettier cases to defeat these definitions:

- Start with a situation where there is a belief that fits the definition ("true belief + *some third condition"*) but is false due to bad luck
    - E.g. Henry's belief "there's a barn" when he is looking at the fake barns
- Then change the situation to one where the belief fits that definition ("true belief + *some third condition"*) but is *true due to good luck*
    - E.g. Henry's belief "there's a barn" when he is looking at the one real barn
- In the second case, the belief will still fit the definition ("true belief + *some third condition"*) because it is basically the same as the first case
- But the second case won't be knowledge, because it's only true due to luck

Zagzebski argues that this formula will always provide a means to defeat any definition of knowledge that takes the form "true belief + some third condition" (whether that third

condition is justification, formed by a reliable process, or whatever). The reason for this is that truth and the third condition are simply *added* together, but not *linked*. The fact that truth and the third condition are not linked leaves a gap where lucky cases can incorrectly fit the definition.

Zagzebski's definition of knowledge

The issues resulting from the gap between truth and the third condition motivate Zagzebski to do away with the "truth" condition altogether. Instead, Zagzebski's analysis of knowledge is that James knows that P if:

- James believes that P
- James's belief that P arises from an act of intellectual virtue

However, in Zagzebski's analysis of knowledge, the "truth" of the belief is kind of implied by the idea of an act of intellectual virtue. This can be shown by drawing a comparison with *moral* virtue.

An act of *moral* virtue is one where the actor both *intends* to do good and *achieves* that goal. For example, intending to help an old lady across the road but killing her in the process is not an act of moral virtue because it doesn't achieve a virtuous goal (despite the virtuous intent). Likewise, helping the old lady across the road because you think she will give you money is not an act of moral virtue - even though it succeeds in achieving a virtuous goal - because your intentions aren't good.

Intellectual virtue is similar: you must both have the correct *motivation* (e.g. you want to find the truth) and succeed as a result of that virtue (i.e. your belief turns out to be true).

Virtues motivate us to pursue what is good. In the case of knowledge, good knowledge is also true. Secondly, virtues enable us to achieve our goals (in the same way a virtuous i.e. good knife enables you to cut) and so intellectual virtues would enable you to reliably form true beliefs.

Sosa's virtue epistemology

Another virtue epistemology definition of knowledge would be:

- P is true
- James believes that P
- James's true belief that P is a result of James exercising his intellectual virtues



Philosopher Ernest Sosa illustrates the meaning of this last condition with the following archery analogy: he says that a virtuous shot in archery has the following three properties:

- Accurate, i.e. it hits the target
- Adroit, i.e. the archer is skilful and shoots the arrow well
- Apt, i.e. the arrow hit the target *because* it was shot well

This last condition – aptness – differentiates Sosa's understanding of knowledge from those criticised by Zagzebski earlier (page 34). According to Sosa, for something to qualify as knowledge the belief must be true *as a direct consequence of* the believer exercising their intellectual virtues – it must be *apt*. Aptness provides a link between truth and the third condition that rules out Gettier-style situations where the belief is only true as a result of luck.

Returning to the fake barn county example, Sosa's virtue epistemology could (correctly) say Henry's belief "there's a barn" in fake barn county would *not* qualify as knowledge – despite being true and formed by a reliable method – because it is not apt. Yes, Henry's belief is accurate (i.e. true) and adroit (i.e. Henry has good eyesight etc.), but he only formed the true belief because of luck, not because he used his intellectual virtues.

Responses to virtue epistemology

Children and animals

As mentioned in more detail in the reliabilism section above (Page 33), a potential criticism of virtue epistemology is that it appears to rule out the possibility of young children or animals having knowledge. Babies and animals seem incapable of possessing intellectual virtues such as rationality or caring about the truth, and yet they seem capable of possessing knowledge of many things: a baby *knows* the sound of its mother's voice, a dog may *know* when it's time for dinner. This suggests that intellectual virtues are not necessary knowledge as virtue epistemology claims.

THE INFALLIBILIST DEFINITION OF KNOWLEDGE

Infallibilism argues that for a belief to count as knowledge, it must be justified in such a way that it is *certain*. In other words, for something to qualify as knowledge, the justification must be such that it cannot be doubted. This is the approach Descartes takes, for example, in his sceptical arguments (page 64).

Infallibilism is basically the tripartite definition of knowledge (page 27), but with a *much* stronger understanding of the "justification" condition. According to infallibilism, many of the things we would ordinarily consider as justifications are not. For example, you may

think your memory of various maps and internet searches and just general common sense justifies your belief that "Paris is the capital of France" – but infallibilism would disagree. You might be misremembering, or the map you looked at might be wrong, or you might be dreaming. You may even be a brain in a vat.

We saw earlier that the gap between "truth" and some third condition of knowledge (e.g. "justification") leaves room to construct Gettier cases. Infallibilism avoids this problem because it links justification and truth so closely that such a gap is impossible: certain justification *guarantees* truth.

Returning to the original Gettier case (see page 28), infallibilism can *correctly* say that Smith's belief "the man who will get the job has 10 coins in his pocket" does not count as knowledge because the justification does not guarantee certainty. Smith might have misheard the interviewer, for example, or Smith might even be a brain in a vat and Jones may not even exist! Either of these scenarios – however unlikely – raise the possibility of doubt and so Smith's belief would not count as knowledge according to infallibilism.

Responses to infallibilism

Too strict

Although infallibilism correctly says Gettier cases do not count as knowledge, infallibilism also says many other things do not count as knowledge. In fact, infallibilism says that pretty much *everything* fails to qualify as knowledge!

It seems perfectly correct to say I *know* many basic facts such as "I know that my name is Christian" and "I know that Paris is the capital of France". However, infallibilism would say I don't know any of these things because they can all be doubted. Pretty much any belief can be doubted, as Descartes demonstrates in his three waves of doubt (page 64), and so there are pretty much zero beliefs that count as knowledge!

We can use the example of "I know that Paris is the capital of France" above to illustrate why this is a problem for infallibilism:

- If infallibilism is true, then I do not know "Paris is the capital of France"
- But I do know "Paris is the capital of France"
- Therefore, infallibilism is false

These examples show that the infallibilist definition sets the bar for knowledge way too high. According to our ordinary understanding of "knowledge", certainty is not *necessary* for knowledge, and so infallibilism fails as a definition of knowledge.

PERCEPTION AS AS A SOURCE OF KNOWLEDGE

OVERVIEW

This epistemology topic looks at how we can gain knowledge from perception. We consider three theories of perception. They are:

1. **Direct Realism**
2. **Indirect Realism**
3. **Idealism**

The theories disagree over such issues as whether a mind-independent external world exists (realism vs. idealism) and also the way we perceive it (direct vs. indirect). Each theory also has various arguments for and against it. These key points are summarised below:

	Direct Realism	Indirect Realism	Idealism
External world	Mind-independent	Mind-independent	None: Objects are *mind-dependent*
Perception method	Direct	Indirect (via sense data)	Direct
Philosophers		Locke Russell	Berkeley
Problems	Illusion Hallucination Perceptual variation Time lag	Veil of perception (scepticism) Berkeley's attack on primary and secondary qualities	Solipsism Illusion Hallucination

DIRECT REALISM

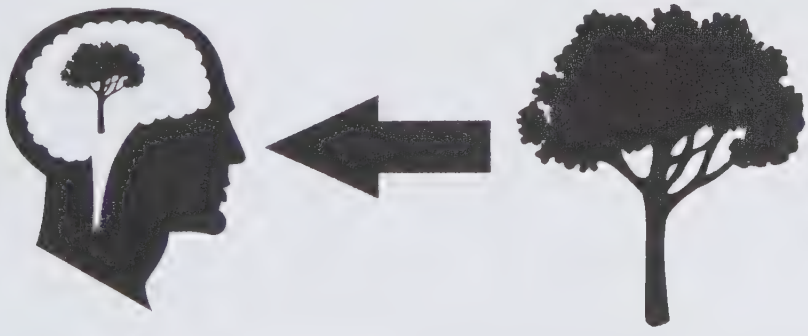
Realism is the view that a mind-independent external world exists. In other words, realists think that objects such as trees and tables and rocks and so on really do exist regardless of whether or not there are any humans around to perceive these objects.

So, *direct* realism is the view that:

- The external world exists independently of the mind (i.e. realism)
- And we perceive the external world *directly* (hence, direct)

Basically, direct realism says that what you perceive is the world itself. When you look at and perceive a tree, for example, you are directly perceiving a tree that exists "out there" in the world independently of your mind. You are also perceiving its properties (size, shape, smell, etc.).

So, direct realism says the *immediate objects of perception* are mind-independent objects and their properties.



Direct realism is often thought of as the common sense theory of perception. When asked what you see, you describe the external object *itself*, not your *perception* of it.

For example, when asked:

> *"What do you see?"*

You might respond:

> *"A table."*

You wouldn't respond with:

> *"Brown patches of sense data in a rectangular arrangement."*

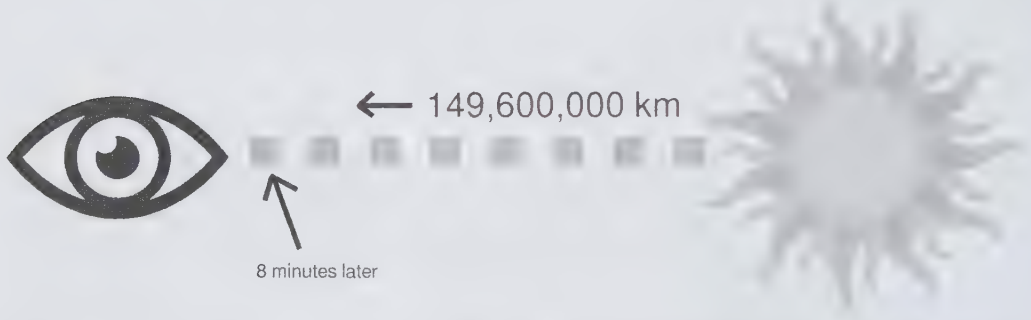
So, at least on the face of it, perceptual experience presents itself to us as mind-independent objects. However, there are several problems for this common sense account of perception.

Responses to direct realism

Time lag

Direct realism says that we perceive the external world directly as it is. But the time lag argument highlights an apparent difference between the external world and our perception of it.

Because it takes time for light and sound waves to travel from objects to us, we can argue that we do not perceive those objects *directly*. For example, the sun is 149,600,000 km from earth and light travels at 299,792,458 metres per second. This means it takes approximately 8 minutes for light to reach earth. So, when you look at the sun, you are seeing it as it *was* 8 minutes ago - i.e. you are not perceiving it *directly* as it is now.



Stars are even further away and so the star you perceive at night could be long gone and yet you would still perceiving it. So, we clearly are not perceiving these objects directly.

The argument is that because it takes time for light waves (and sound waves) to travel from objects to us, we never perceive those objects directly. Instead, what we perceive is light and sound waves that come from mind-independent objects but not mind-independent objects themselves, as direct realism claims.

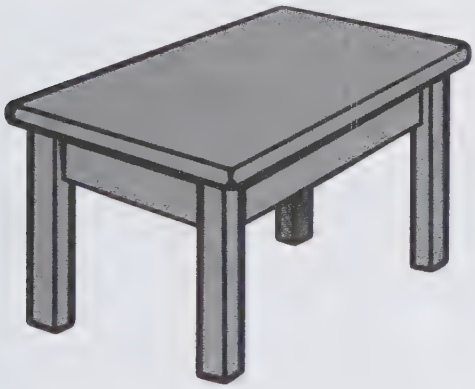
*Possible **response** to the time lag objection:*

The direct realist can respond that the time lag objection misunderstands the theory. Direct realism is a claim about *what* we perceive (mind-independent objects), whereas the time lag objection is an observation about *how* we perceive (via light waves, sound waves, etc.). Direct realism can still claim that when we perceive an object, we perceive it directly - there's nothing in the middle, such as sense data (page 44) - but it is a direct perception of the world as it *was* moments ago.

Perceptual variation

Philosopher Bertrand Russell (1872-1970) provides examples of perceptual variation. These perceptual variations present a further challenge to direct realism because they are examples of differences between reality and our perception of it.

For example, from directly overhead, a table may appear to be *rectangular* in shape. But if you're looking at the table from a different angle, it may look *kite-shaped*. But the table can't really be both shapes - so at least one of these perceptions must be different from the table itself.





Russell gives another example: when I stand on one side of the room, a shiny wooden table may appear to have a white spot where the light is reflecting off it. But we know that the table doesn't *really* have a white spot on it - it's just a trick of the light. So again, this example suggests a distinction between reality (no white spot) and our perception of it (white spot).

Examples of perceptual variation like this highlight differences between mind-independent objects and our perceptions of them. But according to direct realism, we perceive the external world *directly* so there should be no such differences between perception and reality.

So, in short, the perceptual variation argument against direct realism is something like this:

- Direct realism says what we perceive is the object itself
- Perceptions change (e.g. depending on where you stand)
- The object itself does not change
- So, if our perception changes but the object does not, then what we are perceiving is not the object itself
- So, direct realism is false

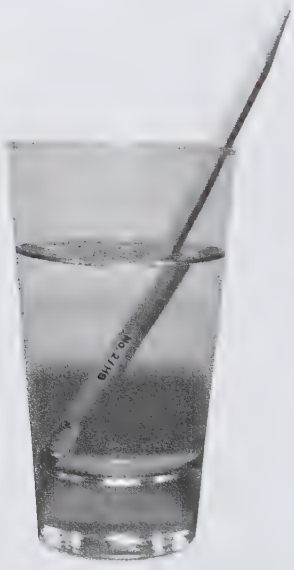
*Possible **response** to the perceptual variation objection:*

Direct realism can respond by refining the theory and introducing the idea of *relational* properties.

A relational property is one that varies in relation to something else. For example, being to the left or right of something (e.g. "the cupboard to the left of the fridge") is a real property that something can have – but it varies relative to other objects. Similarly, we could say that certain properties vary according to the perceiver. *Appearing* kite-shaped is a real property a table can have relative to certain perceivers. To say the table *appears* kite-shaped is not the same as saying the table *is* kite-shaped. In other words, the object *itself* does not change, but the *perceiver* does – and thus the relational properties between the two change.

Illusions

Again, direct realism says that we perceive the external world directly as it is. But illusions are situations where reality (i.e. the external world) differs from our perception of it. For example, when a pencil is placed in a glass of water, we perceive it as crooked. But it isn't *really* crooked.



If direct realism is true, the external world would be exactly as we perceive it. However, with illusions, there is an obvious difference between reality and our perception of it. This suggests that we are not perceiving reality directly and, hence, that direct realism is false.

*Possible **response** to the illusion objection:*

Just as in response to perceptual variation, the direct realist could reply that the pencil has the *relational* property of *looking* crooked to certain perceivers even though it isn't really crooked. However, this response is less convincing against the argument from hallucination below.

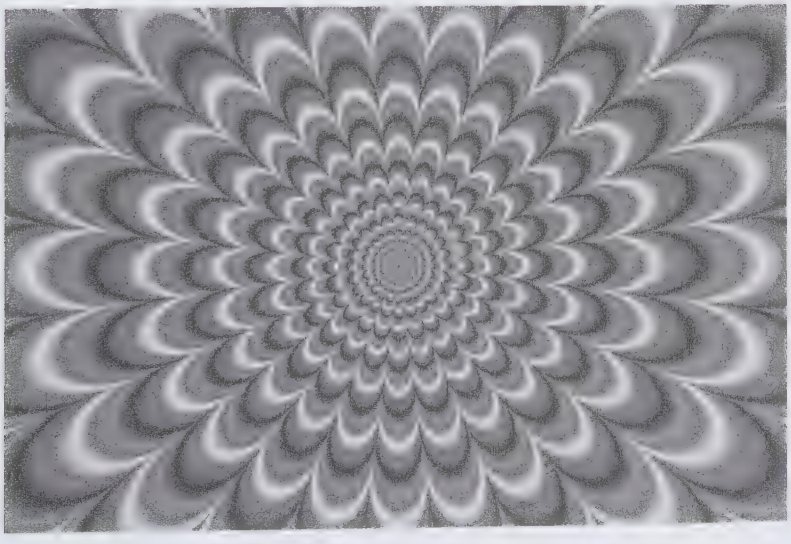
Hallucination

Hallucinations represent an even greater challenge to direct realism than illusions.

During hallucinations – perhaps as a result of being ill or taking drugs – we perceive things that seemingly have no basis in reality whatsoever. For example, someone might take a hallucinogenic drug such as LSD and then perceive elves or gremlins or other such creatures that do not exist in reality.

The direct realist's reply to the illusion objection above does not work for hallucinations: it's one thing to say that a pencil can have the relational property of *looking* crooked – it's another thing entirely to say, for example, that the table has the property of looking like it has gremlins on it! Hallucinations seemingly have no basis in reality whereas illusions, such as the crooked pencil, at least have *some* basis in reality. Whatever is causing the gremlin hallucination can't be the external world – at least not directly – because there is no correspondence between the external world and the perception whatsoever.

Again, direct realism says that we perceive the external world directly as it is. But if these elves or gremlins do not exist in the external world, then we cannot be perceiving the external world directly. So, hallucinations show that direct realism must be false.



INDIRECT REALISM

Indirect realism agrees with direct realism that a mind-independent external world exists – so it's another *realist* theory.

But whereas direct realism says we perceive this external world *directly* as it is, indirect realism says that what we perceive is *sense data.* This sense data is said to be caused by and represent the mind-independent external world. So, we do perceive the mind-independent external world, but we do so *indirectly* via sense data.

In other words, indirect realism is the view that:

- The external world exists independently of the mind (hence, realism)
- But we perceive the external world *indirectly*, via sense data (hence, indirect)



Most philosophers who argue for indirect realism start with the problems for direct realism. The many differences between reality and our perception of it, they say, make direct realism implausible. So, the arguments against direct realism (page 40) can be used as arguments for indirect realism.

Sense data

Sense data is often described as the content of perceptual experience. It's not a physical thing, it exists in the mind and is private to the individual perceiving it.

However, despite existing in the mind, sense data is said to be caused by and to represent mind-independent physical objects. This enables indirect realism to posit the existence of mind-independent objects (i.e. realism) while avoiding the problems with direct realism described above. Differences in perceptual variation (see page 41), for example, can be explained by differences in sense data: the reason I see a white spot on the table from one angle but not another, for example, is because the sense data differs depending on the perceiver. The object itself stays the same, though, even though the sense data changes.

Locke's primary and secondary qualities

Philosopher John Locke's (1632-1704) distinction between primary and secondary qualities is a kind of early version of indirect realism. Primary qualities are analogous to the mind-independent external world, whereas secondary qualities are analogous to sense data:

Primary qualities	Secondary qualities
Properties inherent in the object itself	Powers of an object to cause sensations in humans
Objective	Subjective
• Size	• Colour
• Shape	• Taste
• Motion	• Smell
• Number	• Feel

Locke uses various examples to illustrate the distinction between primary and secondary qualities. One such example is porphyry: a red and white stone. Locke says that when you prevent light from reaching porphyry, "its colours vanish". However, the primary qualities – size, shape, etc. – remain.

The distinction between primary and secondary qualities can be used to support indirect realism. Like sense data, this distinction explains the difference between reality (primary qualities) and our perception of it (secondary qualities).

*Possible **response** to the primary and secondary quality distinction:*

See Berkeley's criticism of the distinction between primary and secondary qualities (page 50).

Responses to indirect realism

Mind-independent objects cannot be like sense data

Again, indirect realists say that we perceive (mind-dependent) sense data and that this sense data resembles mind-independent objects. But philosopher George Berkeley questions how it is possible for one kind of thing (sense data) to resemble a completely different kind of thing (mind-independent objects).

Firstly, sense data is constantly changing, but mind-independent objects are not. For example, we saw in the argument from perceptual variation (page 41) that the sense data of a table might look square from above but kite-shaped from a different angle. Despite

this, indirect realism would say that the table *itself* doesn't change – only the sense data. But if this is the case, then how can the sense data be anything like the object itself? One type of thing (sense data) is constantly changing, whereas the other kind of thing (the mind-independent object) is stable and unchanging. That's a big difference!

Further, if we only ever experience (mind-dependent) sense data, how can we say this sense data *resembles* mind-independent objects? For example, let's say I am looking at a table and perceiving it as square – i.e. that my perception is square. But how can the squareness *of my perception* be somehow similar to the apparent squareness *of the table itself?* Again, a perception is a completely different type of thing to a mind-independent object.

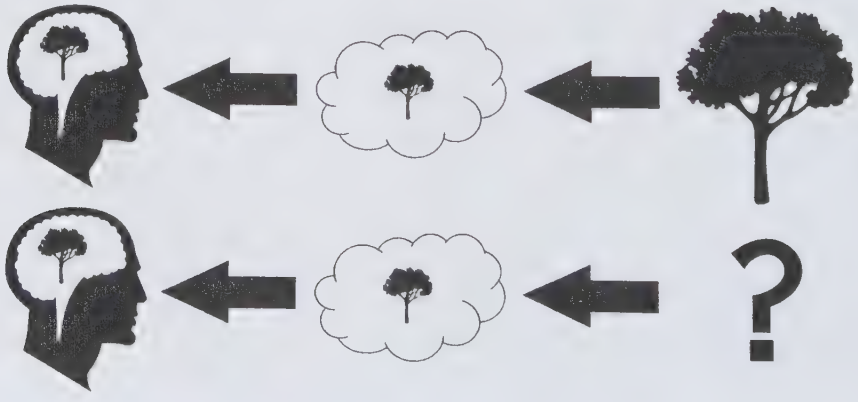
*Possible **response** to mind-independent objects cannot resemble sense data:*

Berkeley's argument above was originally targeted at *Locke's* indirect realism. It's Locke who says sense data *resembles* mind-independent objects, but more modern indirect realism says sense data *represents* mind-independent objects.

An analogy for this difference would be something like: a map *represents* a geographical area, but it doesn't *resemble* it. The geographical area is 3D and consists of grass and rocks and water and so on, whereas the map is a 2D piece of paper with some symbols on it. The 2D map and the 3D area are two completely different kinds of thing so they don't *resemble* each other. However, the map nevertheless *represents* the 3D area. Likewise, indirect realism can say sense data *represents* mind-independent objects even though sense data and mind-independent objects are two completely different kinds of thing.

Scepticism and the veil of perception

Look at the two diagrams below. What would be the difference *to the perceiver* between situations 1 and 2?



The answer, surely, is nothing! From the perceiver's perspective, it doesn't make any difference if there is an external object causing the sense data or not – their perception will be exactly the same in both cases.

This presents a problem for indirect realism: Indirect realism says that perception works like situation 1: the tree causes sense data that represents the tree. But if we only perceive the sense data, *how could we know there really is a mind-independent tree causing this sense data?* Reality could be like situation 2 and there would be absolutely no difference from the perspective of the perceiver. If we only perceive sense data and not the object itself, then there is no way of telling if the sense data is an accurate representation of the external world – or even that there is an external world at all! There exists a *veil of perception* between the object and the sense data that we cannot get past.

The argument, then, is this: indirect realism claims that there is a mind-independent external world, but cannot possibly provide any evidence for this claim because all we ever perceive is sense data.

Indirect realist responses to the sceptical problem

Several indirect realist philosophers have responded to the problem of the veil of perception above. The syllabus lists three responses to the veil of perception (and we also cover these responses in the scepticism discussion starting from page 73).

Russell: External world is the best hypothesis

Bertrand Russell concedes that there is no way we can conclusively get past the veil of perception and prove the existence of the mind-independent external world. However, Russell argues that the existence of a mind-independent external world is the *best hypothesis* (best explanation) for our perceptions.

Russell's argument is as follows:

- Either:
  - A: Mind-independent objects exist and cause my sense data
  - B: Mind-independent objects do not exist and do not cause my sense data
- I can't *prove* A or B definitively
- So, I have to treat A and B as hypotheses
- A is a better explanation of my experience than B
- So, mind independent objects exist and cause my sense data

Russell uses an example of a cat to explain why he thinks hypothesis A is a better explanation for his perceptions than hypothesis B. Suppose I see a cat sitting on a sofa and then later I see it sitting on the floor, the two explanations are:

- A: The cat exists independently of me and moved from one position to the other (passing through a range of intermediate positions) when I wasn't looking
- B: The cat stopped existing when I looked away and sprang into existence in a new position later on when I look again

Hypothesis B isn't really an explanation at all – it would mean it's basically just random that I have a perception of a cat in one place and then later in another. Hypothesis A, though, *connects* the two perceptions and explains why they happen the way they do.

Similarly, hypothesis A also explains other things like why the cat gets hungry between one meal and the next: the cat really exists independently of my perceptions and undergoes physical changes when I'm not looking.

Russell concludes that hypothesis A provides a better explanation of such perceptions than hypothesis B: it just makes more sense to posit mind-independent objects that undergo physical changes when you're not looking than it does to say that perceptions disappear and then reappear in different states for no reason.

Locke: The involuntary nature of perception

John Locke provides two responses to the veil of perception. Locke's first argument for the existence of a mind-independent external world is that perception is *involuntary*.

He first describes how he can close his eyes and *imagine* various sensations – such as the light of the sun, the smell of a rose, or the taste of sugar – at will. In contrast, when he opens his eyes and actually looks at the sun, *he cannot avoid* having the perception of the sun produced in him.

So, whereas imagination and memory are produced by the mind itself, perception must be produced by something *external* to the mind. Hence, Locke concludes that external objects exist and are what cause perception.

Locke/Trotter Cockburn: Coherence of different senses

Locke's second response to the veil of perception is the coherence of the different senses. He describes how, if you *see* a fire and doubt whether it is real or not, you can put your hand to the fire and also *feel* that it exists. The heat and pain produced by feeling the fire confirms your visual perception: you can confirm the existence of what you see with your eyes by feeling it with your hands.



Locke then gives another example to illustrate this point: writing words on paper. It's not possible to make words appear on a piece of paper by simply imagining it – you have to actually write them. And then, once you have written the words on the paper, it's not

then possible to see a blank piece of paper or different words to those you have written. You can even pass the paper to someone else to read what is on the page back to you, which confirms your visual perception of the words with what you hear. All this suggests that the words really do exist on the page independently of my mind.

Catherine Trotter Cockburn (1679-1749) made a similar argument to Locke. She describes how different kinds of perception - e.g. sights and sounds - occur together. For example, we associate visual perceptions of dogs with the sounds they make (e.g. barking). These associations enable us to predict one experience from another. For example, if you *see* a dog, you can predict what sort of *noise* it makes. Likewise, if you *hear* barking, you can predict what the source of that noise will *look* like. Hearing and seeing are two completely different *kinds* of experience, and yet they almost always occur together like this. Trotter Cockburn argues this is because the different senses are perceiving the same, mind-independent, object.

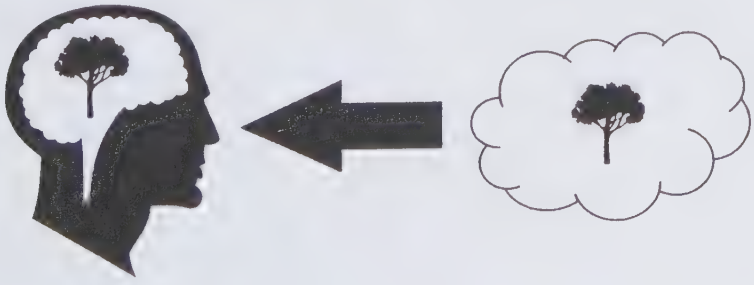
IDEALISM

Idealism differs from the other two theories we have looked at so far in that it rejects the existence of a mind-independent external world. In other words, idealism rejects *realism* (both direct and indirect).

According to idealism, unless an object is being perceived by a mind, it doesn't exist. In other words, there is no such thing as mind-independent objects. Instead, idealism claims that what we perceive are *mind-dependent ideas*.

So, idealism claims that:

- There is no external world independent of minds
- Instead, what we perceive are ideas (which is what reality is)



George Berkeley: *Three Dialogues Between Hylas and Philonous*

Bishop George Berkeley (1685-1753) is the most famous proponent of idealism. His idealism is captured by the phrase *"esse est percipi"* which translates as "to be is to be perceived". This phrase means that for something to exist, it must be perceived. This also means that, if something is not being perceived, then it does not exist.

Berkeley offers various arguments against the existence of a mind-independent external world. These arguments include a variation of the sceptical argument against indirect realism described earlier (page 46). Berkeley questions how – if realism is true – we can link up our perception of objects with the objects that supposedly cause our perceptions. Berkeley concludes that we can't ever get past the veil of perception to prove the existence of anything behind it.

The syllabus lists two of Berkeley's arguments against the existence of a mind-independent external world: his "master" argument, and his attack on Locke's distinction between primary and secondary qualities.

Attack on Locke's primary/secondary quality distinction

Earlier, we looked at Locke's distinction between primary and secondary qualities (page 45). Recall that Locke said primary qualities are mind-independent, whereas secondary qualities are not:

> *"Take away the sensation [...] and all colours, tastes, odours, and sounds [...] vanish and cease, and are reduced to their causes"*

Based on this extract, Locke seems to be saying that secondary qualities are mind-*dependent*. Berkeley agrees with Locke on this, and provides his own arguments to further support the claim that secondary qualities are mind-dependent.

But Berkeley goes further than Locke and argues that *primary* qualities are also mind-dependent. Berkeley's arguments for this are very similar to the perceptual variation examples used to argue against direct realism earlier (page 41). He describes various examples where primary qualities – such as size, shape, and motion – differ depending on the perceiver:

- Something that looks small to me may seem large to a small animal
- Something that looks small at a distance may seem large close up
- A smooth surface may look jagged under a microscope
- An object that appears to be moving quickly to humans may appear to be moving slowly to a fly

Size, shape, and motion are all supposed to be primary qualities, but these examples show how our perception of them differs depending on the circumstances.

Given these variations, Berkeley argues that we can't say these objects have one single size, shape, or motion independent of perception. This shows that primary qualities are just as mind-dependent as secondary qualities are.

Berkeley also argues that everything we perceive is either a primary quality or a secondary quality. Bringing this all together, we get the following argument:

- When we perceive an object, we don't perceive anything in addition to its primary and secondary qualities
- So, everything we perceive is either a primary quality or a secondary quality
- Secondary qualities are mind-dependent
- Primary qualities are also mind-dependent
- Therefore, everything we perceive is mind-dependent

So, Berkeley uses Locke's primary and secondary quality distinction to prove that everything we perceive is mind-dependent. If everything we perceive is mind-dependent, this implies that there is no such thing as a mind-independent external world (and so realism is false).

The master argument

The second of Berkeley's arguments for idealism is known as the master argument.The dialogue (between Hylas and Philonous) for the master argument can be summarised as:

> *Philonous: Try to think of an object that exists independently of being perceived.*
>
> *Hylas: OK, I am thinking of a tree that is not being perceived by anyone.*
>
> *Philonous: But that's impossible - you're perceiving it right now! You might be imagining a tree in a solitary place with no one perceiving it but you're still thinking about the tree. You can think of the idea of a tree, but not of a tree that exists independently of the mind.*

So, Berkeley's master argument is essentially that the very idea of mind-independent objects (i.e. realism) is inconceivable. His reasoning is that whenever we try to think of an unperceived and mind-independent object we are perceiving it and so it is not mind-independent. In other words, as soon as we think of a mind-independent object, it becomes mind-dependent. So, the idea of a mind-independent object is inconceivable.

*Possible **response** to the master argument:*

We can respond that Berkeley's master argument conflates the *idea* of an object with the object itself.

Of course it's impossible *to think* of something that's mind-independent – because thought itself is mind-dependent! But just because we can't think of something that is mind-independent, it doesn't follow from this that a mind-independent object is *itself* impossible. So, even if Berkeley has shown that the idea of an unperceived tree is impossible, it doesn't follow that an unperceived tree is itself impossible.



Just because we can't have an idea of a mind-independent tree, it doesn't mean a mind-independent tree is itself impossible

Clearly, all *ideas* are mind-*dependent* and so any idea about an object will also be mind-dependent. But just because *ideas* about objects are mind-dependent, it doesn't follow that *the object the idea is about* is also mind-dependent.

God as the cause of perceptions

Despite arguing against the existence of a mind-independent external world, idealism does not lead to the same veil of perception problem that indirect realism does (see page 46). The veil of perception disappears when we realise that words like "physical object" refer to *bundles of ideas* and not mind-independent objects (as realism assumes). Ideas *are* reality according to idealism, and so it doesn't make sense for there to be a veil of perception between these ideas and reality.

But this still leaves the question of what *causes* these ideas. Berkeley's answer is God, and his argument for this is as follows:

- Everything we perceive is mind-dependent (as argued above)
- There are 3 possible causes of these perceptions:
  - Ideas
  - My own mind
  - Another mind

- It can't be ideas, because ideas by themselves don't cause anything
- It can't be my own mind, because if I was the cause of my own perception then I'd be able to control what I perceived
- Therefore, the cause of my perception must be another mind
- Given the complexity, variety, order, and manner of my perceptions, this other mind must be God

Positing God as the cause of perceptions also resolves a potential objection to idealism: If, as Berkeley says, "to be is to be perceived" then this means that objects don't exist when we are not perceiving them. But this doesn't seem to be true. For example, if I leave an apple in a closed drawer for a few months, when I come back it will be withered and have changed form. How can this be possible unless it existed in the drawer when I wasn't looking? Berkeley's response is that the apple (and all "physical objects") constantly exist in the mind of God.

Responses to idealism

Problems with the role of God in idealism

Berkeley's conception of God seems to contradict his claim that what we perceive are ideas in God's mind.

For example, God is said to be eternal and unchanging, but my perceptions constantly change and go in and out of existence. So how can these *changing* perceptions exist in God's mind if God himself is *unchanging*?

Further, God is said not to feel sensations such as pain, but my perceptions often include such sensations. So, how can my perception of pain be an idea in God's mind if God himself cannot feel pain?

*Berkeley's **response** to this objection:*

Berkeley considers these objections and responds by clarifying his theory. What we perceive, he says, are *copies* of ideas that exist in God's *understanding*. Berkeley says that although God does not perceive sensations such as pain or change Himself, He *understands* what it is for us to undergo these experiences. Our changing perceptions are what God actively wills us to perceive.

Solipsism

Solipsism is the view that one's mind is the only thing that exists. There are no other minds and no mind-independent objects.

We saw earlier that Berkeley argues that mind-independent objects do not exist and that everything we perceive is mind-dependent (page 50). These arguments could also support the view that *other minds* do not exist either. After all, perception gives no reason to believe anything exists beyond our own experience – and this includes the existence of other minds. If we only ever perceive ideas, as Berkeley claims, we have no reason to believe in mind-independent objects or other minds. Thus, idealism leads to solipsism.

*Berkeley's **response** to this objection:*

Berkeley does provide an argument for at least one other mind: God's mind (page 52). Further, we might argue, as Mill does (page 205), that we can *infer* the existence of other minds by analogy to our own experience.

Illusion (again)

Idealism is a direct theory of perception: Berkeley says that we perceive ideas *directly* and that these ideas are what reality is. Thus, idealism makes no distinction between appearance (perception) and reality.

But if there is no distinction between perception and reality, how can Berkeley explain the argument from illusion from earlier (page 42)? For example, a pencil in water may *look* crooked, but we don't think the pencil is crooked *in reality*. Idealism has to say the pencil really is crooked – but this seems obviously false. So, idealism must be wrong.

*Berkeley's **response** to this objection:*

Berkeley argues that, yes, there is some sense in which the pencil *really is* crooked. After all, there are no such things as mind-independent objects, only ideas. If we are perceiving the idea of a crooked pencil, that idea is just as real as the idea of a non-crooked pencil when you take it out of the water.

Berkeley does say, though, that sometimes such perceptions can be misleading. For example, just because the pencil looks crooked when it is in water, it would be a mistake to assume that the pencil would also *feel* crooked, or that it would look crooked in other conditions.

Hallucination (again)

Hallucinations present a similar problem for idealism as illusions.

Again, idealism makes no distinction between our perceptions and reality – Berkeley says that "to be is to be perceived". But if there is no distinction between appearance and reality, this implies that hallucinations are just as real as ordinary perceptions are. This seems obviously false though – hallucinations are clearly not real, whereas ordinary perceptions are. So, again, idealism must be wrong.

Further, if God is the cause of our perceptions as Berkeley claims (page 52), why would God cause us to perceive hallucinations?

*Berkeley's **response** to this objection:*

Hallucinations and dreams, argues Berkeley, are products of *imagination* rather than perception. Even though imagination is normally voluntary, there are cases – i.e. hallucinations and dreams – where imagination is involuntary. Whereas God is the cause of ordinary perception, in hallucinations and dreams, we are the cause of our perceptions (albeit involuntarily).

Berkeley also provides two ways in which we can tell hallucinations and dreams from ordinary perceptions. Firstly, hallucinations and dreams are "dim, irregular, and confused" whereas ordinary perceptions are "more vivid and clear". Secondly, hallucinations and dreams are not connected with our ordinary lives and experience.

REASON AS A SOURCE OF KNOWLEDGE

OVERVIEW

This topic examines whether all our knowledge comes from perception or whether there are other sources of knowledge. There are really two separate debates in this topic, and these are:

- Innatism vs. empiricism
- Rationalism vs. empiricism

Broadly speaking, **empiricism** says that all knowledge comes from experience. And so, innatism and rationalism are two ways we can reject this claim:

- **Innatism** says that we are born with some knowledge already (and innate knowledge obviously doesn't come from perceptual experience because you haven't had any perceptual experience yet when you are born).
- **Rationalism** says that we can acquire some knowledge purely through intuition and deduction (i.e. we can acquire knowledge purely by thinking rather than through perceptual experience).

A lot of the discussion (in particular rationalism) draws on the following terms:

A priori / a posteriori

A priori and a posteriori are two different kinds of *knowledge:*

A priori knowledge	A posteriori knowledge
Knowledge that can be acquired without experience of the external world, *through thought alone*	Knowledge that can only be acquired from *experience* of the external world
E.g. working out what 900 divided by 7 is	E.g. doing an experiment to discover the temperature at which water boils

This reason as a source of knowledge section is mostly concerned with a priori knowledge, whereas the perception as a source of knowledge (page 38) was concerned with a posteriori knowledge.

In general, rationalists and innatists are fans of a priori knowledge. Empiricists, in contrast, don't rate a priori knowledge and argue that many claims of a priori knowledge are actually a posteriori instead.

Analytic / synthetic

Analytic and synthetic are two different kinds of *truths:*

Analytic truth	Synthetic truth
True in virtue of the *meaning of the words*	True in virtue of *how the world is*
Examples: • "A bachelor is an unmarried man" • "Triangles have 3 sides" • "1+1=2"	Examples: • "Grass is green" • "Water boils at 100°c" • "Paris is the capital of France"

Analytic truths cannot be denied without resulting in a logical contradiction. To say, "not all bachelors are unmarried", for example, is to misunderstand the word bachelor – the concept of a married bachelor does not make sense. Similarly, one cannot coherently imagine a triangle with 4 sides because the very idea involves a contradiction.

Denial of a synthetic truth *does not* lead to a logical contradiction, though. For example, we can coherently imagine red grass, which is a denial of the synthetic truth "grass is green". Though experience tells us grass is not actually red in reality, there is no *logical* contradiction in the idea of red grass.

INNATISM VS. EMPIRICISM

Innate knowledge is knowledge you're born with and so doesn't require experience (a posteriori) to be known. In other words, innate knowledge is a priori knowledge.

The debate between rationalism and innatism is about whether we have innate knowledge or not:

- Innatism says we have some innate knowledge
- Empiricism says we do not have any innate knowledge

It's also important to note that the kind of knowledge we are considering in this innate knowledge discussion is innate *propositional* knowledge (see page 26). It's uncontroversial, for example, that babies are born with innate *ability* knowledge such as

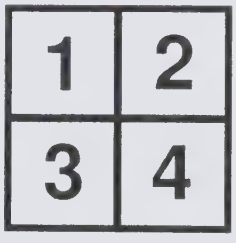
knowing how to breathe - both innatism and empiricism would accept this. So, the debate is about whether humans are born with any innate knowledge of facts.

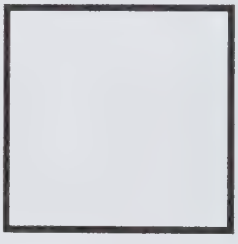
Innatist arguments

Plato: Meno's slave

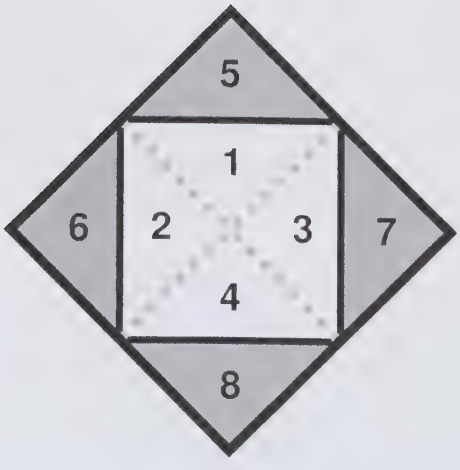
Plato (~420 BC-347 BC) argues that all learning is a form of recalling knowledge from before we're born. So, in other words, we're born with innate knowledge but we just need to remember it.

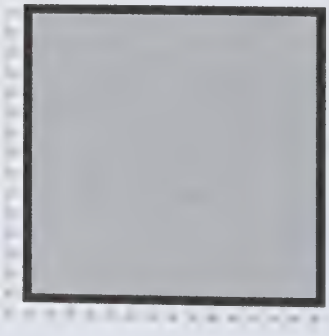
To prove his theory, Plato shows how Meno's slave - a boy who has never been taught geometry - *knows* a geometry proof that a square with an area of 2A will have sides equal to the diagonal of a square with an area of 1A.





1A
(4 square feet)





2A
(8 square feet)

The key facts of the argument are as follows:

- Socrates draws a square on the ground that is 2 feet x 2 feet
- Meno's slave agrees its area is 4 square feet
- Socrates then draws another square on the ground that has an area of 8 square feet
- Socrates then asks: What are the lengths of the sides?
- Meno's slave incorrectly guesses 4 feet initially (the area would be 16 square feet, not 8)
- But Socrates asks Meno's slave a series of questions
- Meno's slave answers the questions correctly and realises that the sides of a square with an area of 8 square feet will be equal to the *diagonal* of the original 2 feet x 2 feet square

Again, Meno's slave has never been taught geometry, i.e. he has no *experience* of geometry. Despite this, Meno's slave is able to correctly answer Socrates' questions (or at least correct his mistakes).

And so, because Meno's slave has no *experience* of geometry, Plato argues that his correct knowledge here must be *innate*.

Leibniz: Necessary truths

Another argument for innate knowledge comes from Gottfried Wilhelm Leibniz (1646-1716), who argued that knowledge of *necessary truths* must be innate.

A necessary truth is something that *must* be true – in every possible world. But how can we know what must be true in every possible world if we only have experience of *this* particular world? Leibniz would argue that the answer is that knowledge of necessary truths is innate.

Contingent truths	Necessary truths
What *is* the case	What *must be* the case
Could have been false in some other possible world	True in every possible world
E.g. "Paris is the capital of France" is a contingent truth because they could have chosen to make Lyon or some other city the capital instead.	E.g. "2+2=4" is a necessary truth because there is no possible world where "2+2" equals 7 or some number other than 4.

Note: For the purposes of A-level philosophy, you can think of necessary truths as similar to analytic truths (page 57).

Leibniz notes how a posteriori experience can only tell us about specific instances. For example, experience can tell us that adding *these* 2 apples to *these* 2 two apples gives us *these* 4 apples. Yet we know that the underlying truth is much deeper than that. It's not just that *in this particular instance* the two sets of apples added together to total 4 apples. We know that *in every instance* adding 2 apples to 2 apples will give us 4 apples. We know that there will never be an instance where you add 2 apples to 2 apples and get 5 apples. In other words, we know that "2+2=4" is a *necessary* truth.

No amount of experience can tell us how things *must* be because even if you conduct 100000 experiments, you never know what could happen in experiment 100001. However unlikely it may seem, the experience of 100000 tests does not prove that it's *impossible* for something different to happen next time.

And yet despite this, we know that it *must always be true* that "2+2=4". We know that it's *impossible* that you could ever add 2 and 2 and get anything other than 4. But this knowledge can't come from experience, argues Leibniz, because experience can only tell us what *is* true, not what is *necessarily* true.

Leibniz argues that we know necessary truths such as "2+2=4" by paying close attention to "what is already in our minds". In other words, Leibniz is saying that knowledge of necessary truths is innate.

Empiricist arguments

John Locke: *Essay Concerning Human Understanding*

Locke's arguments against innate knowledge

Innate knowledge would be universal

Locke argues that if we did have innate knowledge then every human would have such knowledge.

So, for example, everyone would know the theorem of geometry that Meno's slave realises in Plato's example above (page 58). But, Locke argues, "children and idiots" do not possess such knowledge – they don't know the theorems of geometry, for example. So, this knowledge is not universal and therefore, Locke argues, *not innate*.

Argument against innate concepts

Locke's second argument against innate knowledge is an argument against the existence of innate *concepts*.

If we lack innate concepts then we also lack innate *knowledge*, because propositional knowledge relies on concepts. For example, you can't know that "1+1=2" without first having the concepts "1", "+", and "2". So, if Locke succeeds in disproving the existence of innate concepts, he will also succeed in disproving the existence of innate knowledge.

Locke gives two examples to reject the existence of innate concepts:

- Observation of newborn babies suggests they do not have any concepts beyond those experienced in the womb (e.g. the concepts of warmth or pain)
- God is often used as an example of an innate concept (e.g. as seen in Descartes' trademark argument on page 66) but babies do not have this concept. Further, there have been many atheist societies throughout history that did not have the concept of God.

*Possible **response** to Locke's arguments against innate concepts:*

Leibniz agrees with Locke that innate knowledge requires innate concepts. But Leibniz argues that it's possible to have innate concepts and yet not be *conscious* of them.

Some of Leibniz's own examples of innate concepts are logical concepts such as identity (e.g. "a = a") and impossibility (e.g. "it's impossible for both *a* and *not a* to be true"). Obviously, a newborn baby can't verbally articulate these thoughts, but this doesn't mean the concept isn't there. We innately know these concepts – even if we can't articulate them – and they are essential to all thought, whether we consciously recognise them or not. Over time, we learn to recognise these concepts and make them explicit, but they were always there in the mind.

Similarly, just because some people and societies may lack a *word* for "God", this doesn't mean they lack the concept. It may take experience to consciously develop the concept of God, but the concept itself can't come from experience because it goes *beyond* experience. For example, the concept of God is the concept of an infinite being, but nothing in experience shows us this concept of infinity.

Tabula rasa

It follows from Locke's rejection of innatism that all knowledge (and concepts) must come from experience. Locke famously argues that the mind at birth is a *"tabula rasa"* – a blank slate. Locke argues that the mind at birth contains no ideas, thoughts, or concepts. Instead, knowledge comes from two types of experience:

- Sensation: Our sense perceptions – what we see, hear, smell, taste, etc.
- Reflection: Experience of our own minds – thinking, wanting, believing etc.

Simple, complex, and abstract general ideas

Locke also gives an explanation of how humans can form complex concepts, such as the concept of God, entirely from experience. This is his theory of simple, complex, and abstract general ideas.

When I look at a clear sky, my sensation of blue might give me the *simple* concept of blueness. Likewise, when I'm outside in winter, my sensation of cold might give me the *simple* concept of coldness.

Complex concepts are made up of the building blocks of simple concepts. For example, my concept of the ocean could consist of *both* the simple concepts of blue and cold above. Pretty much everything is a complex concept made up of simple concepts to differing degrees – for example a chair might consist of many simpler concepts (e.g. brown, hard, wooden, small, etc.) and yet they all form the same thing: *this* chair. This chair is a complex concept because it contains many simpler concepts within it.

But complex ideas can go beyond specific instances of things (e.g. *this* chair) to form abstract ideas (e.g. chairs *in general*). For example, chair #1 may be wooden and have 4 legs, but chair #2 may be metal and have 3 legs. Chair #3 may be a red plastic stool. We can form the *general* concept of chair by ignoring the irrelevant features of each chair (e.g. colour, material, number of legs) and abstracting the features common to all the individual chairs.



We form the complex idea of chair from experience of specific instances of chairs

Similarly, we form abstract concepts such as beauty or justice by abstracting from experience. For example, we see a beautiful lake, a beautiful painting, a beautiful person, etc. and over time we abstract the common features from these experiences to form the abstract concept of beauty. Locke claims all concepts – from the simple to the complex – are derived from experience in some way, and so are not innate.

RATIONALISM (INTUITION AND DEDUCTION) VS. EMPIRICISM

Rationalism vs. empiricism is a debate about what kind of knowledge we can acquire through a priori *intuition and deduction:*

- (Rational) intuition: The ability to know something is true just by thinking about it
  - E.g. Descartes' cogito argument below (page 65)
- Deduction: A method of deriving true propositions from other true propositions (using reason)
  - E.g. deducing statement 3 from statements 1 and 2 below:
    1. If A is true then B is true
    2. A is true
    3. Therefore, B is true

Earlier (page 56), we defined rationalism as the view that we can acquire some knowledge purely through intuition and deduction. This is slightly inaccurate, though, because even an empiricist would admit there is *some* knowledge that can be known purely through intuition and deduction. For example, you don't need empirical experience to know "2×75=150" because it is an analytic truth.

So really, what the debate between rationalism and empiricism is about is whether we can acquire knowledge of *synthetic truths* using intuition and deduction:

- Empiricism says *all* a priori knowledge is of analytic truths (i.e. there is no synthetic a priori knowledge)
- Rationalism says *not all* a priori knowledge is of analytic truths (i.e. there is at least one synthetic truth that can be known a priori using intuition and deduction)

Most of the time, empiricism holds true. If you take any synthetic truth, such as "water boils at 100°c", it seems impossible that we could learn it without some a posteriori experience of the world (e.g. an experiment).

So, most synthetic truths are known a posteriori. But the question is whether this relationship holds true *all* the time or just *some* of the time. Rationalism says the latter: there is *at least one* synthetic truth that can be known a priori through intuition and deduction.

Rationalist arguments

Rene Descartes: *Meditations*

Rene Descartes (1596-1650) attempts to establish various synthetic truths using intuition and deduction alone.

The synthetic truths Descartes argues for are:

- I exist
- God exists
- The external world exists

Again, intuition and deduction are a priori methods of gaining knowledge. So, if Descartes' arguments work, then they support the rationalist claim that some synthetic truths can be known a priori.

Three waves of doubt

As we saw in definitions of knowledge, Descartes is an infallibilist (page 36). Descartes only wants knowledge if it is *certain* and so before attempting to establish what is true, Descartes begins by doubting everything he thinks he knows. If it's possible to doubt it, then he does.

These sceptical arguments have come to be known as the three waves of doubt. They are:

- Illusion
- Dreaming
- Deception (the evil demon)

First off, Illusion. I can doubt the reliability of my sense experience as it has deceived me in the past. For example, a pencil in water may look crooked even though it isn't, or from a distance an object might look different to how it actually is.

But this isn't reason to doubt *everything*. Illusions don't happen *that* often, and when they do we can generally tell. The very fact that we're able to recognise illusions *as* illusions shows that they are weird exceptions to otherwise normal experience.

The second wave of doubt goes further, though. I might think I'm awake when I'm actually dreaming. So, I might believe I'm looking at a computer screen, but if I'm simply *dreaming* that I am, then that belief is false.

This is a stronger level of doubt: everything you perceive could be false, because in a dream all your perceptions are imaginary and yet you believe they are real. However, says Descartes, even when you're dreaming there are still *basic ideas* that are common to

both dreams and reality. For example, the ideas of space and size and motion are common to both dreams and reality. This suggests that basic truths such as geometry and maths are safe from the second wave of doubt.

But the third wave of doubt undermines even the most basic knowledge: an evil demon could be controlling my entire experience, making me think "1+1=2" is true when actually it's false. So, this third wave of doubt means basically *everything* I think I know can be doubted. An evil demon could be controlling all my thoughts and perceptions and making me have nothing but false beliefs. I can't trust anything – not even maths or basic ideas like space and shape.



This position of extreme doubt is known as *philosophical scepticism* (see the limits of knowledge topic on page 72 for more detail).

In this sceptical position, Descartes rules out pretty much all knowledge. He can't trust his perceptions (because the evil demon may be deceiving him) and so he has only his mind – i.e. intuition and deduction – from which to establish knowledge.

Cogito argument for the existence of the self

So, Descartes is currently in a position of extreme scepticism. But is there anything he can know for certain in this situation?

Yes, says Descartes, this:

dubito, ergo cogito, ergo sum

This translates as:

- I doubt
- Therefore I think
- Therefore I am

The shortened version – *cogito ergo sum* – is probably the most famous phrase in philosophy. Descartes' point is that even if the evil demon is deceiving him on just about everything, Descartes can't doubt that he exists. The reason for this is that even if the

demon is deceiving him, there must be something for the demon to deceive in the first place! The fact that Descartes is able to doubt his own existence is proof that he does, indeed, exist. So, even from the most extreme position of scepticism, Descartes can be certain of one truth: "I exist".

Note that "I exist" is a synthetic truth. And so, if Descartes has proven "I exist" using only intuition and deduction, then he has proven a synthetic truth using a priori methods only. Thus, the cogito argument would prove rationalism is correct.

For empiricist responses, see page 69.

Clear and distinct ideas

For Descartes, the cogito is less of an *argument* and more of a *rational intuition*. This means it's not a matter of deriving the conclusion "I exist" from the premise "I think" but is instead something that the thinker can *see* is true simply by thinking about it. Descartes claims that the proposition "I exist" is a *clear and distinct idea* – something that the thinker can see is true using rational intuition.

Descartes argues that he can trust clear and distinct ideas, and this gives him a sort of platform from which he can reason his way out of his position of extreme scepticism. From this platform, Descartes argues that he can know "God exists" and, later, that "the external world exists".

Trademark argument for the existence of God

Descartes gives more than one argument for the existence of God using intuition and deduction. We cover Descartes ontological argument (page 156) and cosmological argument (page 169) in the metaphysics of God module, so in this section we shall just focus on Descartes' trademark argument.

This argument is called the "trademark" argument because Descartes argues that concept of God is like an innate trademark placed in our minds. A summarised version of the trademark argument is as follows:

1. I have the concept of God
2. My concept of God is the concept of something infinite and perfect
3. But I am a finite and imperfect being
4. The cause of an effect must have at least as much reality as the effect
5. So, the cause of my concept of God must have as much reality as what the concept is about
6. So, the cause of my idea of God must be an infinite and perfect being
7. So, God exists

This summary misses out a few steps but, for the purposes of the exam, this is sufficient detail. Descartes' point is essentially that it's impossible that he himself could have come up with this concept of God – an infinite and perfect being – all by himself. Whatever caused Descartes' idea of God must itself be an infinite and perfect being because, Descartes argues, "the cause of an effect has as much reality as the effect" (i.e. the cause must be as real as the effect).

Again, "God exists" is a synthetic truth (unless ontological arguments work, but that's a topic for page 156) and Descartes has seemingly established it using intuition and deduction – i.e. a priori methods – only. Thus, if Descartes' arguments work here, he has established a synthetic truth using a priori means and shown that rationalism is correct.

For empiricist responses, see page 69.

Argument for the existence of the external world

Descartes then continues his process of intuition and deduction to argue that because God exists, his perception can be trusted. These perceptions are of an external world of physical objects and, because Descartes can trust his perceptions, he can trust the external world exists.

The key points of the argument are as follows:

1. I have perceptions of an external world with physical objects
2. My perceptions cannot be caused by my own mind because they are involuntary (Descartes' argument for this is similar to e.g. Locke's argument on page 48)
3. So, the cause of my perceptions must be something external to my mind
4. God exists (see trademark argument above)
5. If the cause of my perceptions is God and not the physical objects themselves, then God has created me with a tendency to form false beliefs from my perception (because premise 1)
6. But God is a perfect being by definition (see e.g. Descartes' ontological argument on page 156) and so would not create me with a tendency to form false beliefs from my perceptions
7. So, I can trust my perceptions
8. So, given premises 1 and 7 above, I can know that an external world of physical objects exists

Once again, "the external world exists" is a synthetic truth – there's no logical contradiction in the idea of the world never existing. Also, Descartes' reasoning here is a priori: he can't use his perceptions to prove the external world exists because the evil demon argument ruled that out earlier (page 64). Instead, Descartes uses reasoning

outside of perception to show that his perceptions can be trusted and that the external world exists.

So, "the external world exists" is another synthetic truth proved through a priori intuition and deduction.

For empiricist responses, see page 70.

Empiricist arguments

David Hume: *An Enquiry Concerning Human Understanding*

Hume's fork

David Hume (1711–1776) was an empiricist. Hume's fork is the claim that there are only two kinds of knowledge: relations of ideas and matters of fact.

There is a lot of overlap with Hume's distinctions here and the analytic/synthetic and a priori/a posteriori distinctions we looked at earlier (page 56). However, Hume was writing before these distinctions were a thing, which is why there is this duplicate terminology.

Relations of ideas	Matters of fact
"either intuitively or demonstratively certain" (i.e. an analytic truth)	*"The contrary of every matter of fact is still possible; because it can never imply a contradiction"* (i.e. a synthetic truth)
"discoverable by the mere operation of thought, without dependence on what is any where existent in the universe" (i.e. they are known a priori)	Can't be established purely by thought and thus require empirical observation to establish their truth (i.e. they are known a posteriori)
The key feature of a relation of ideas is that it cannot be denied without a contradiction	The key feature of a matter of fact is that there is no logical contradiction in it being false

The key difference between a relation of ideas and a matter of fact is that the denial of a matter of fact is possible, but the denial of a relation of ideas is impossible. For example, "grass is green" is true as a *matter of fact* because it's logically possible for this to be false: we can imagine red grass or purple grass, for example. In contrast, it is logically impossible to deny a *relation of ideas* such as "triangles have 3 sides". The very idea of a triangle with 4 sides, for example, is inconceivable. So, unlike "grass is green", it's not logically possible for "triangles have 3 sides" to be false.

Hume argues we can know the truth of relations of ideas just by thinking about them – *thought alone* reveals that it's impossible for "triangles have 3 sides" to be false. In contrast, we have to use empirical experience to verify matters of fact – we can't know

"grass is green" without empirical verification because it's logically possible that grass is red or purple.

Applying Hume's Fork to some of Descartes arguments, we can argue that they rely on *matters of fact*. But, according to Hume's Fork, matters of fact are a posteriori. Thus, if Hume's Fork is correct, then it shows that Descartes' arguments are not entirely a priori and thus fail to establish rationalism.

Responses to rationalism

Response to Descartes' cogito argument

Descartes concludes "I exist" from "I think", but Hume argues that this doesn't necessarily follow.

Hume points out that what we actually experience is a constantly changing array of thoughts and perceptions. However, we never actually experience a *self*. There are thoughts, yes, but this does not guarantee that there is a constant *thing* that actually thinks those thoughts, or that this *thing* is the same thing from one thought to the next. For example, the evil demon could just be creating a bunch of disembodied false thoughts that "I exist".

In other words, there is no logical contradiction in the idea of a thought without a thinker, and so you cannot deduce "I exist" from "I think" using a priori reasoning only.

Responses to Descartes' arguments for the existence of God

See the Metaphysics of God section for more responses to Descartes' ontological argument (page 157) and cosmological argument (page 169).

Causation requires experience

Both Descartes' trademark argument (page 66) and Descartes' cosmological argument (page 169) rely on a premise that is something like "everything has a cause" or "every cause must have at least as much reality as the effect". However, if we show that knowledge of causation requires *experience*, then we can show that these arguments are not a priori and thus fail to establish rationalism.

To show this, we can argue that the premise "everything has a cause" *can* be denied without contradiction. It's a weird idea, but there is no *logical* contradiction in the idea of an uncaused event. So, applying Hume's fork (page 68), we can argue that this shows "everything has a cause" is a *matter of fact* rather than a relation of ideas. And, according to Hume, all matters of fact require empirical experience (i.e. they cannot be known through intuition and deduction alone).

So, Hume's fork suggests Descartes' cosmological and trademark arguments are not purely a priori. And if Descartes' arguments are not a priori, then they do not demonstrate the existence of synthetic a priori knowledge.

The concept of God comes from experience

In his trademark argument (page 66), Descartes claims that his concept of God cannot come from experience and must be innate. However, if we can show this concept does come from experience, then we show that Descartes' argument is not a priori and thus fails to establish rationalism.

We've already seen Locke's argument that the concept of God is not innate (page 60). Hume agrees with Locke on this and provides his own explanation of how we can form the concept of God from experience.

Descartes claims it's impossible that he – a finite and imperfect being – could come up with the concept of God – an infinite and perfect being – by himself. But Hume disagrees. Hume argues that we can create concepts such as infinity and perfection by taking the finite and imperfect ideas we get from experience and extending these "beyond all limits". For example, experience gives us the concept of goodness and so to create the concept of a perfectly good being (God), we take that concept and extend it beyond all limits. Similarly, the concept of infinity is not something we experience directly, but it is something we can create by negating the concept of finite. In other words, "infinite" just means "not finite".

If Hume is right that the concept of God comes from experience, then Descartes' trademark argument is not a priori and thus fails to establish rationalism.

Responses to Descartes' argument for the existence of the external world

Relies on Descartes' argument that God exists

Descartes' proof of the external world relies on the premise that God exists and that God is the cause of his perceptions. But if we can cast doubt on Descartes' arguments for these claims and whether they are a priori (see above), then we cast doubt on his argument for the existence of the external world too.

Causation (again)

Like Descartes' cosmological and trademark arguments, Descartes' proof of the external world relies on knowledge of causation. For example, Descartes argues that because his perceptions are involuntary they must be *caused* by something external to his mind.

But, as mentioned above (page 69), knowledge of causation requires *experience*. Hume would say there is no logical contradiction in the idea of a perception that has no cause

and so we cannot know a priori that our perceptions are caused by something external. Because Descartes' argument here relies on a posteriori knowledge of causation, he fails to prove the synthetic truth that "the external world exists" a priori and thus fails to prove rationalism.

THE LIMITS OF KNOWLEDGE

OVERVIEW

This topic looks at **philosophical scepticism** and considers whether it is possible to know anything at all.

We've already touched on a lot of the arguments involved in this discussion already – the main one being Descartes' three waves of doubt (page 64). Descartes' third wave of doubt – the evil demon argument – is an example of global/philosophical scepticism because it casts doubt on pretty much everything we ordinarily consider to be knowledge.

The point of considering extreme sceptical scenarios like the evil demon is to determine what, if anything, we can know. We've already covered several potential responses to extreme scepticism including Descartes' own response (page 66), Russell's hypothesis that the external world is the best hypothesis (page 47), and reliabilism (page 33).

Philosophical doubt vs. ordinary doubt

In our everyday lives, we doubt things all the time. For example, you might be unsure whether a friend's birthday is the 17th or the 18th of August, or what time the philosophy exam is, or you might doubt your memory of a fact such as "Paris is the capital of France". This is "normal incredulity" or "ordinary doubt" in the language of the philosophy syllabus – it is doubt about some *specific* claim (or group of claims). *Philosophical* doubt goes beyond such ordinary doubt and provides reasons to doubt pretty much *everything* we think we know because of global sceptical scenarios.

There are a bunch of global sceptical scenarios in philosophy and popular culture, such as:

- Descartes' evil demon
- The Matrix
- The simulation hypothesis
- Perfect virtual reality
- Brain in a vat

This last one – brain in a vat – comes up quite often in philosophy. The scenario is this: if all your experience is just electrical signals interpreted by your brain, then you wouldn't

be able to tell the difference if you were a disembodied brain in a vat being fed these same electrical signals *artificially*.



So, in the brain in a vat scenario, you might think "I'm outside walking my dog" but in reality you are just a disembodied brain in a vat being fed electrical signals that make it *feel* like you are outside walking your dog. So, your belief "I'm outside walking my dog" would be false.

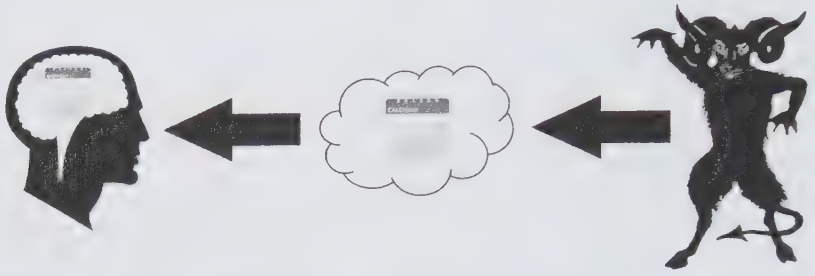
All the global sceptical scenarios above play the same role here: they are situations where everything you believe could be false and there would be no way of knowing. Global scepticism thus undermines all our usual justifications for belief.

SCEPTICISM

Descartes' evil demon

We've already covered Descartes' three waves of doubt in the reason as a source of knowledge section (page 64). The third wave of doubt - the evil demon argument - is an example of global scepticism.

In cases of ordinary doubt, you can often resolve them by referring to some sort of justification. If I doubt which day my friend's birthday is, for example, I can check my calendar to make sure. Once I see their name under the 17th, I can be confident my belief that "my friend's birthday is on the 17th" is true.



But global scepticism scenarios like Descartes' evil demon go beyond such ordinary doubt and cast uncertainty over basic propositions such as this. If all my experience is just an illusion created by an evil demon, then I can't trust my calendar - I can't even trust my visual perception. For all I know, my friend doesn't even exist and my perceptions of him were just an illusion created by the evil demon. If I can't even be sure my friend *exists*, then I definitely can't be sure what day their birthday is.

You can run the same sort of argument for pretty much any ordinary knowledge proposition:

- "I know that grass is green"
  - Nope, the evil demon was tricking you and grass is actually purple
- "I know that Paris is the capital of France"
  - Nope, the evil demon was tricking you - the world you think you know was just an illusion created by the evil demon. France doesn't even exist, let alone Paris
- "I know that 2+2=4"
  - Nope, the evil demon is messing with you again. 2+2 actually equals 5 and each time you add the two numbers together, the evil demon messes with your mind and makes you think the answer is 4

All of these scenarios are *possible* - and there is no way you would be able to tell otherwise. Thus, the possibility of the evil demon casts doubt over everything we ordinarily consider to be knowledge. We could be constantly making mistakes and there is no way of telling otherwise.

The global sceptic thus concludes that *any knowledge* is impossible.

Responses to scepticism

Descartes' own response to scepticism

We covered Descartes' own responses in the knowledge from reason section earlier.

Starting from page 65, we saw how Descartes' cogito argument shows that, even if he is being deceived by the evil demon, he can at least be certain of the proposition "I exist".

Descartes then goes on to argue that he can also know that "God exists" (see e.g. his arguments on pages 66, 156, and 169).

Having established that God exists, and that God would not allow him to be globally deceived, Descartes concludes that he can trust his perceptions and so he can trust that "the external world exists" (page 67). If Descartes can trust his perceptions, then his perceptions justify his knowledge of ordinary propositions, and thus Descartes has defeated global scepticism.

Russell's response to scepticism

In the indirect realism topic, we saw Russell's response to the sceptical challenge that says the existence of the external world is the best hypothesis (page 47). Adjusting that argument slightly for the evil demon scenario, we could say the following:

- Either:
  - A: The external world exists and causes my perceptions
  - B: An evil demon exists and causes my perceptions
- I can't *prove* A or B definitively
- So, I have to treat A and B as hypotheses
- A is a better explanation of my experience than B
- So, mind independent objects exist and cause my perceptions

*Possible **response #1** to the claim that the existence of the world is the best hypothesis:*

Russell's claim that we can't prove A or B definitively was kind of Descartes' whole point! Descartes wasn't trying to *prove* the evil demon exists, he was just saying that it's *possible* – i.e. that it was a viable hypothesis – and so we can't trust our ordinary knowledge.

However, Russell could respond to Descartes that the *possibility* of the evil demon hypothesis does not mean knowledge is impossible. Just because we can't know for *certain* that we're not being deceived by an evil demon, that does not mean we can't have knowledge. Descartes is assuming an infallibilist definition of knowledge (page 36) here – he's assuming that we have to know we're not being deceived by an evil demon in

order to have knowledge. But Russell could respond that *certainty* is not necessary: sure, we *might* be being deceived by an evil demon and we can't know either way, but as long as we're *not* being deceived and our beliefs are true then our ordinary (uncertain) justifications are sufficient for knowledge.

*Possible **response #2** to the claim that the existence of the world is the best hypothesis:*

Another way the sceptic could respond to Russell's argument is to dispute that the external world is the best hypothesis at all. If we are being deceived by an evil demon, or are in the matrix, or are brains in vats, then our experience would appear exactly the same to us as if it was caused by the external world. Thus, we have no grounds to prefer one hypothesis over the other: the available evidence supports both hypotheses equally.

Locke's response to scepticism

We also saw in the indirect realism section that Locke provides 2 responses to the sceptical challenge. These were:

1. Perception, unlike imagination, is *involuntary* – I can't choose what I see or hear. This suggests perception is caused by something external to my mind (page 48)
2. My different perceptions (e.g. sight and touch) are *coherent* – I can touch the fire with my hand to verify what I see with my eyes. This suggests there is a common reality that causes both perceptions (page 48)

*Possible **response** to Locke's arguments against scepticism:*

For the first point, we can respond that even if Locke succeeds in proving that *something* external is causing his perceptions, he doesn't succeed in proving that this perception is in any way an accurate representation of the external world. The external something causing his (involuntary) perceptions could be an evil demon that is deceiving him rather than a world of mind-independent objects.

The response to Locke's second point is similar: just because our different senses are coherent, it doesn't do much to prove they are representative of reality. An evil demon could be creating coherent experiences – e.g. the demon could deceive you into *hearing* dogs barking at the same time as *seeing* dogs – and there is no way you could tell otherwise. Locke gives the example of *seeing* a fire and then confirming its existence by *touching* it and feeling heat – but surely an evil demon capable of creating a visual perception of a fire would be capable of creating a feeling of heat when you touch the fire too! So, the coherence of the different senses fails to rule out deception by evil demon.

Berkeley's response to scepticism

The evil demon argument never really gets going against idealism because idealism doesn't make a distinction between perceptions and reality: "to be is to be perceived" as

Berkeley says. It doesn't make sense to claim that the idea you are perceiving is somehow different from "reality" because, according to idealism, ideas *are* reality.

Berkeley realises his perceptions must be caused by something external to him and, given the complexity of these perceptions, Berkeley concludes that this cause must be the mind of God (see page 52). You can think of Berkeley's conception of reality as kind of like a *benevolent* version of the evil demon: God is causing his perceptions, but He is not deceiving him in doing so, because these ideas *are* reality.

*Possible **response** to Berkeley's argument against scepticism:*

This response assumes idealism is the correct theory of perception. But idealism faces its own problems, such as hallucination (page 54).

If we accept that hallucinations are not "real" perceptions, then the sceptical problem re-emerges: it is possible that *all* our perceptions are hallucinations (perhaps caused by an evil demon) and thus none of our perceptions are real. This gives us reason to doubt all our ordinary knowledge, because it could all be formed from hallucinations.

Reliabilist responses to scepticism

The reliabilism definition of knowledge (page 33) says that knowledge is true belief formed via a reliable method. Assuming I'm not being actually deceived by an evil demon, then my perception would count as a reliable method of gaining knowledge because my perceptions reliably cause me to form true beliefs.

Of course, I can't prove either way whether I'm being deceived by an evil demon - but that doesn't matter. I don't have to *know* my perception is reliable in order for it to count as a reliable method – it just is.

Consider the following two scenarios:

- Scenario 1: I am *not* a brain in a vat
  - My perception is a reliable method because I'm living in the real world and am perceiving it accurately
  - My perception leads me to the belief "I have hands"
  - My belief is true, because this is scenario 1 and I'm not a brain in a vat
  - So I have a true belief formed via a reliable method that "I have hands"
  - So, according to reliabilism, I *know* "I have hands" in scenario 1
- Scenario 2: I *am* a brain in a vat
  - My perception is not a reliable method because I'm a brain in a vat being fed artificial stimuli
  - My perception leads me to the belief "I have hands"

- But my belief is false, because this is scenario 2 and I'm actually a brain in a vat
- So I have a false belief formed via an unreliable method that "I have hands"
- So, according to reliabilism, I *do not* know "I have hands" in scenario 2

We can't know whether we're in scenario 1 or 2 – *but we don't have to*. The point is: *if we're in scenario 1,* then we can have knowledge of ordinary propositions such as "I have hands" because our perception would count as a reliable method.

Yes, this means that we can't *know that we know* such propositions, but we don't have to. As long as we're not in some weird sceptical scenario like the evil demon or brain in a vat, then knowledge is possible if "knowledge" is defined as true belief formed via a reliable method.

*Possible **response** to reliabilist arguments against scepticism:*

This response assumes reliabilism is the correct definition of knowledge. But if we go with some other definition of knowledge – such as infallibilism or justified true belief – then we can't provide proper justification for ordinary knowledge because we can't justify that we're not in some sceptical scenario. So, according to most other definitions of knowledge, I cannot know "I have hands" without first defeating the sceptical challenge.

EPISTEMOLOGY – MODULE SUMMARY

What is knowledge?

- We are defining *propositional* knowledge (i.e. knowledge of facts), not acquaintance or ability knowledge. A good definition will consist of conditions that are:
  - Individually necessary (i.e. if something doesn't have this condition then that thing isn't knowledge)
  - Jointly sufficient (i.e. anything that meets all these conditions would be an instance of knowledge)
- **Definition 1: The tripartite definition of knowledge as Justified, True, Belief**
  - Problem: Gettier cases, e.g.:
    - Smith and Jones are interviewing for the same job
    - Smith hears the interviewer say "I'm going to give Jones the job"
    - Smith also sees Jones count 10 coins from his pocket
    - So, Smith forms the belief that "the man who will get the job has 10 coins in his pocket"
    - But Smith gets the job, not Jones
    - And, by sheer coincidence, Smith also happens to have 10 coins in his pocket
    - Smith's belief that "the man who will get the job has 10 coins in his pocket" is a justified true belief that is *not* knowledge (because it's only through luck that he's right). Therefore, the JTB definition is not sufficient because you can have a JTB that is not knowledge
- **Definition 2: Justified true belief + no false lemmas**
  - In the Smith/Jones example above, Smith's belief is inferred from the *false lemma* (i.e. false premise/false belief) that "Jones will get the job". If we rule out false lemmas we avoid the Gettier case above.
    - Problem: Fake barn county:
      - Henry is driving through fake barn county looking at barns and (falsely) thinking: "there's a barn"
      - But one time he looks at the only real barn and thinks: "there's a barn"

        - This time, his belief is true and justified and not inferred from a false lemma. The JTB+N definition would say it is knowledge but it clearly isn't knowledge because it's just lucky that he's right this time. So, the JTB+N definition is not sufficient either
- **Definition 3: Reliabilism: True belief formed via a reliable method** (e.g. good eyesight and good memory are reliable methods)
    - Problem: Fake barn county again. Henry has perfect eyesight, good memory, etc. and so forms his true belief via a reliable method. But it still isn't knowledge because it's still just lucky that he's right this time
- General problem for all the definitions above: Zagzebski: We can always come up with Gettier cases if we define knowledge as true belief + *something*. Here is the formula:
    - Start with a scenario where someone forms a false belief that has all the extra conditions (justified, no false lemmas, reliable, etc. Let's call this JTB+ABC) but is false due to bad luck
    - Now repeat the scenario where the belief is true due to good luck
    - The belief will be a JTB+ABC, but won't be knowledge because it will just be lucky
- **Definition 4: Infallibilism: True belief that is justified in such a way that it is impossible to doubt**
    - Problem: Too strict. By this definition we can't know "grass is green" or "London is the capital of England" even though we clearly do know such things. In other words, certainty/the impossibility of doubt is not necessary for knowledge
- **Definition 5: Zagzebski's virtue epistemology: A belief arising from an act of intellectual virtue**
    - "Intellectual virtue" has two components:
        - Pursues what is good (in the case of knowledge, good = true)
        - Enables it to be successful (beliefs are true *because* of the way they are formed)

Knowledge from perception

- **Direct realism:** The immediate objects of perception are mind-independent objects (i.e. we perceive the external world directly as it is)
  - Problems:
    - Time lag: e.g. we perceive stars as they were millions of years ago, not as they currently are. Therefore, we are not perceiving reality directly
      - Reply: This confuses *how* we perceive with *what* we perceive
    - Perceptual variation: e.g. a table looks square from a birds-eye view but kite shaped from a different angle. It can't be both and so at least one perception does not represent reality
      - Reply: (Mind-independent) *relational* properties, e.g. the table has the property of looking kite-shaped relative to a perceiver's position, in the same way whether an object is left or right of you depends on where you are standing
    - Illusion: E.g. we perceive a stick as crooked when it's in water, when in reality it is not crooked
      - Reply: Relational properties again. The stick has the property of *looking* crooked from a certain perspective
    - Hallucination: e.g. when you take drugs or are ill you perceive things that do not exist in reality
- **Indirect realism:** The immediate objects of perception are (mind-dependent) sense data, which are caused by and represent mind-independent objects
  - Locke: Primary and secondary qualities
    - Primary: Inherent in the object itself (e.g. size, shape, motion, number) - i.e. objective, mind-independent
    - Secondary: Powers of objects to cause sensations in humans (e.g. colour, taste, smell, feel) - i.e. subjective, mind-dependent
  - Problems:
    - Scepticism/veil of perception: How do we know there are mind-independent objects that cause our sense data/perceptions?
      - Indirect realist replies to scepticism:
        - Locke: Involuntary nature of experience: when I imagine things, I can *choose* what I perceive, but when I look at things in the real world, I can't choose what I perceive. This suggests

the cause of my perceptions is something external to me

- Locke (again): Coherence of different senses: we can confirm the same things with different senses. E.g. when you *see* a fire, you can touch it with your hand and *feel* it too. Or, you can write a letter on a piece of paper and see it, and someone can also read that same letter back to you and you can hear it. This coherence suggests there is a common reality causing both perceptions
- Russell: External world is the best hypothesis: the hypothesis that my cat exists and walked from the floor to the sofa when I wasn't looking is a more likely explanation for why the cat is now on the sofa than the explanation that the cat just disappeared and reappeared on the sofa randomly

- **Idealism:** The immediate objects of perception are mind-dependent ideas (i.e. mind-independent objects don't exist – ideas *are* reality)
  - Arguments for idealism:
    - Berkeley's attack on Locke's primary/secondary quality distinction: Primary qualities (i.e. size, shape, motion) vary just as much as secondary qualities (i.e. colour, taste, etc.). For example, something that looks *small* (primary quality) to me might look *big* to an ant. Given these variations, we can't say objects have one single size, shape, or motion *independent of perception*. This shows that primary qualities are just as mind-dependent as secondary qualities are.
    - Berkeley's master argument: the idea of an unperceived mind-independent object is inconceivable. E.g. try to think of a tree that no one is perceiving - you can't, because the very act of thinking about it would mean you are perceiving it!
      - Reply: The *idea* of an unperceived tree is inconceivable, yes, but that doesn't mean it is impossible *in reality*. Clearly, all ideas are mind-dependent and so any idea about an object will also be mind-dependent. But just because *ideas* about objects are mind-dependent, it doesn't follow that *what the idea is about* is also mind-dependent.
  - Arguments against idealism:
    - Idealism leads to solipsism (the view that one's own mind is the only thing that exists): If everything is mind-dependent, how do I know anything else exists? It could all just be dependent on my mind

- Reply: Berkeley proves the existence of at least one other mind: God's mind:
  - There are 3 possible causes of these perceptions: ideas, my own mind, or another mind
    - It can't be ideas because ideas by themselves don't cause anything - ideas need a mind to think them
    - It can't be my own mind because if I was the cause of my own perception then I'd be able to control what I perceived
    - So, the cause of my perception must be another mind
      - And, given the complexity and manner of my perceptions, this other mind must be God

- Problems with the role of God in Berkeley's idealism: God is said to be eternal and unchanging, but my perceptions constantly change and go in and out of existence. Also, God doesn't feel pain, but my perceptions often include pain. So, how can my perception of pain be an idea in God's mind (as Berkeley claims) if God Himself cannot feel pain?
  - Reply: Berkeley says what we perceive are copies of ideas that exist in God's understanding. Although God does not perceive pain or change Himself, He *understands* what it is for us to undergo these experiences. Our changing perceptions are what God actively wills us to perceive.
- Illusions: Idealism says that in perceiving ideas, we are perceiving reality directly. But we may perceive a pencil in water as looking crooked, but we don't think the pencil is actually crooked in reality. So, idealism is false: our perception is different from reality
  - Reply: Berkeley would just say that our perception of the crooked pencil is just as real as the perception of the normal pencil
    - Reply to reply: But what about hallucination and dreaming? Are we really supposed to believe that my hallucination of gremlins or my dream about flying is as real as my ordinary perceptions?
      - Berkeley's reply: No, hallucination and dreaming are caused by our own minds (i.e. imagination), and so we can see that these perceptions are not real. Also, hallucinations/dreams are "dim, irregular, and confused"

Knowledge from reason

- Two separate debates: innatism vs. empiricism and rationalism vs. empiricism.
- **Innatism vs. empiricism**
  - Innatist arguments: We are born with some innate (propositional) knowledge
    - Plato: Meno's slave has never been taught geometry and yet he knows that a square with an area of 2A will have sides that are the length of the diagonal of a square with an area of 1A.
    - Leibniz: Experience only tells us what *is* the case, not what *must be* the case. However, we know that necessary truths such as "1+1=2" must always be the case, and so this knowledge can't come from experience.
  - Empiricist arguments: We are born a blank slate with no innate knowledge:
    - Locke: Innate knowledge would be universal, but there is no evidence that children and idiots have knowledge of facts such as the theorems of geometry (i.e. Plato's example of supposedly innate knowledge).
    - Locke #2: Innate knowledge would require innate concepts but there is no evidence that babies have any concepts beyond those that come from experience.
- **Rationalism vs. empiricism**
  - Rationalist arguments: There is at least one synthetic truth that is knowable a priori (i.e. using intuition and deduction)
    - Descartes' cogito argument: I think, therefore I exist. "I exist" is a synthetic truth proven purely through a priori means
      - Empiricist response: "I exist" does not necessarily follow from "I think" – there could be a thought without a thinker
    - Descartes' trademark argument: I have the concept of God as an infinite and perfect being. The cause of an effect must be as real as the effect. Therefore the cause of my concept of God must be an infinite and perfect being and so God exists. "God exists" is a synthetic truth proven purely through a priori means
      - Empiricist response #1: Descartes' premise that "the cause of an effect must be as real as the effect" is a matter of fact and requires empirical experience to know. Therefore the trademark argument is not a priori
      - Empiricist response #2: The concept of God comes from experience

and so the trademark argument is not a priori

- Descartes' argument for the existence of the external world: God is a perfect being and so would not create me with a tendency to form false beliefs. My perceptions cause me to believe there is an external world and so God would not allow me to be deceived about this, so I can trust that the external world exists. "The external world exists" is a synthetic truth proved through a priori means
  - Empiricist response #1: Descartes' argument for the existence of the external world relies on his previous arguments that "God exists". If these arguments fail, the argument for the existence of the external world does too
  - Empiricist response #2: Descartes cannot know a priori that his perceptions are caused by something external to him, because causation is not a relation of ideas but a matter of fact – there is no contradiction in the idea of uncaused perceptions

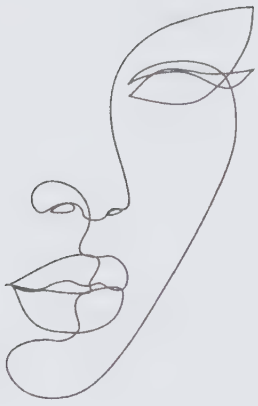
- Empiricist arguments: All knowledge of synthetic truths is a posteriori
  - Hume's fork: There are two (and *only* two) types of knowledge:
    1. Relations of ideas: Cannot be denied without contradiction and so can be proven purely by thought (i.e. analytic a priori)
    2. Matters of fact: Can be denied without contradiction and so require empirical verification to confirm they are true (i.e. synthetic a posteriori)

The limits of knowledge

- Ordinary doubt: Doubting whether you know some specific thing, e.g. I'm not sure I locked the door when I left and so I'm not sure my belief "the door is locked" is true. Or "I'm not sure if his birthday is on the 17th or the 18th"
- Global scepticism/philosophical doubt: Disbelieving/doubting *all knowledge* because of possible sceptical scenarios. Sceptical scenarios are possible situations that undermine all ordinary justifications and mean you could be wrong about everything without knowing you are wrong. For example:
  - Brain in a vat: You're actually just a disembodied brain in a vat being stimulated artificially to believe you are living in the real world. All your beliefs about the real world are false
  - Descartes' evil demon: An evil demon is causing all your perceptions and tricking you into believing you're living in the real world when you're not.

All your beliefs about the real world are false because the evil demon tricks you about everything

- Responses to global scepticism:
  - Descartes: I know I exist (cogito argument), I know God exists (trademark argument), and I know God wouldn't let me be constantly deceived. So, I can trust my perceptions and the knowledge that comes from my perceptions
  - Russell: External world is the best hypothesis. Sceptical scenarios are possible, sure, but the hypothesis that my perceptions are caused by and resemble mind-independent objects is a better explanation
  - Locke: Perceptions are involuntary (I can't choose what I perceive) and so this shows they are caused by something external to me. Also, my perceptions are coherent (e.g. sights and sounds occur together), which suggests a common reality to both
  - Berkeley: If idealism is true, there is no distinction between perception and reality. So, sceptical scenarios where perception is somehow *different* from reality never get going with idealism
  - Reliabilism: Knowledge = true belief formed via a reliable method. If I'm not a brain in a vat (or some other sceptical scenario), then my perception counts as a reliable method and my ordinary beliefs will be true. I can't know whether I'm in a sceptical scenario or not, but I don't need to know my perception is reliable in order for it to be reliable



MORAL PHILOSOPHY

Moral philosophy – sometimes called ethics – is about right and wrong, good and bad. This module covers three competing ethical theories as well as several practical applications of these theories to moral issues. This module also looks at metaethics: what moral judgements, such as "murder is wrong" actually mean and whether they are capable of being true or false.

NORMATIVE ETHICAL THEORIES

OVERVIEW

Ethics is the study of morality – i.e. right and wrong, good and bad. The A-level philosophy syllabus looks at three ethical theories:

1. **Utilitarianism**
2. **Kant's deontological ethics**
3. **Aristotle's virtue ethics**

Each theory provides a framework intended to guide moral behaviour. We can apply these theories to ethical dilemmas such as "is it OK to steal?"

For utilitarian theories, what matters is the *consequence* of an action. If stealing a loaf of bread, say, prevents your family from dying of starvation, then the annoyance of the shopkeeper is likely to be outweighed by your happiness that your family is still alive. So stealing the bread is morally permissible.

Kant's deontological ethics takes a rule-based approach. According to Kant, we have a *duty* to follow moral laws that are universal. He provides two tests to determine what these laws are. Based on these tests, Kant would say stealing is always wrong, regardless of the consequences.

Aristotle's virtue theory takes a different approach to both Kant and utilitarianism. Kant and utilitarianism both give formulas for what good *actions* are, whereas Aristotle is more concerned with what a good *person* is. So, if a virtuous person would not steal in a particular set of circumstances, then it is wrong to steal in those circumstances.

	Utilitarianism	Kant's deontological ethics	Aristotle's virtue ethics
Summary	Maximise pleasure, minimise pain	Act according to rules you could wish everybody follow	Virtues are character traits that help one act morally
Problems	Difficult to calculate	Tweaking the maxim	No clear guidance
	Ignores moral status of relationships (e.g. family)	Ignores consequences	Circularity
	Tyranny of the majority	Other valuable motivations	Competing virtues
	Ignores preferences	Conflicts between duties	Difference between eudaimonia and moral good
	Doctrine of swine	Non-hypothetical imperatives	

UTILITARIANISM

Utilitarian ethical theories are *consequentialist*. This means they say that it's the consequences of an action that make it either right or wrong. The most obviously relevant consequences are pain and pleasure. So, generally speaking, utilitarian theories say we should act to maximise pleasure and minimise pain.

There are three different versions of utilitarianism on the syllabus: act, rule, and preference utilitarianism.

Act utilitarianism

The simplest form of utilitarianism is act utilitarianism. Act utilitarianism says that in each situation we should choose the action that maximises happiness. Jeremy Bentham (1748-1832) is generally regarded as the creator of act utilitarianism and he says:

> *"The greatest happiness of the greatest number is the foundation of morals and legislation."*

Bentham's utilitarianism is also known as *hedonistic* or *quantitative* utilitarianism. It says morally good actions are those that maximise happiness/pleasure (and minimise pain) in each specific instance – there are no rules or moral values beyond this.

So, if we want to know what to do in a given situation, we should compare all the possible courses of action and choose the one that maximises happiness. Happiness is the *only* good, and so it doesn't matter if you're lying or stealing or killing as long as that action results in the greatest happiness.

For example, an act utilitarian would argue that it is morally right for a poor person to steal from a rich person because the money would cause more happiness for the poor person than it would cause unhappiness for the rich person. Similarly, an act utilitarian would argue that murder is justified if your victim is himself a murderer who would kill 10 people in the future.

The felicific calculus

Bentham's act utilitarianism is called *quantitative* utilitarianism because it's about *quantifying* happiness – adding up all the happiness and subtracting all the pain – and then deciding how to act based on the numbers. He provides the "felicific calculus" as a way to calculate utility in this way.

The felicific calculus has seven variables:

- Intensity: how strong the pleasure is
- Duration: how long the pleasure lasts

- Certainty: how likely the pleasure is to occur
- Propinquity: how soon the pleasure will occur
- Fecundity: how likely the pleasure will lead to more pleasure
- Purity: how likely the pleasure will lead to pain
- Extent: the number of people affected

So, for example, if two different courses of action lead to two different intensities of pleasure, then (all else being equal) the ethically right course of action is the one that leads to the more intense pleasure. Likewise, the calculus dictates that we should prefer a longer-lasting pleasure than a shorter one. It gets complicated, though, when comparing e.g. a more intense but short-lasting pleasure vs. a less intense but longer lasting pleasure (see below).

Anyway, the felicific calculus should (in theory) provide a means to *quantify* total happiness: add up all the pleasures and subtract all the pains. Act utilitarians would agree that acting morally is about making this number as big as possible.

Responses to act utilitarianism

Difficult to calculate

Bentham's felicific calculus seems impractically complicated to use every single time you have to make a decision.

For example, how are we supposed to quantify each of the seven variables? Are we supposed to hook everyone up to brain scanners to measure how intense their happiness is? This is obviously not possible. But even if we could objectively measure the intensity of happiness, how do we then compare this against the other six variables? For example, how do we decide between a longer-lasting dull pleasure and a short-lived but more intense pleasure?

Utilitarianism demands we make such calculations every time we act. Such calculations seem, if not impossible, a highly impractical way to decide how to act.

The moral status of particular relationships

Certain people – namely, friends and family – are more important to us than others. But act utilitarianism is concerned only with the greatest good for the greatest number. This means, then, that there are no grounds to justify acting to maximise the happiness of the people you know and care about over random people on the street.

This means, for example:

- The £10 you spent buying your mum a birthday present made her happy, sure, but it would have made Joe Bloggs in Mozambique happier. So buying your Mum

a birthday present was morally wrong according to utilitarianism, because it didn't maximise happiness.

- The time you spent with your friends made them happy, but volunteering at the local soup kitchen would have increased the greatest good for the greatest number more effectively. So, you acted wrongly by spending time with your friends because this did not maximise the greatest happiness for the greatest number.

You get the idea. If we sincerely followed act utilitarianism to the letter then we would never be morally permitted to spend time or money on our loved ones.

These examples can be developed into two possible objections:

- That utilitarianism is too idealistic and doesn't work in practice.
- That certain relationships such as family, friends, and the local community have a unique moral status (for example, that parents have a moral duty to look after their children before anyone else) but that act utilitarianism forces us to ignore these moral obligations.

Tyranny of the majority

There are some things that just seem inherently wrong regardless of the consequences. However, if it makes enough people happy, act utilitarianism would have to say such things are morally correct.

For example:

- Imagine a society of 1 million people where 95% of the population enjoy seeing innocent people tortured.
- In such a society, act utilitarianism would justify taking an innocent person and torturing them on live TV or something for the amusement of the 95% of people who enjoy that.
- Obviously torturing 1 innocent person causes that individual a great deal of unhappiness, but his unhappiness is massively outweighed by the happiness of the 950,000 people watching on TV.
- So, not only would utilitarianism not condemn torturing this innocent person – utilitarianism would say it's wrong *not to* torture him! In other words, torturing the innocent person is a *good* thing because it creates the greatest happiness for the greatest number.
- But this is obviously false – it's *just wrong* to torture innocent people, regardless of the consequences.

*Possible **response** to the claim that utilitarianism leads to the tyranny of the majority:*

Although this is a consequence of *act* utilitarianism, *rule* utilitarianism could argue that the rule "don't punish innocent people" leads to greater happiness than not following this rule (see page 94).

Ignores preferences (Nozick's experience machine)

There are situations where we might *prefer* something even if it makes us less happy, and situations where we might prefer something *not* happen even though it would make us more happy.

Philosopher Robert Nozick's experience machine thought experiment illustrates this point: Imagine you could be plugged into a virtual reality machine that simulates the experience of a perfect life. In other words, the machine *maximises your happiness and minimises your pain*. Once plugged in, you don't know you're in a virtual reality and you believe your perfect life is completely real.

Despite maximising happiness, many people would prefer *not* to enter the experience machine. These people would prefer to live a real life and be in contact with reality even though a real life means less happiness and more pain compared to the experience machine.

According to act utilitarianism, everyone should enter the experience machine (because all that matters is maximising happiness). But it seems morally wrong to ignore a person's preferences and, for example, force them to enter the experience machine in order to increase happiness.

This example illustrates a problem with Bentham and Mill's *hedonism* (the idea that happiness and pleasure are the only things of value). We realise there are things in life more important than simple pleasure – such as being in contact with reality – but act utilitarianism ignores our preferences for these things.

*Possible **response** to the claim that utilitarianism ignores preferences:*

Preference utilitarianism would argue that we should respect people's preferences even if doing so does not maximise happiness (see page 95).

Other utilitarian philosophies

In addition to act utilitarianism above, the syllabus also lists two other forms of utilitarianism:

- Rule utilitarianism: We should follow *general rules* that maximise happiness (even if there are specific instances where following these rules does not maximise happiness)

- Preference utilitarianism: We should act to maximise people's *preferences* (even if these preferences don't maximise pleasure and minimise pain)

John Stuart Mill: *Utilitarianism*

John Stuart Mill (1806-1873) was a student of Jeremy Bentham and utilitarianism. But Mill rejected Bentham's *quantitative* approach in favour of a more *qualitative* approach and made arguments for a sort of *rule* utilitarianism rather than *act* utilitarianism. Mill also provided a metaethical "proof" of utilitarianism:

Mill's "proof" of utilitarianism

Mill argues that happiness is not only a good thing, it is the *only* good. Other things we desire – e.g. truth and freedom – are part of what happiness is to us. Mill provides the following argument as "proof" of these claims:

1. The only proof that something is desirable is that people desire it
2. No proof can be given why the general happiness is desirable other than that each person desires their own happiness
3. The fact that happiness is *desirable* is "all the proof the case admits of" that happiness is a *good* thing
4. All our other values (e.g. truth, freedom, dignity) constitute what makes us happy – we only value them *because* they increase happiness
    - In other words, satisfying these other values are part of what it means to be happy
5. So, not only is happiness good, it is the *only* good

*Possible **response** to Mill's proof of utilitarianism:*

G.E. Moore argues that Mill's proof of utilitarianism commits the fallacy of equivocation (particularly in premise 3 above) in equating happiness with good. The two things may be *correlated*, but this does not show they are the *same* thing (see the naturalistic fallacy on page 114 for more detail).

Higher and lower pleasures (qualitative hedonism)

Earlier, we saw how Bentham's felicific calculus seeks to *quantify* happiness (page 89). However, critics of this purely quantitative approach argue that this makes utilitarianism a "doctrine of swine" in that it reduces the value of human life to the same simple pleasures felt by pigs and other animals.

In response to this objection, Mill rejects Bentham's felicific calculus and argues that not all pleasures and pains are equally valuable. Mill argues that there is a *qualitative*

difference between the "higher" pleasures of thought, feeling, and imagination compared to the "lower" pleasures of the body and the senses.

Mill argues that people who have experienced both higher and lower pleasures always prefer higher pleasures. Compare, for example:

- The guilty dopamine hit of eating junk food or mindlessly watching YouTube videos and playing video games (lower pleasures)
- The deep sense of accomplishment after completing some creative project you've been working on for months or the meaningful happiness that comes from spending quality time with loved ones (higher pleasures)

Mill argues that humans always prefer higher pleasures over lower pleasures because they value *dignity* – and dignity is an important component of happiness (see Mill's proof of utilitarianism above). So, returning to the "doctrine of swine" objection, Mill famously says:

> *"It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied."*

Rule utilitarianism

Rule utilitarianism focuses on the consequences of *general rules* rather than the consequences of specific actions. So, whereas act utilitarianism argues we should act to maximise happiness *in each individual case*, rule utilitarianism argues that we should act according to *general rules* that maximise happiness (even if following these rules doesn't lead to maximum happiness in every instance).

Some philosophers interpret John Stuart Mill as a rule utilitarian. For example, Mill says:

> *"though the consequences in the particular case might be beneficial—it would be unworthy of an intelligent agent not to be consciously aware that the action is of a class which, if practiced generally, would be generally injurious, and that this is the ground of the obligation to abstain from it."*

So, for example, there may be instances where stealing would lead to greater happiness than not stealing. However, if people were always allowed to steal in such instances, it would cause a constant level of anxiety and uncertainty that you might get your property stolen. This constant anxiety would be "generally injurious" and result in less happiness overall. Thus, rule utilitarianism could argue that we should follow the general rule "don't steal".

This sort of reasoning provides a possible response to the tyranny of the majority objection described above (page 91). For example, rule utilitarianism could argue that living in a society where you knew innocent people were regularly tortured would lead to

greater unhappiness overall. Sure, the majority might enjoy seeing other people tortured on live TV, but they would also live in constant fear that they could be the next innocent victim chosen for torture!

*Possible **response** to rule utilitarianism: rule fetishism*

Although rule utilitarianism may avoid some of the morally undesirable outcomes of act utilitarianism, placing too much emphasis on rules can lead to a different set of morally undesirable outcomes. For example, is it really sensible to say "never steal" or "never lie" even when doing so saves someone's life? Rule fetishism – the stringent adherence to rules simply because they are rules – leads to morally questionable outcomes. This objection is also discussed in relation to Kant's deontological ethics (see page 99).

Preference utilitarianism

Preference utilitarianism is a *non-hedonistic* form of utilitarianism. This means that instead of maximising happiness (hedonistic utilitarianism), we should act to maximise people's *preferences*.

This provides a response to the experience machine objection to act utilitarianism (page 92). Act utilitarianism says we should shove everyone into the experience machine – whether they want to go in or not – because doing so would maximise their happiness. However, preference utilitarianism can reject this by saying we should respect people's preference to live in the real world (even if living in the real world means less happiness).

A related example would be carrying out the wishes of the dead. It can't increase the happiness of a deceased person to carry out their will (because they're dead). However, if a deceased person expressed a preference for their money to be donated to the local cat shelter, say, then it seems there is a moral obligation to honour this preference. Act utilitarianism, though, would say we should ignore this preference and just spend the money in whichever way maximises happiness. But it seems wrong to just ignore people's preferences like this. Preference utilitarianism, though, can avoid this problem and say we have a moral obligation to respect the *preferences* of the dead.

Preference utilitarianism can also tie in with Mill's distinction between higher and lower pleasures (page 93). Mill claims that higher pleasures are just inherently more valuable than lower pleasures, but preference utilitarianism can explain this in terms of preference: we *prefer* higher pleasures over lower pleasures, and so should seek to maximise those.

KANT'S DEONTOLOGICAL ETHICS

In a sentence, Immanuel Kant's (1724-1804) ethics says: "treat others how you would like to be treated yourself".

However, it gets a bit more complicated than that. Kant's theory is quite long-winded and has several definitions to remember, but the key points are:

- The only thing that is good "without qualification" is *good will*
- Good will means *acting for the sake of duty*
- You have a duty to follow the moral law
- Moral laws are *universal*
- You can tell if a maxim (rule) is universal if it passes the *categorical imperative*
    - Two tests for whether a maxim passes the categorical imperative:
        - It must not lead to a contradiction in conception
        - It must not lead to a contradiction in will
- Finally, Kant also says never to treat people as means to an end (the humanity formula)

The good will

Good will means acting for the sake of duty. Acting for the sake of duty, according to Kant, is the only thing that is good "without qualification".

So, if you save someone's life because you expect to be financially rewarded, this action has no moral worth. You're acting for selfish reasons, not for the sake of duty. However, if you recognise your duty to save someone's life and act because of this duty then that choice *does* have moral worth – regardless of whether it achieves the desired outcome.

Duty

Deontology (as in Kant's deontological ethics) is the study of duty.

Kant argues that we each have a duty to follow the moral law. The moral law, according to Kant, is summarised by the categorical imperative.

The categorical imperative

Kant describes the categorical imperative as follows:

> *"Act only according to that maxim whereby you can at the same time will that it should become a universal law without contradiction."*

Maxims are basically rules or laws. So, in plain English, Kant is saying: only follow rules that you could wish *everyone* would follow.

Kant gives two ways to test whether you can wish your maxim be followed universally: contradiction in conception and contradiction in will.

Test 1: Contradiction in conception

For a law to be universal, it must not result in a *contradiction in conception*. A law leads to a contradiction in conception if it would somehow be self-contradictory for everyone to follow it.

For example, Kant would say the maxim "it's OK to steal" leads to a contradiction in conception if applied universally. The reason for this is as follows:

- If stealing was always OK then you could just take whatever you wanted from someone whenever you wanted
- But if you could just take other people's stuff whenever you wanted, the concept of ownership wouldn't exist because everyone would have just as much "right" to an object as the "owner" did
- But if the concept of ownership doesn't exist, then stealing wouldn't even be possible! After all, you can't steal something from someone if they don't own it in the first place!
- So, if it's *always* OK to steal, then it isn't even *possible* to steal
- So, the maxim "it's OK to steal" leads to a contradiction in conception when applied universally

Kant says that if a maxim leads to a contradiction in conception, you have a *perfect duty* not to follow that maxim. So, going back to the example above, Kant would say we have a perfect duty not to steal.

Test 2: Contradiction in will

Assuming the maxim does not result in a contradiction in conception, we must then ask whether the maxim results in a *contradiction in will*. In simple English, a contradiction in will basically means to expect everyone else to follow a rule that you yourself don't.

For example, there is no contradiction *in conception* (see above) if everyone followed the maxim "don't help others in need". The idea of a world where no one helps anyone else may be unpleasant, but it is not self-contradictory.

However, there is a contradiction *in will* with this maxim. The reason for this is that there are times where we are in need and want other people to help us. So, we can't rationally

will "don't help others in need", because it contradicts our own desire to receive help when we need it.

The way Kant actually phrases this point is in terms of "means" and "ends". Kant says we will certain *ends* (i.e. goals) and that part of willing these ends means also willing the *means* to achieve these ends. Returning to the example of "don't help others in need", this leads to a contradiction in will because there are possible scenarios where it is impossible to achieve our ends without the means of help from others. We can't rationally will our goal/end without also willing the *means* to achieve that goal – and help from others may be the only possible means to achieve it.

Kant says that if a maxim leads to a contradiction in will, then we have an *imperfect duty* not to follow that maxim.

The humanity formula

Kant also gives another formulation of the categorical imperative, called the humanity formula:

> *"Act in such a way that you always treat humanity [...] never simply as a means, but always at the same time as an end."*

Treating someone as a means to your own end means to use them. So, Kant is basically saying don't use people.

An example of using someone as a means to an end would be tricking someone into marrying you to get their money. If you pretend to love someone to marry them and take their money, then you are treating them as a *means* to make money, which is wrong. Instead, Kant says, we should always treat people as ends in themselves.

According to Kant, it's the deception that is the problem in the marrying for money example above. The reason the deception is a problem is that it undermines the rational agency of the other party. By withholding your true intentions, you prevent the other party from rationally pursuing their own ends (e.g. to find a loving partner).

However, if you were honest with the other party, they are able to make an informed choice on whether or not this fits with their ends. In principle, Kant would say there's nothing wrong with marrying someone for money – as long as both parties are *aware* this is what is going on (and so both parties are able to rationally pursue their own goals). For example, the other person's goal might be to get married to *anyone*, regardless of whether it's true love or not. In this case, both parties can use each other for mutual benefit and thus (rationally) pursue their own ends.

Responses to Kant's deontological ethics

Tweaking the maxim to justify any course of action

Kant argues that ignoring a perfect duty leads to a contradiction in conception. As we saw in the stealing example (page 97), the very concept of private property couldn't exist if stealing was universally permissible. So, the maxim "it's OK to steal" leads to a contradiction in conception.

However, by tweaking the maxim slightly, we can avoid this contradiction in conception. For example, instead of my maxim being "it's OK to steal", I could claim my maxim is "to steal *when your name is Christian and it's Wednesday*" or "to steal *from stores that begin with the letter A*".

Both of these maxims *can* be universalised without leading to a contradiction in conception. They would apply rarely enough that there would be no breakdown in the concept of private property. And so, by defining maxims cleverly, it seems possible to justify *any* course of action using the categorical imperative.

Possible ***response*** *to the objection that we can avoid duties by tweaking our maxim:*

Kant would likely argue that modifying your maxim in this way is cheating because the extra conditions – such as my name or the name of the store – are irrelevant to this situation. Kant would argue that the categorical imperative is concerned with the *actual* maxim I am actually acting on and not some arbitrary one I just made up. If we apply the categorical imperative honestly to our true maxim, which is "it's OK to steal", it would lead to a contradiction in conception.

Ignores consequences

There is a strong (utilitarian) intuition that consequences are important when it comes to moral decision making. This intuition is drawn out by ethical dilemmas such as the trolley problem:



Many people would have the utilitarian intuition that it is morally acceptable to murder 1 person if doing so somehow saved 1000 people. However, Kant says murder is wrong (except in self defence).

Another example is stealing. Many people would have the utilitarian intuition that it's morally acceptable to steal food in some situations – for example, stealing food to save your starving family's life. However, Kant says stealing is *always* wrong, and so you should just let your family starve to death. However, there is a strong utilitarian intuition that Kant's response here is wrong.

And again, a similar outcome is discussed in the lies section of applied ethics (page 135). Kant argues that we have a perfect duty not to lie – even if telling a lie would save someone's life.

These are all examples where consequences seemingly make it morally justifiable – morally good, even – to go against the categorical imperative. These examples show that following Kant's ethics too strictly leads to morally absurd outcomes that ignore the greater good.

*Possible **response** to the objection that consequences are important to morality:*

In response to these examples, Kant would insist that we still have a moral duty to follow the categorical imperative even if this leads to undesirable consequences such as unhappiness or even loss of life.

Kant would insist that the good will is the only thing that is good without qualification and that consequences (such as happiness) are not always good without qualification. For example, happiness derived from making others unhappy or from torturing animals isn't good. In contrast, the good will (i.e. choosing your actions for the sake of duty) is always good without qualification. Morality is not about anticipating the consequences of our actions, it is only about following duty.

Ignores other valuable motivations

In the discussion of the good will (page 96), we saw how Kant argues that acting for the sake of duty is the only thing that is good without qualification. In other words, being motivated by duty is the source of moral worth.

But imagine a close friend is ill in hospital. You pay them a visit because you genuinely care about them and want to make sure they're OK. According to Kant, this motivation (concern for your friend) has no moral value because the only source of moral value is to be motivated by *duty*.

In contrast, if you had no genuine concern for your friend but begrudgingly went to visit them purely because you had a duty to do so, this *would* have moral value according to

Kant. But this seems backwards: surely it is better to go and see your friend because you genuinely care about them rather than just because you have a duty to!

Another example to illustrate this point is a father spending time with his children:

- Father A doesn't love his children but recognises he has a duty to spend time with them and so acts according to that duty
- Father B loves his children and sincerely enjoys spending time with them and so he does this regardless of whether or not this is his duty

According to Kant, father A is morally praiseworthy, but father B is not. Again, this seems backwards: surely it is better to spend time with your family because you *love* them and enjoy their company than to do so purely out of *duty!*

*Possible **response** to the objection that Kant's ethics ignores the value of other motivations:*

Kant would respond by making a distinction between acting *out of* duty and acting *in accordance with* duty. There is nothing wrong with being motivated by motivations such as love, but we shouldn't choose how to act *because* of them. We should always act *out of* duty, but sometimes what we want to do anyway is *in accordance with* our duty. In other words, we should always be motivated by duty but if our duty happens to coincide with other motivations such as love then that's a bonus.

Conflicts between duties

Kant argues that it's never acceptable to violate our duties – they are categorical. But this raises the question of what to do in cases of a conflict of duties.

For example, Kant would say we have a duty to never lie. But what happens if you make a promise to someone but then find yourself in a situation where the only way to keep that promise is by telling a lie? Whichever choice you make you will seemingly violate one of your duties.

*Possible **response** to the possibility of conflicting duties:*

Kant acknowledges the possibility of conflicting *obligations*. If someone makes a promise, they create an obligation for themself to keep that promise. But the *grounding* of an obligation someone creates *for himself* is weaker than the grounding of moral duty, which is categorical. So, in the example above, the moral duty never to lie overrules the obligation to keep a promise.

As for a true conflict of *duties*, Kant argues that such a thing is impossible. Our moral duties are objective and rational and so it is inconceivable that they could conflict with one another. If it *appears* that there is a conflict in our duties, it must mean we have made a mistake somewhere in formulating them. After all, you can't *rationally* will a

maxim to become a universal law if it conflicts with another law you rationally will – because that would be contradictory.

So, if we think through our duties carefully, Kant says, a true conflict is irrational and inconceivable. Applied to the example above, Kant could say we shouldn't make a promise that could conflict with our moral duties.

Foot: *Morality as a system of hypothetical imperatives*

Philippa Foot argues that moral imperatives are not *categorical* in the way Kant thinks they are.

To show why, we first need to understand the distinction between *categorical* and *hypothetical* imperatives:

Hypothetical imperatives	Categorical imperatives
Qualified by an "if" statement	Not qualified by an "if" statement, they apply universally
Examples: • "You should do your homework *if you want to do well in the exam*" • "You should leave now *if you want to catch the 2:30 train*"	Examples: • "You shouldn't steal" • "You shouldn't tell lies"

Foot describes various *non-hypothetical* (but not categorical) imperatives. For example, the rules of etiquette say "handshakes should be brief", but there is no hypothetical condition for this rule. You can try to make the rules of etiquette hypothetical, e.g. "handshakes should be brief *if you don't want to seem weird*", but we still wouldn't say that makes it OK for someone who doesn't care how he's perceived to hold people's hands for 5 minutes when he shakes hands. So, etiquette seems non-hypothetical.

Yet despite being a non-hypothetical imperative, Kant wouldn't say that "handshakes should be brief" is a *categorical* imperative in the same way that the laws of morality are. For one thing, Kant would say the rules of morality by themselves give reason to act a certain way, and thus apply to everyone and anyone all the time. For Kant, following the moral law is a matter of being rational. In contrast, you might not always have rational reason to follow the rules of etiquette (e.g. you don't have reason to follow the rules of etiquette if you don't care about looking weird).

However, Foot claims that Kant's categorical imperatives are no more categorical than the rules of etiquette. For example, there is nothing *irrational* (as Kant would claim)

about not following a rule like "don't steal" if you don't accept the rule in the first place. Similarly, it is not *irrational* to want other people to behave in a way that is different to how you yourself behave.

So, according to Foot, Kant's categorical imperatives are no more powerful than non-hypothetical imperatives such as the rules of etiquette. In other words, there's no *real* reason why we should follow categorical imperatives in the same way there's no real reason why you shouldn't carry on holding someone's hand after a handshake. The only difference, says Foot, is how we *feel* about each imperative: we just feel that the rules of morality are somehow inescapable in a way the rules of etiquette aren't.

ARISTOTLE'S VIRTUE ETHICS

The first thing to say about Aristotle's (384 BC–322 BC) virtue ethics is that it addresses a different question to Kant's deontological ethics and utilitarianism. Instead of approaching moral philosophy from the perspective of "what should I do?" *(action-centred)*, Aristotle addresses a question more like "what sort of person should I be?" *(agent-centred)*. So, instead of defining a good person as someone who does good actions, Aristotle would define good actions as those done by good people.

Aristotle makes a bunch of different points that can seem a bit unconnected at times. However, the key points to remember are:

- *Eudaimonia* = the good life for human beings
- The good life for a human being must consist of something *unique* to human beings
- Human beings are rational animals, and *reason* is their unique characteristic activity (their *ergon*)
    - So, the good life for humans (eudaimonia) will be one full of actions chosen according to reason
- *Virtues* are character traits that enable us to act according to reason
    - More specifically, virtues are the mid point between a vice of deficiency and a vice of excess
- We develop virtues through habit and training, similar to learning a skill

Eudaimonia

You can have good food, good friends, a good day. Aristotle's ethical enquiry is concerned with the good *life* for a human being. The word Aristotle uses for this is *eudaimonia,* which is sometimes translated as "human flourishing".

Aristotle has in mind the good life for a human in a broad sense. Eudaimonia is not just about following moral laws, or being happy, or being successful – it's about all these things together (and more). It's a good life in the *moral* sense as well as in the sense that it's the kind of desirable and enjoyable and successful life you would want for yourself.

Eudaimonia is a property of someone's life taken as a whole. It's not something you can have one day and then lose the next. Good people sometimes do bad things, but this doesn't make them bad people. Likewise, people who have good lives (eudaimons) can sometimes have bad days.

Aristotle says that eudaimonia is a *final end*. We don't try to achieve eudaimonia as a means to achieve some goal but instead it is something that is valuable for its own sake.

Ergon, arête, and virtue

While fleshing out this concept of eudaimonia, Aristotle uses the words arête and ergon. These roughly translate as:

- Ergon: The function or characteristic activity of a thing
- Arête: A property or virtue that enables a thing to achieve its ergon

For example, a knife's ergon is to cut things. And a good knife has the *arête* of sharpness because this enables it to cut things well.



A knife's ergon is to cut, and the arête of sharpness helps the knife fulfil its ergon

Aristotle argues that eudaimonia must consist of something unique to humans. The ergon of humans, says Aristotle, is to use *reason*. Reason is what makes us unique from trees, plants, books, knives, animals – everything else in the world.

However, this does not mean that we achieve eudaimonia by doing nothing but idly thinking and reasoning. Instead, Aristotle's claim is that humans always *choose their actions* for some reason – good or bad. So, what Aristotle actually says is that the good life for a human being (eudaimonia) is one full of actions chosen according to good reason.

Virtues are character traits that enable us to choose our actions according to good reason. So, just as the arête of sharpness enables a knife to fulfil its ergon to cut things, the arête of virtues help humans fulfil their ergon, which is to choose actions according to reason. A bit like eudaimonia, virtues are not something you have one day but not the next. If someone has a virtuous character but slips up one day and does something unvirtuous, this doesn't make them a bad person. Likewise, a bad person whose character is prone to vice doesn't suddenly develop virtuous character through committing one virtuous act. So, again, virtues are *character traits* – they are part of what we are. So, working backwards:

- Virtues are character traits that enable us to fulfil our function/ergon
- Our function/ergon is to use reason
- Choosing actions according to reason enables us to achieve eudaimonia (i.e. a good life)
- Eudaimonia is a final end (i.e. something that is good for its own sake)

The doctrine of the mean

Aristotle's doctrine of the mean (also called the golden mean) provides more detail about what virtues traits actually are. The doctrine of the mean says that virtues are the intermediate or average (the mean) between two extremes.

For example, if you never stand up for yourself then you are cowardly (vice of deficiency). But if you go too far the other way and start fights with anyone for the slightest reason then you are reckless (vice of excess). But if you have a courageous character, you will stand up for yourself when it is appropriate to do so. So, courage is a virtuous character trait because it is in between the two extremes.

Some other examples:

Vice of deficiency	Virtue	Vice of excess
Cowardice	Courage	Recklessness
Shy	Modest	Shameless
Stingy	Liberal	Wasteful
Self-denial	Temperance	Self-indulgence
Surly	Friendly	Obsequious
Unirascible	Good-tempered	Short-tempered
Boorish	Witty	Buffoonish

Again, virtues are *character traits*. So, for example, just because you have a good-tempered character *in general* this doesn't mean you should always be good-tempered *in every situation*.

There are times when anger is the appropriate (and virtuous) response. The virtuous person is not someone who *never* feels angry or other extremes of emotions. Instead, the virtuous person is someone whose character disposes them to feel these extreme emotions *when it is appropriate to do so*.

As Aristotle says:

> *"For instance, both fear and confidence and appetite and anger... may be felt both too much and too little, and in both cases not well; but to feel them at the right times, with reference to the right objects, towards the right people, with the right motive, and in the right way, is what is both intermediate and best, and this is characteristic of virtue."*

The skill analogy

Acquiring virtues is somewhat analogous to acquiring *skills* such as learning to ride a bike or play the piano:

- Nobody is born knowing how to play the piano, but we are born with the *capacity* to know how to play the piano. Likewise, nobody is born virtuous, but they have the capacity to become virtuous
- You don't learn to play the piano by simply reading books and just studying the theory, you have to actually *do it*. Likewise, it's not enough to just read and learn about virtue, you have to actually act virtuously until it becomes part of your character.
- When you first start learning to play the piano, you follow the rules and try not to press the wrong keys – but you don't really understand what you're doing. In the case of virtue, we start by teaching children rules for behaviour (e.g. "don't hit people" or "if you haven't got anything nice to say, don't say anything at all") and they just follow these rules because they're told to – not because they understand why.
- But as you progress with playing the piano, you become able to play automatically without thinking and, eventually, you might become so comfortable playing the piano that you're able to improvise and understand what sounds good, what doesn't, and why. Likewise, by following rules for acting virtuously, we begin to do so automatically (i.e. it becomes part of our character). Further, we begin to understand what virtue is and this enables us to improvise according to what the situation demands.

Phronesis

Another of Aristotle's terms is *phronesis*. Phronesis translates as "practical wisdom" and basically means a general understanding of good that enables someone to think through, understand, and act according to what is good.

In Aristotle's words:

> *"Practical wisdom, then, must be a reasoned and true state of capacity to act with regard to human goods."*

The components of phronesis are:

- A general idea of what is good and bad as relates to eudaimonia
- The ability to apply this general idea to specific situations and "deliberate" (i.e. think through) what is required in terms of attitude and action
- The ability to act on this deliberation

Phronesis is not the sort of thing you can learn from books – it's *practical* wisdom. Real life is complicated and nuanced and so it's not possible for a set of rules (e.g. utilitarianism or Kant's deontological ethics) to capture what good is in every situation. Instead, it takes *experience* to develop a practical wisdom that understands what is good and is able to apply it to the particular situation.

Voluntary and involuntary action

Aristotle says we should only praise or condemn actions if they are done *voluntarily*. In other words, you can't criticise someone for acting unvirtuously if their actions weren't freely chosen.

Aristotle makes a distinction between *voluntary* and *involuntary* actions and says that we should only judge a person's voluntary actions. He makes a further distinction between two different kinds of involuntary action:

- Voluntary: Acting with full knowledge and intention
- Involuntary:
  - Compulsion: Being *forced* to do something you don't want to do – e.g. sailors throwing goods overboard to save the boat during a storm
  - Ignorance: Doing something you don't want to do by *accident* - e.g. slipping on a banana skin and spilling a drink on someone

Responses to Aristotle's virtue ethics

No clear guidance

Aristotle describes virtues in the middle of the two extremes (the doctrine of the mean) and says that this varies depending on the situation. But this isn't very helpful as a practical guide of what to do.

For example, Aristotle would say it is correct to act angrily sometimes – but when exactly? And how angry are you supposed to get before it crosses over from a virtue to a vice of excess?

Kant gives the categorical imperative as a test to say whether an action is moral or not. And even utilitarianism has the felicific calculus. But with Aristotle, we have no such criteria against which to judge whether one course of action is better than another. The doctrine of the mean doesn't give actual quantities, only vague descriptions such as "not too much" and "not too little". If you genuinely don't know what the correct course of action is, virtue theory doesn't provide any actual guidance for how to act.

*Possible **response** to the objection that virtue ethics provides no clear guidance:*

Aristotle could reply that virtue theory was never intended to provide a set of rules for how to act. Life is complicated – that's the whole reason why you need to develop phronesis (page 107) in the first place: so you can act virtuously in the many complicated situations that arise.

Plus, we can still *reflect* whether an action is, for example, courageous or stupid. We could also ask questions like "how could I be more friendly in this situation?" that help us decide how to act. Just because virtue theory doesn't provide a specific course of action, that does not mean it provides no guidance whatsoever.

Circularity

Aristotle can be interpreted as defining virtuous acts and virtuous people in terms of each other, which is circular. This interpretation says Aristotle is arguing something like:

- A virtuous act is something a virtuous person would do
- And a virtuous person is a person who does virtuous acts

Which doesn't really say anything at all. These descriptions are circular and so say nothing meaningful about what a virtuous person or a virtuous act actually is.

*Possible **response** to the objection that virtue ethics is circular:*

The virtue ethicist can reply, however, that Aristotle's definitions are not circular. Aristotle describes the virtuous person in terms of *eudaimonia* (page 104), which is in turn described in terms of, for example, the ergon of reason and choosing actions according

to reason. Virtues, according to Aristotle, are character traits that enable us to fulfil our ergon achieve eudaimonia. So, in short, a virtuous person is ultimately defined in terms of eudaimonia and so the definition is not circular.

Competing virtues

We can imagine scenarios where applying two different virtues (e.g. justice and mercy) would suggest two different courses of action.

For example, if you're a judge and someone has stolen something, you have to choose between the virtue of justice (i.e. punishing the criminal) and the virtue of mercy (i.e. letting the criminal go). You can't choose to do both things, so whichever choice you make will be unvirtuous in some way.

*Possible **response** to the issue of competing virtues:*

Aristotle would reply that such conflicts between virtues are impossible. As mentioned in the no clear guidance objection, virtues are not rigid and unbreakable rules and the correct virtue and in what amount depends on the circumstances. Aristotle would say that phronesis (page 107) would mean knowing what each virtue tells you to do and in what amount according to the particular circumstances of the case. So, for example, the virtuous judge could sentence a person according to justice, but show appropriate mercy if there are extenuating circumstances. For example, the judge might give a lesser sentence to a thief who stole to feed his starving family compared to a thief who stole purely for greed.

Difference between moral good and eudaimonia

According to Aristotle, the good life for a human being is eudaimonia. And, as mentioned, eudaimonia includes many elements beyond simply being moral – such as honour, wealth, and happiness.

However, we often make a distinction between a good life for me (eudaimonia) and a *morally* good life. For example, imagine a nurse who spends her entire life saving lives in some remote country. She doesn't enjoy her work, but does it because she believes it's needed. She's constantly stressed and dies at age 30 from a virus caught while carrying out her work.

With such an example, we tend to have a strong intuition that this nurse's life is *morally* good (she's done nothing but help other people) but she clearly did not achieve *eudaimonia*. Such examples can be used to make either of the following objections:

- That there is a difference between what is morally good and eudaimonia, and so Aristotle's theory fails as an account of moral philosophy
- That Aristotle's theory is only concerned with self interest

*Possible **response** to the issue of the difference between good and eudaimonia:*

In response to the first point, we can argue that Aristotle was never trying to answer the (narrow) question of what a *morally* good life is. Aristotle's inquiry was concerned with the good life *in general*.

But further – and this provides a response to the second point – Aristotle would likely argue that being virtuous and achieving eudaimonia does involve some level of commitment to others. For example, Aristotle would say that being fair, kind, and generous are all important virtues and so helping others and being altruistic is a component of eudaimonia. However, Aristotle would say altruism is not the *only* component of a good life. Or, to put it another way, being morally good is necessary to achieve eudaimonia but not sufficient.

METAETHICS

OVERVIEW

Metaethics is about what moral judgements – e.g. "murder is wrong" – mean and what (if anything) makes them true or false. These are two separate topics, but they overlap significantly. The main debate is about whether mind-independent moral properties, such as goodness and badness, exist:

- **Moral realism:** There are mind-independent moral properties and facts
- **Moral anti-realism:** There is no such thing as mind-independent moral properties or moral facts

But there is also a second (related) debate about what moral judgements mean:

- **Cognitivism:** Moral judgements express cognitive mental states – i.e. beliefs – *that can be true or false*
- **Non-cognitivism:** Moral judgements express *non-cognitive* mental states, which are *not capable of being true or false*

The A-level philosophy syllabus covers five specific metaethical theories:

- Moral realist theories:
  1. **Naturalism (cognitivist):** "Murder is wrong" is true if the act of murder has the *natural* property of wrongness (e.g. because it causes *pain*, which is a natural property)
  2. **Non-naturalism (cognitivist):** "Murder is wrong" is true if the act of murder has the *non-natural* property of wrongness
- Moral anti-realist theories:
  3. **Error theory (cognitivist):** "Murder is wrong" is false because the property of wrongness *doesn't exist* ("murder is good" is also false, because the property of goodness doesn't exist either)
  - Non-cognitivist theories:
    4. **Emotivism:** "Murder is wrong" just means "Boo! Murder!" and so is not capable of being true or false
    5. **Prescriptivism:** "Murder is wrong" means "Don't murder people!" and so is not capable of being true or false

The arguments for the theories are often linked with each other. For example, the arguments against naturalism are used as arguments for non-naturalism, and the arguments against non-naturalism are used as arguments for error theory.

And because each theory fits into these two categories (cognitivism/non-cognitivism and realism/anti-realism), you can argue against that theory by attacking either category. For example, error theory is both cognitivist *and* anti-realist, so you can argue against error theory by arguing for non-cognitivism or realism (or both). Similarly, because emotivism is both non-cognitivist and anti-realist, you can argue against emotivism by arguing for realism or cognitivism (or both) – and so on.

MORAL REALISM

Realist metaethical theories argue that mind independent moral properties (e.g. goodness, badness, rightness, wrongness) and facts (e.g. "murder is wrong") exist. In other words, "good" is a mind-independent property like "square" or "green" or "over 2 metres tall".

So, when we make moral judgements – e.g. "murder is wrong" – moral realism says these statements are literally true or false in the same way "grass is green" is literally true. Murder has the (mind-independent) property of wrongness in the same way grass has the property of greenness.

But while moral realists all agree that mind-independent moral properties exist, they can disagree about what *kind* of properties they are:

- Moral naturalism says moral properties are *natural* properties
- Moral non-naturalism says moral properties are *non-natural* (think non-physical) properties

Both moral naturalism and moral non-naturalism are *cognitivist* theories: they both agree that when people express moral judgements they express cognitive beliefs that are capable of being true or false.

Moral naturalism

Moral naturalism says that moral properties exist and that these moral properties are natural (think physical) properties. Or, in other words, moral naturalism says moral properties can be *reduced* to natural properties. These natural properties exist independently of minds and so moral naturalism is a moral realist theory.

So, according to naturalism, "murder is wrong" expresses a belief that murder is wrong – where "wrong" refers to some natural property.

Utilitarian moral naturalism

The most obvious example of moral naturalism is utilitarianism (page 89). Utilitarians like Mill and Bentham argue that "good" can be reduced to pleasure, and "bad" can be reduced to pain. Pain and pleasure are *natural* properties of the world/our minds: there is nothing spooky or non-natural about them. So, utilitarianism is a form of moral naturalism.

Mill's proof of utilitarianism (page 93) is an example of a *metaethical* argument for utilitarianism. As a quick reminder, the key points of Mill's argument are:

- People desire their own happiness, so *individual* happiness is desirable.
- This is reason to believe the *general* happiness is desirable too.
- The fact that the general happiness is desirable is the only possible proof we could give that the general happiness is *good*
- Further, happiness is the *only* good, because all other values are just part of what makes us happy

The conclusion is a metaethical claim: that "good" *reduces* to happiness. And since happiness is a *natural* property, it follows that "good" reduces to a natural property.

*Possible **response** to Mill's proof of utilitarianism:*

See the naturalistic fallacy below (page 114).

Virtue ethics moral naturalism

Although utilitarianism is the obvious form of moral naturalism, some philosophers see virtue ethics as a form of moral naturalism too.

For example, Aristotle's discussion of ergon/function (page 104) can be interpreted as a discussion of *natural* facts about human beings. We might argue that it is a natural fact that the function of human beings is to use reason – in the same way it is a natural fact that the function of a knife is to cut things. There is nothing spooky or non-natural about claiming that the function of a knife is to cut things, and likewise there is nothing spooky or non-natural about claiming that the function of human beings is to use reason. So, there are natural facts about function. Similarly, there are natural facts about traits (virtues) that enable things to perform their function well.

So, on this reading, "good" reduces to a set of natural facts about function and performing that function well. Something is good if it performs its function well, and virtues are good because they enable a thing to perform its function well. For example,

being courageous is good because being courageous helps humans fulfil their function (to act correctly according to reason).

Responses to moral naturalism

- The naturalistic fallacy (see page 114)
- The is-ought problem (see page 124)
- The verification principle (see page 121)

Moral non-naturalism

Moral non-naturalism says that moral properties exist and that they are *non-natural* properties. So, whereas moral naturalism says moral properties can be reduced to natural properties, non-naturalism says that moral properties are *basic:* they cannot be reduced to anything simpler. But these basic moral properties exist independently of our minds and so moral non-naturalism is a moral realist theory.

So, according to non-naturalism, "murder is wrong" expresses a cognitive belief that murder is wrong – where "wrong" refers to a non-natural property.

G.E. Moore: *Principia Ethica*

The naturalistic fallacy

G.E. Moore (1873-1958) argues for moral non-naturalism. His book, *Principia Ethica*, begins with criticisms of moral naturalism. Moore coins the term "naturalistic fallacy" to describe the fallacy (i.e. bad reasoning) of equating goodness with some natural property, as Mill does in his "proof" of utilitarianism (page 93).

For example, Moore would say it is a fallacy to conclude that drinking beer is *good* from the fact that drinking beer is *pleasurable* because goodness and pleasure are two completely different kinds of properties; good is a *moral* property, whereas pleasure is a *natural* property. Even if goodness and pleasure often accompany each other (e.g. doing good acts causes people to feel pleasure) it does not and could not follow that goodness and pleasure are the same thing. No matter how closely they are correlated, moral properties are a completely different thing to natural properties.

So, Moore argues that moral properties – such as "good" and "bad" – cannot be reduced to natural properties . In fact, moral properties can't be reduced at all. They are basic.

Note: Moore's naturalistic fallacy is very similar to Hume's is/ought problem (page 124). However, Hume takes the problem to show that moral properties don't exist (moral anti-realism), whereas Moore intends to show that moral properties do exist (moral realism) but that they cannot be reduced to anything simpler, such as natural properties.

The open question argument

Moore's second argument against moral naturalism is the open question argument. Moore argues that it is an open question whether "pleasure" and "good" are the same thing:

- Closed question: "Is good good?" or "is pleasure pleasure?"
- Open question: "Is pleasure good?"

Moore argues that if goodness and pleasure really were the same thing (as naturalism claims), it would be a closed question to ask "is pleasure good?". In other words, Moore is arguing that if naturalism was true and "good" meant the same thing as "pleasure", it wouldn't make sense to ask "is pleasure good?" because it would be like asking "is pleasure pleasure?".

*Possible **response** to the open question argument:*

Moore's argument here clearly doesn't work because there are plenty of examples where two things are in fact the same thing but it is still an open question whether they are.

For example, "water" and "H2O" refer to the same thing, but it is still an open question to ask "is water H2O?" So, it's possible that "pleasure" and "good" could refer to the same thing and yet it still be an open question to ask "is pleasure good?"

Intuitionism

Non-naturalism claims moral properties are non-natural properties, but this raises the question of how we acquire *knowledge* of them. Moore's answer to this question is intuitionism.

Naturalism can explain our knowledge of moral properties as similar to knowledge of other natural properties. For example, there is nothing spooky about how we know "torturing people causes pain" because "pain" is a natural property. And so, if pain is bad (as utilitarian naturalism claims), then there's nothing spooky about how we can know "torturing people is bad" - you have a natural explanation for natural knowledge. But non-naturalism cannot make this same move: knowledge of non-natural properties is difficult to explain because non-natural moral properties are not like ordinary physical properties (e.g. pain, happiness, blueness, roundness, or largeness) that can be perceived and measured.

Moore's answer to this issue is *intuitionism*. He argues that, via a faculty of rational intuition, we can directly reflect on the truth of moral judgements such as "murder is wrong". The truth or falsehood of such moral judgements is said to be self evident because of this intuition.

*Possible **response** to intuitionism:*

As part of his epistemic argument from queerness (page 118) J.L. Mackie rejects Moore's intuitionism specifically. Mackie rejects intuitionism as a "lame answer" to the question of how we acquire knowledge of moral properties because it doesn't really explain anything. Mackie argues that to say we acquire knowledge of moral properties via intuition is just to say it is "utterly different from our ordinary ways of knowing everything else."

Responses to moral non-naturalism

- Argument from queerness (see page 118)
- The verification principle (see page 121)

Responses to moral realism

- Argument from relativity (see page 117)
- Hume's arguments against cognitivism (see page 123)
- The is/ought problem (see page 124)
- The verification principle (see page 121)

MORAL ANTI-REALISM

Anti-realist metaethical theories argue that mind-independent moral properties (e.g. goodness, badness, rightness, wrongness) do not exist. It follows from this that there are no such things as true moral facts (e.g. "murder is wrong"). The syllabus looks at three anti-realist metaethical theories:

- Error theory (cognitivist) says moral judgements are cognitive statements but mind-independent moral properties don't exist
- Non cognitivist theories:
    - Emotivism says moral judgements are non-cognitive statements that express feelings of approval or disapproval
    - Prescriptivism says moral judgements are non-cognitive statements that are intended as instructions as well as expressions of approval/disapproval

Error theory

Error theory says that when we make moral judgements we express cognitive (i.e. true or false) beliefs about the external world. However, because error theory claims that moral properties *don't exist,* all these moral judgements are false.

So, according to error theory, "murder is wrong" expresses a cognitive belief that murder has the property of wrongness - but "wrongness" doesn't exist and so the statement is false. Similarly, "murder is good" is *also* false, because "goodness" doesn't exist either. Again: error theory says *all* moral judgements are false.

J.L. Mackie: *Ethics: Inventing Right and Wrong*

Error theory = cognitivism + anti-realism. And so, J.L. Mackie's (1917-1981) case for error theory has two separate parts: arguments for cognitivism and arguments for anti-realism.

Mackie's arguments for anti-realism are the most important here. He argues that differences in moral beliefs across cultures suggest mind-independent moral properties don't exist, and that mind-independent moral properties would have to be unlike any other mind-independent properties we know of.

Arguments for cognitivism

Mackie begins by defending a cognitivist interpretation of moral language. He provides various examples and arguments to suggest that moral discussion in both philosophical contexts and ordinary language is cognitivist - i.e. that it aims to make true/false claims about the world.

In *philosophical* contexts, Mackie argues that moral philosophy throughout history has tended to assume objective moral values. Plato and Kant, for example, didn't argue that morality was subjective of a matter of feeling - they argued that certain moral claims are *objectively* true or false.

Further, Mackie argues, our *ordinary* language assumes cognitivism as well. To illustrate this, Mackie uses the example of someone facing the moral dilemma of whether or not to take a job undertaking research related to bacteriological warfare. In trying to resolve a moral dilemma like this, you don't ask how you *feel* about it - you ask whether such action is wrong *in itself.*

Argument from relativity

Having established that moral judgements are cognitive, Mackie then argues for moral anti-realism – i.e. that mind-independent moral properties do not exist.

Mackie's first argument for anti-realism starts by pointing out that there are variations in moral beliefs between cultures (and between the same culture at different time periods). For example:

- Some cultures are polygamous, other cultures think this is wrong
- Some cultures think eating certain animals is wrong, other cultures eat those animals
- Historically, many cultures have kept slaves, but most modern cultures think this is wrong

If moral realism is correct, there would only be one objectively correct answer to all of these issues. Why, then, is there so much disagreement on these issues? If there is one true answer, why is it that some cultures grasp the moral truth but others just can't?

Compare moral knowledge with knowledge of basic scientific and mathematical truths, such as "1+1=2". Every culture would agree this is true. The different cultures might have different words or learn it at different times, but they all arrive at *the same* belief on this issue – because "1+1=2" is objectively true.

This convergence does not appear to happen with *moral* belief, though. Cultures that are independent of each other tend to form independent moral practices and beliefs. We can explain this disagreement in one of two ways:

1. One culture has, for some reason, discovered objective moral reality while the other hasn't
2. Each culture has different conditions and a different way of life and has developed their own moral beliefs in response to that

Mackie's argument is that 2 is the more plausible account. If moral realism were true, you wouldn't expect to see such divergent moral beliefs. Or, in other words, if there really were mind-independent moral properties and facts, you would expect every culture to discover them in a similar way to how every culture has discovered objective truths such as "1+1=2".

Arguments from queerness

Mackie then provides a further argument that alleged moral properties are so strange (or "queer" to use Mackie's term) that we should reject their existence. In particular, Mackie is attacking the idea of *non-natural* moral properties, such as those posited by Moore (page 114). Mackie argues that such non-natural moral properties would be:

- *Metaphysically* queer: If mind-independent moral properties exist, they would need to have "intrinsic prescriptivity". For example, "good" things would need to somehow have *to-be-doneness* built into them and "bad" things would have *not-*

to-be-doneness built into them. But it's not clear how this is possible (scientifically, metaphysically, or otherwise).

- *Epistemically* queer: If mind-independent moral properties exist, then it is a total mystery how we would acquire knowledge of them. Whereas *natural* knowledge can be explained *naturally*, moral knowledge can't be explained in the same way and instead requires spooky hypotheses such as Moore's intuitionism.

Metaphysically, moral properties would be unlike any other properties we know of (in science or otherwise). If we are to believe that certain actions have the mind-independent moral property of "good" – such as giving to charity – then those actions must somehow have *to-be-doneness* built into them, Mackie argues. But how is this possible? Similarly, if something had the property of "bad", then that thing would have *not-to-be-doneness* built into it. But, again, that just sounds weird. It seems impossible to make sense of how ordinary, objective, natural properties of the world could relate to subjective human motivations in this way. It doesn't make sense how a physical object or action could *in itself* demand we act a certain way.

Mackie also questions how moral properties relate to natural properties. For example, how does the *natural* fact that someone stole something relate to the *moral* fact that what they did was wrong? Imagine someone stealing from a shop, and then imagine someone stealing from a shop *and that this action has the property of wrongness.* What's the difference between the two cases? What does the property of *wrongness* add to the situation that wasn't already there?

Epistemically, moral properties would also be weird – it's a total mystery how we could know about them. Whereas knowledge of *natural* properties and facts can be explained *naturally*, knowledge of *moral* properties seems somehow spooky. How do we know that stealing is wrong, for example? Or that helping others is good? Do we *see* wrongness? Or *touch* goodness? Obviously not.

In short, Mackie says that the queerness of alleged moral properties – both metaphysical and epistemic – is evidence that moral properties do not actually exist.

Responses to error theory

- Moral naturalism (see page 112)
- Intuitionism (see page 115)
- Non-cognitivism (see below)
- Moral nihilism (see page 127)
- Moral progress (see page 128)

Non-cognitivism

If you remember from the overview (page 111), non-cognitivists believe that moral judgements such as "murder is wrong" express *non-cognitive* mental states. Non-cognitive statements do not aim to describe reality and so are not supposed to be taken as either true or false.

Cognitive statements	Non-cognitive statements
Aim to literally describe how the world is	Do not aim to literally describe how the world is
Are true or false	Are neither true or false
Examples: • "Water boils at 100°c" • "Triangles have 3 sides" • "The Eiffel Tower is in Paris" • "The Eiffel Tower is in Berlin"	Examples: • "Ouch!" • "Boo!" • "Hooray!" • "Don't do that!"

It follows that all non-cognitivist metaethical theories are also *anti-realist* theories: if moral judgements do not aim to express true or false statements, then there is no need for mind-independent moral properties that would make those statements true or false.

The syllabus lists two specific non-cognitivist metaethical theories: emotivism and prescriptivism.

Emotivism

Emotivism says that moral judgements express feelings and emotions of approval and disapproval.

For example, when a referee gives a bad decision during a football match and the crowd shout "Booooo!", that "Booooo!" statement is clearly neither true or false – it is simply an *expression* of the crowd's negative feelings towards the referee's decision. Thus, "Booooo!" is a non-cognitive statement.

According to emotivism, it's a similar story with moral judgements. When someone says "murder is wrong!", what they really mean is something like "Boo! Murder!" And likewise, when someone says "giving money to charity is good", what they are really expressing is something like "Hooray for giving money to charity!".

A.J. Ayer: Verification principle

A.J. Ayer (1910-1989) was an emotivist, arguing that:

> *"If I say to someone, "you acted wrongly in stealing that money" […] I am simply evincing my moral disapproval of it. It is as if I had said, "You stole that money," in a peculiar tone of horror."*

But the key argument on the syllabus – Ayer's verification principle – can serve as an argument for moral non-cognitivism in general.

Ayer's verification principle says cognitive (/meaningful) statements must be either:

- An analytic truth (e.g. "a triangle has 3 sides"), or
- Empirically verifiable (e.g. "water boils at 100°c")

Any statement that does not fit either of these descriptions is non-cognitive (/meaningless) according to the verification principle. And Ayer argues that *moral judgements* fail the verification principle.

Firstly, "murder is wrong" is clearly not an analytic truth because there is no *logical* contradiction in the statement "murder is *not* wrong" – it's not like a married bachelor or a 4-sided triangle.

Secondly, Ayer also argues that "murder is wrong" is not empirically verifiable either – both on the naturalist and non-naturalist interpretations:

- Naturalism would argue that we could prove that murder causes pain, anger, etc. However, Ayer argues that this is not the same as proving murder is *wrong* (see Moore's naturalistic fallacy on page 114 for a similar argument). Hence, Ayer rejects naturalism: we can empirically verify that murder causes *pain*, say, but we cannot empirically verify that murder is *wrong*.
- Ayer also argues that there is no way to empirically verify the presence of *non-natural* properties. Even if the act of murder did possess the non-natural property of wrongness, how could we ever *prove* this? You can't see or test for the non-natural properties, and so their existence is not empirically verifiable. So, Ayer also argues against non-naturalism too: the existence of non-natural properties cannot be empirically proven.

So, moral judgements are not analytic truths, nor are they empirically verifiable. Therefore, according to the verification principle, moral judgements are non-cognitive/meaningless.

*Possible **response #1** to the verification principle: reject the verification principle*

We could respond that Ayer's verification principle fails its own test. The claim that "a statement is only meaningful if it is analytic or empirically verifiable" is *itself* neither an

analytic truth or empirically verifiable! Therefore, according to its own criteria, the verification principle is meaningless.

However, Ayer would likely respond that the verification principle was never intended as an empirical hypothesis or typical cognitive statement. Instead, the verification principle is a *definition* of what a meaningful/cognitive statement is. But we do not necessarily have to accept this definition as correct: we may argue against Ayer that there are, in fact, meaningful cognitive statements that are neither analytic truths nor empirically verifiable. And we might further argue that moral judgements are examples of such cognitive statements.

*Possible **response #2** to the verification principle: accept the terms of the verification principle*

Alternatively, we can accept the verification principle as the criteria for cognitive/meaningful statements and argue that moral judgements are either analytically true or empirically verifiable. The former is a difficult argument to make, but you could argue that Mill's proof of utilitarianism (page 93) is empirical *verification* of the moral judgement that happiness is good.

Hume: *Treatise of Human Nature*

Hume has been interpreted as an emotivist like Ayer. Like Ayer, Hume's main arguments are for moral non-cognitivism, but Hume also makes claims to the effect that moral judgements express emotions.

Hume's fork (again)

If you remember from epistemology, Hume's fork (page 68) argued there are just two types of *judgements of reason*:

- Relations of ideas
- Matters of fact

A judgement of reason is effectively the same thing as a *cognitive* judgement (the labels cognitivism and non-cognitivism didn't really exist in Hume's day). And so, Hume is saying that anything that is not a judgement of reason – i.e. a relation of ideas or a matter of fact – is a *non-cognitive* statement.

Firstly, Hume argues, moral judgements are not *relations of ideas*. Remember, to deny a relation of ideas leads to a contradiction, like saying "triangles do not have 3 sides". But there is no contradiction in saying "murder is not wrong", for example, and so the moral judgement "murder is wrong" cannot be a relation of ideas.

Further, Hume argues, moral judgements are not *matters of fact* either:

> *Take any action [said] to be vicious: Wilful murder, for instance. Examine it in all lights, and see if you can find that matter of fact, or real existence, which you call vice. In which-ever way you take it, you find only certain passions, motives, volitions and thoughts. There is no other matter of fact in the case. The vice entirely escapes you, as long as you consider the object. You never can find it, till you turn your reflection into your own breast, and find a sentiment of disapprobation, which arises in you, towards this action. Here is a matter of fact; but 'tis the object of feeling, not of reason. It lies in yourself, not in the object."*

Matters of fact are true in virtue of experience. But here, Hume is saying we have no experience of wrongness (vice) *in the object itself*. Instead, the only experience we have is of certain attitudes or feelings *within ourselves*.

So, Hume's argument here is *very* similar to Ayer's verification principle above (page 121), except it uses different words. The argument can be summarised as:

- There are only two types of judgements of reason *(i.e. two types of cognitive judgements):*
  - Relations of ideas *(i.e. analytic truths)*
  - Matters of fact *(i.e. empirically verifiable)*
- Moral judgements are not relations of ideas
- Moral judgements are not matters of fact
- Therefore, moral judgements are not judgements of reason *(i.e. they are non-cognitive)*

Again, if moral judgements are not judgements of reason, then they are not cognitive judgements. So, Hume's fork supports non-cognitivism in the same way Ayer's verification principle does.

*Possible **response** to Hume's fork:*

The responses to Hume's fork are very similar to the responses to Ayer's verification principle above: we can either reject the fork itself, or accept the fork and argue that moral judgements are either relations of ideas or matters of fact.

Moral judgements motivate action

According to Hume, judgements of reason/cognitive judgements don't motivate us to act in any way. For example, my cognitive belief that "today is Monday" – by itself – doesn't motivate me to act in any way. It's simply a belief I have. Instead, says Hume, it's *emotions* and *desires* that motivate us to act. For example, my *desire* not to be late to Monday's philosophy class might motivate me to leave the house and walk to class.

According to Hume, moral judgements seem more like the walking to class example because they motivate action. For example, my belief that "murder is wrong" will motivate me not to murder, or my belief that "giving money to charity is good" will motivate me to give money to charity.

So, Hume's argument here is essentially:

- Moral judgements motivate action
- Judgements of reason (i.e. cognitive judgements) cannot motivate action
- Therefore, moral judgements are not judgements of reason (i.e. cognitive judgements)

If moral judgements are not judgements of reason, then they are *non-cognitive* judgements. So, this argument further supports non-cognitivism.

*Possible **response** to the argument judgements motivate action:*

We might respond by rejecting Hume's claim that moral judgements motivate action. For example, I could argue that my (cognitive) belief that "murder is wrong" does not – by itself – motivate me to act in any particular way. Instead, it is my *desire* to be a good person that motivates me to not murder people.

In other words, my motivation to act morally can be explained by the cognitive moral judgement that something is bad + the non-cognitive *desire* not to be a bad person.

The is/ought problem

Hume makes a distinction between two types of claim: statements about what is and statements about what ought to be:

Claims about what *is*	**Claims about what *ought to be***
Factual claims about what is the case	Value judgements about what is good and bad
Examples: • "That is an act of torture" • "Smith murdered Jones"	Examples: • "You ought not torture" • "Smith shouldn't have murdered Jones"

Hume argues that there is a *gap* between the two kinds of claim: that you cannot logically derive ought statements like "you ought not torture" from statements about what is, such as "that is an act of torture".

For example, to say "you shouldn't torture animals because it hurts them!" derives a statement about what *ought* to be (you shouldn't torture animals) from a statement about what *is* (torturing animals causes pain). But Hume says this does not – *and could*

not – follow. You cannot argue from "is" to "ought" because they are two completely different kinds of claims – like comparing apples and oranges.

We can use this distinction to support non-cognitivism: We might argue that the *reason* we cannot derive "ought" statements from "is" statements is because the former type of statement is non-cognitive while the latter is cognitive. "Is" statements like "Smith murdered Jones" are capable of being true or false, whereas "ought" statements like "Smith shouldn't have done that" are expressions of emotion that are not capable of being true or false.

*Possible **response** to the is/ought problem:*

We may respond that there are examples of valid moral reasoning that suggest you *can* derive value judgements from factual judgements. For example:

1. If murder is wrong, then you shouldn't murder people
2. Murder is wrong
3. Therefore, you shouldn't murder people

If this argument is sound – and it appears to be – then this suggests that the value judgement in premise 2 is capable of being true or false and thus that moral reasoning from "is" to "ought" is possible (this argument for cognitivism is discussed further on page 126).

Prescriptivism

Prescriptivism says moral judgements express *prescriptive* instructions that aim to guide behaviour.

An example of a prescriptive statement in ordinary life would be something like "shut the door". The statement "shut the door" is clearly neither true or false – it is simply an *instruction* to do something. Thus, prescriptive statements such as "shut the door" are also non-cognitive statements.

According to prescriptivism, it's a similar story with moral judgements. When someone says "murder is wrong!", what they (primarily) mean is something like "don't murder people!"

R.M. Hare: *The Language of Morals*

R.M. Hare (1919-2002) agrees with emotivism (page 120) that moral judgements express (non-cognitive) attitudes. But Hare argues this misses the *force* of moral judgements. Although there is some sense in which moral judgements express emotion, the *main* point is to guide conduct. For example, "stealing is wrong" implies the *imperative* "don't steal".

As well as the above analysis of moral judgements, Hare provides an analysis of general value terms such as "good", "bad", "right", and "wrong". Hare argues that the meaning of these terms is not simply descriptive but that they also *commend* or *criticise*. Hare uses an example of a "good" strawberry to illustrate this: A purely *descriptive* analysis of "good strawberry" might reduce its meaning to "sweet and juicy strawberry". But this is clearly not the *only* thing I mean when I say "this is a good strawberry" because there are ways in which we use language that conflict with this analysis. For example, I might say "this is a good strawberry *because* it is sweet and juicy" – but this statement wouldn't make sense on the purely descriptive analysis because it would be like saying "this is a sweet and juicy strawberry because it is sweet and juicy". So, according to Hare, "good strawberry" does not simply describe, it also *commends* the strawberry.

But in order to commend (or criticise) something, we must *assume* a certain set of *standards*. In the example above, the standards against which I commended the strawberry were "sweet and juicy". However, these standards are not objective – I simply assume that's what "good" means for strawberries. There are no facts I can point to that determine that set of standards is correct or incorrect.

So, returning back to *moral* value judgements, these work in a similar way to the strawberry: When I say "she is a *good* person" I am assuming a certain set of moral standards (in this case some standards for how *people* should be) and commending that person against those standards. Again, there are no *facts* that determine which set of moral standards are correct, and so Hare's position here is anti-realist.

According to Hare, commendation or criticism against a set of standards is the primary meaning of "good" and "bad" respectively. Such commendation or criticism is intended to provide (imperative) guidance on how others should act and be.

Responses to non-cognitivism

Note: The following arguments apply to non-cognitivist theories (i.e. prescriptivism and emotivism) but not cognitivist theories (e.g. error theory, naturalism, and non-naturalism)

Moral argument and reasoning

Non-cognitivism appears to be at odds with how we typically use moral judgements because we often use moral judgements as part of moral *reasoning*. For example:

1. If murder is wrong, then paying to have someone murdered is wrong
2. Murder is wrong
3. Therefore, paying to have someone murdered is wrong

This seems like a *valid* argument - the conclusion seems to follow logically from the premises.

But if non-cognitivism is correct, it's hard to see how this is a valid argument at all. It seems the meaning of "murder is wrong" changes between the first and second premises:

- In premise 2, "murder is wrong" is by itself and so, according to non-cognitivism, should be understood as a non-cognitive statement.
- But in premise 1, "murder is wrong" is presented as a *cognitive* statement. It's presented as something that, *if true*, implies something else. However, according to non-cognitivism, "murder is wrong" *cannot be true or false* because it is a non-cognitive statement.

The problem for non-cognitivism, then, is that people regularly "embed" moral judgements in statements like in premise 1 above – they say things like "if murder is wrong then" and "if stealing is bad then". But if non-cognitivism is correct, it's hard to make sense of why people do this. Why would people embed moral judgements in sentences like this if moral judgements were incapable of being true or false?

When moral judgements are embedded in the context of moral argument and reasoning, it seems obvious that moral judgements are intended as *cognitive* statements rather than non-cognitive statements.

Mackie's arguments against non-cognitivism

As mentioned earlier, Mackie provides further examples of how both ordinary and philosophical moral discussion is cognitivist (see page 117 for more details).

Responses to moral anti-realism

Note: The following problems apply to all anti-realist theories (i.e. error theory, prescriptivism, and emotivism)

Moral nihilism

If moral anti-realism is true, it can be argued that this leads to moral nihilism: the view that no actions are inherently right or wrong. Moral anti-realism means there's nothing true about moral judgements such as "murder is wrong", which raises the question of why anyone should bother to be "moral" at all.

*Possible **response #1** to the claim that moral anti-realism leads to moral nihilism: there are non-cognitive reasons to behave "morally"*

Non-cognitivists can respond that just because there's no inherent right or wrong, people still have moral attitudes and feelings. And the realisation that moral values are just expressions of feelings doesn't mean we should (or could) stop having these moral

feelings. It's also somewhat self-defeating to be a moral nihilist according to non-cognitivism. After all, living as if there are no moral values is itself an expression of a certain attitude or feeling.

*Possible **response #2** to the claim that moral anti-realism leads to moral nihilism: there are practical reasons to behave "morally"*

Cognitivist anti-realist theories (i.e. error theory) have a harder time responding to the charge of moral nihilism. One response could be to just accept the charge of moral nihilism and argue that, though undesirable, this doesn't make error theory any less true. There may also be practical reasons to behave *as though* some moral judgements are true. For example, if you were always stealing from your friends, chances are they wouldn't remain friends with you for very long.

*Possible **response #3** to the claim that moral anti-realism leads to moral nihilism: so what?*

Finally, the anti-realist could simply accept that moral anti-realism leads to nihilism and ask: *so what?* We may not *want* moral nihilism to be true – it may be an undesirable consequence – but just because we don't *want* moral nihilism to be true, that doesn't prove that anti-realism is actually false.

Moral progress

Our moral values have changed over time. For example, it was considered morally acceptable to keep slaves back in the time of Plato but it's not today. If we accept that such changes are examples of *moral progress*, then we can make this argument:

1. If moral anti-realism is true, then there would be no moral progress
2. But there has been moral progress
3. Therefore, moral anti-realism is false

*Possible **response** to the moral progress argument:*

Moral anti-realism can reply that this argument begs the question – i.e. it assumes its conclusion in its premises. To say there has been moral progress (premise 2) assumes that moral anti-realism is false – it assumes there are objective moral facts against which we can say our moral knowledge has progressed or not. But why should the anti-realist accept that there's been objective moral progress when he doesn't accept the existence of objective morality in the first place?

However, the moral realist could potentially define moral progress in less question-begging ways. For example, you could say morality has become more *logically consistent*.

APPLIED ETHICS

OVERVIEW

Applied ethics takes the ethical theories we studied previously and applies them to practical moral issues. The A-level philosophy syllabus also mentions applying metaethical theories to these issues too.

The syllabus looks at four issues to apply ethics to:

1. **Stealing**
2. **Simulated killing**
3. **Eating animals**
4. **Telling lies**

STEALING

It's pretty obvious what stealing is. But to frame it in a philosophical way, people often say that individuals have property *rights* – i.e. that they have rights over certain things. To steal is to violate these property rights.

Utilitarianism and stealing

For act utilitarianism, whether or not it is acceptable to steal something will depend on the situation. There is no moral right to property over and above its utilitarian benefits and so if an act of stealing results in greater happiness then it would be morally acceptable to steal.

For example, act utilitarianism would say it is acceptable for a starving person to steal food if it saves their life, because the victim's loss is outweighed by the thief's benefit. Similarly, an act utilitarian could argue that it's morally acceptable for very poor people to steal from very rich people (like Robin Hood), because the rich person's loss of happiness is insignificant compared to the thief's gain of happiness.

However, rule utilitarianism could argue that although there may be *specific instances* where stealing leads to greater happiness, having a *general rule* of "don't steal" leads to greater happiness overall. John Stuart Mill makes a similar argument in his discussion of

justice and property rights: A society that permitted stealing would be one in which everyone would live in constant fear of being robbed. Without property rights, people would feel insecure that they would be stolen from. This distrust and fear would lead to a less happy society than one in which stealing isn't allowed, and so a rule utilitarian could argue that we should never violate a person's property rights – even in situations where doing so would maximise happiness.

Kant's deontological ethics and stealing

Kant would argue that it is never acceptable to steal for the following reasons:

- The categorical imperative says: "act only according to maxims you can will would become a universal law"
- My maxim is: "to steal"
- If I will stealing to be a universal law, then anyone could steal whenever they wanted
- But if anyone could steal whenever they wanted, the very concept of personal property wouldn't exist (because if anyone is entitled to just take my property from me in what sense is it mine?)
- And if there is no such thing as personal property, the very concept of stealing doesn't make sense (because you can't steal something from someone if it isn't theirs to begin with)
- So, willing that everyone followed the maxim "to steal" leads to a contradiction in conception
- So, we have a perfect duty not to steal
- So, stealing is (always) wrong

Aristotle's virtue ethics and stealing

Aristotle says there are some actions that *never* fall within the golden mean – and stealing is one of them. According to Aristotle, stealing is an injustice because it deprives a person what is justly and fairly theirs.

Even in extreme cases, such as stealing £1 from a billionaire to buy bread to save a starving child who will otherwise die, Aristotle would probably still say that stealing is wrong. Aristotle makes a distinction between unjust *actions* and unjust *states of affairs*. A starving child is an example of an unjust state of affairs – an unfortunate situation – but that's just the way the world is sometimes. According to Aristotle, it is much worse to deliberately and freely choose to commit unjust actions – even if you are doing so to counteract unjust states of affairs.

That said, virtue ethics philosophers don't have to agree with Aristotle on *everything*. You could perhaps disagree with Aristotle that stealing never falls within the golden mean and argue that there are extreme cases where stealing becomes morally acceptable. Knowing when would require practical wisdom.

Metaethics and stealing

- Moral realism:
  - Naturalism: "Stealing is wrong" is true if stealing has the *natural* property of wrongness (e.g. because it causes people to be upset)
  - Non-naturalism: "Stealing is wrong" is true if stealing has the *non-natural* property of wrongness
- Moral anti-realism:
  - Error theory: "Stealing is wrong" is false because the property of wrongness *doesn't exist*
  - Non-cognitivism:
    - Emotivism: "Stealing is wrong" just means "Boo! Stealing!" and so is not capable of being true or false
    - Prescriptivism: "Stealing is wrong" means "Don't steal!" and so is not capable of being true or false

SIMULATED KILLING

Simulated killing is about fictional death and murder, such as in video games and films. It's not about actually killing people (which is more obviously wrong).

This might seem like a non-issue at first – how can just watching or pretending to kill someone be bad? But there are all sorts of moral dimensions you can discuss. For example:

- The difference between watching a killing (e.g. in a film) and playing the role of the killer (e.g. in a video game)
- The effects simulated killing has on a person's character (e.g. whether exposure to simulated killing makes them more violent)
- Whether simulated killing is wrong in itself (in the UK, for example, video games involving rape and paedophilia are illegal even though they're just simulations – why not murder too?)

Utilitarianism and simulated killing

The obvious response of act utilitarianism would be that simulated killing is morally acceptable. After all, the person watching the film or playing the video game gets some pleasure from the simulated killing, and the person being killed doesn't actually suffer because it's only a simulation. So, in this situation, simulated killing results in a net gain of happiness.

But from a wider perspective, there are ways simulated killing could possibly decrease happiness. For example, if exposure to simulated killing makes a person more likely to kill someone *for real,* then maybe this pain would outweigh the happiness? Maybe simulated killing makes people more violent in general?

There are all sorts of studies on this topic, often with conflicting conclusions. If nothing else, these considerations support the difficult to apply objection to utilitarianism. However, if there was an obvious and irrefutable study that showed simulated killing makes people significantly more likely to murder in real life, then *rule* utilitarianism could make the case that simulated killing is wrong.

Kant's deontological ethics and simulated killing

Kant would most likely have no major objection to simulated killing. Murdering people in video games does not lead to a contradiction in conception, or a contradiction in will, or violate the humanity formula. In other words, simulated killing does not go against the categorical imperative.

Aristotle's virtue ethics and simulated killing

A key idea of Aristotle's ethics is that virtues are *character traits*. According to Aristotle, being a good person is not just about knowing what the virtues are, it's about acting on them until they become automatic and part of your character.

So, Aristotle might argue that if someone spends a lot of time playing video games that involve simulated killing, they may develop bad habits and bad character. For example, repeatedly killing fictional innocent people in a game may make someone increasingly unkind or unjust.

On the other hand, Aristotle might argue that killing fictional people is not *actually* unjust or unkind. After all, they're not real, and so there's no real injustice. Doing unjust acts develops the vice of injustice, but simulated killing is technically not an unjust act.

Like with any application of Aristotle's ethics, you also have to consider the context – there are no "one size fits all" rules. A virtuous person might partake in simulated killing occasionally as a form of entertainment or because they enjoy the competitive challenge

of gaming. In that context, simulated killing might not be unvirtuous. But doing nothing with your life except killing people in video games just because you love killing people is probably not virtuous.

Metaethics and simulated killing

- Moral realism:
  - Naturalism: "Simulated killing is wrong" is true if simulated killing has the *natural* property of wrongness
  - Non-naturalism: "Simulated killing is wrong" is true if simulated killing has the *non-natural* property of wrongness
- Moral anti-realism:
  - Error theory: "Simulated killing is wrong" is false because the property of wrongness *doesn't exist*
  - Non-cognitivism:
    - Emotivism: "Simulated killing is wrong" just means "Boo! Simulated killing!" and so is not capable of being true or false
    - Prescriptivism: "Simulated killing is wrong" means "Don't do simulated killing!" and so is not capable of being true or false

EATING ANIMALS

Utilitarianism and eating animals

A conclusion of Bentham's view that happiness is the only good is that utilitarian principles must be extended to animals (because animals can feel pleasure and pain just as humans can). So, Bentham says:

> *"The question is not, Can they reason?, nor Can they talk? but, Can they suffer? Why should the law refuse its protection to any sensitive being?"*

Peter Singer, another utilitarian philosopher, develops this line of thinking further. To privilege human pain and pleasure over animals is *speciesist*, he says (in a similar way to how privileging men over women, say, is *sexist).* So, utilitarians such as Bentham and Singer say that eating animals is wrong because the pain caused to the animals outweighs the pleasure of those who eat them.

There are, however, ways a utilitarianism could potentially justify eating animals.

According to utilitarianism, an action is good if it results in maximises pleasure and minimises pain. And you could argue that, if it wasn't for farming animals for food, many animals would never have existed and so would never have been able to experience pleasure and pain in the first place. So, if the animals farmed for food have an overall happy life and a painless death, then eating animals is morally justifiable because it results in a net increase of pleasure. An implication of this view is that farming conditions and practices are important: Farming animals in cramped or uncomfortable conditions where they have unhappy lives, say, would be morally wrong according to utilitarianism.

Kant's deontological ethics and eating animals

Kant's categorical imperative is only intended to apply to *rational beings*. Animals, Kant would say, do not have rational will and so are excluded from the categorical imperative.

There is seemingly no contradiction in conception and no contradiction in will that results from the maxim "it's OK to eat animals". Further, the humanity formula only says don't treat *humanity* as a means to an end because humans have a rational will and so are ends in themselves. But animals, Kant would say, do not have a rational will and so can be treated solely as means.

One potential issue this raises is the question of non-rational humans. Newborn babies and people with certain disabilities, for example, do not have a rational will. But surely Kant wouldn't want to say it's OK to eat such humans or treat them solely as a means?

Aristotle's virtue ethics and eating animals

Aristotle's discussion of eudaimonia is concerned with the good life for *humans* specifically. Animals, unlike humans, are not capable of reason and so eudaimonia doesn't apply to them. So, Aristotle would likely not see any issue with eating animals.

But some more recent virtue ethics philosophers take a different approach to this. Cora Diamond, for example, argues that there are different practices surrounding the eating of animals – some of which are virtuous and some of which aren't. Yes, animals are a different kind of being to humans and so we shouldn't treat their happiness as *equal* (as utilitarians like Singer say), but they are nevertheless living beings that can have good and bad lives. To completely ignore this, as some factory farming practices do, demonstrates the vices of callousness and selfishness. In contrast, rearing your own chickens and treating them humanely – even if you do ultimately eat them – demonstrates the virtues of sympathy and respect.

So, Diamond's virtue ethics could say that eating animals is sometimes acceptable – as long as it is done in the right way and for the right reason.

Metaethics and eating animals

- Moral realism:
  - Naturalism: "Eating animals is wrong" is true if eating animals has the *natural* property of wrongness (e.g. because wrongness is a natural property such as pain)
  - Non-naturalism: "Eating animals is wrong" is true if eating animals has the *non-natural* property of wrongness
- Moral anti-realism:
  - Error theory: "Eating animals is wrong" is false because the property of wrongness *doesn't exist*
  - Non-cognitivism:
    - Emotivism: "Eating animals is wrong" just means "Boo! Eating animals!" and so is not capable of being true or false
    - Prescriptivism: "Eating animals is wrong" means "Don't eat animals!" and so is not capable of being true or false

TELLING LIES

Utilitarianism and telling lies

As always, whether or not act utilitarianism would say it's morally acceptable to lie will depend on the situation. If telling a lie leads to greater happiness, then act utilitarianism would say you should lie. For example, if someone asks you whether they look good but they don't, the utilitarian thing to do is to lie and say "yes" because telling the truth would make the other person unhappy.

But *rule* utilitarianism could argue that a rule to "never lie" would lead to greater happiness than a rule that allows everyone to lie. For example, if everyone was an act utilitarian and always lied to increase happiness/reduce pain, then nobody could trust anything anyone said. And the rule utilitarian could argue that such a society – where no one can trust anyone else's word – would be less happy overall.

Kant's deontological ethics and telling lies

The point of telling a lie is to get the other person to believe something false. But if everyone always told lies, then people wouldn't believe each other. Given this, Kant

would argue that the maxim "it's OK to lie" fails the first test of the categorical imperative because it would lead to a contradiction in conception. If it was always acceptable to lie, the very concept of telling a lie (i.e. saying something is false in order to deceive someone into believing it's true) wouldn't make sense. So, according to the categorical imperative, we have a perfect duty to always tell the truth.

But French philosopher Benjamin Constant challenged Kant's approach of radical honesty by asking whether you should tell a known murderer the location of his victim when asked. Say, for example, a person is escaping an axe-murderer and you let them hide in your house. Shortly after, a crazy looking guy with an axe and blood-stained clothes asks you "where is he?" According to Kant, you should tell the truth: "he's in here." Telling the truth in this situation is seemingly the wrong thing to do, though, which suggests Kant's ethics is a poor account of morality.

Kant's response to this argument is to stick to his guns: you should not lie to the murderer *even to save a life.* One reason for this is that Kant says the moral worth of actions is determined by whether they are done for the sake of *duty*, not the *consequences*. Further, it's impossible to anticipate the consequences of our actions. However, Kant argues that if we do go against duty and try to anticipate the consequences of our actions, then we are then at least partially responsible for those consequences. So say, for example, you lie and tell the axe murderer that "he's next door" and, unbeknown to you the victim has sneaked out and really has gone next door, then you bear some responsibility for the death that results from not following your duty.

Finally, Kant would also argue that telling lies violates the humanity formula of the categorical imperative. To lie to someone is to treat them as a means to an end rather than an end in themselves. The example of marrying someone to get their money (page 98) illustrates this: if the person you lie to knew you were only *pretending* to love them, they might not choose to get married. And by lying about your true feelings, this person is unable to make an informed choice and rationally pursue their own choices.

Aristotle's virtue ethics and telling lies

In his discussion of the "social qualities" of truth and falsehood, Aristotle says:

> *"Falsehood is in itself bad and reprehensible, while the truth is a fine and praiseworthy thing; accordingly, the sincere man, who holds the mean position, is praiseworthy, while both the deceivers are to be censured"*

Aristotle is talking about lying about oneself: on one side *boasting* is a vice of excess, and on the other *false modesty* is a vice of deficiency. Telling the truth - i.e. "the sincere man" - is in the middle (i.e. the golden mean) and so is the virtuous action.

When Aristotle says "falsehood is in itself bad", he appears to be saying that lying is *always* wrong. However, Aristotle later describes *degrees* to which telling lies is bad: lying to protect your reputation, for example, is not as bad as lying to gain money. Given this, you could potentially argue that there may be situations where it is morally acceptable to lie, such as in the example of saving a life above. Again, as always, knowing when circumstances mean it's acceptable to lie will require practical wisdom/phronesis.

Metaethics and telling lies

- Moral realism:
  - Naturalism: "Telling lies is wrong" is true if telling lies has the *natural* property of wrongness
  - Non-naturalism: "Telling lies is wrong" is true if telling lies has the *non-natural* property of wrongness
- Moral anti-realism:
  - Error theory: "Telling lies is wrong" is false because the property of wrongness *doesn't exist*
  - Non-cognitivism:
    - Emotivism: "Telling lies is wrong" just means "Boo! Telling lies!" and so is not capable of being true or false
    - Prescriptivism: "Telling lies is wrong" means "Don't tell lies!" and so is not capable of being true or false

MORAL PHILOSOPHY – MODULE SUMMARY

Utilitarianism

- (Act) utilitarianism: Good = happiness, bad = pain. Morally right actions maximise happiness (also called quantitative hedonistic utilitarianism)
  - Arguments *against* act utilitarianism:
    - Difficult/impossible to calculate: Bentham's utility calculus is impractical. It is impossible to quantify, e.g. the intensity of pleasures but, even if we could do this, how do we compare intensity against the other six variables? For example, how do we decide between a more intense but shorter duration pleasure vs. less intense but longer duration pleasure?
    - Moral status of special relationships: Maximising the *general* happiness means we can't give preference to e.g. our family and friends over strangers. For example, utilitarianism implies that it would be immoral to spend £10 buying your brother a birthday present if that £10 could have made Joe Bloggs in Mozambique happier. But this seems a stupid outcome because we have moral obligations to our family and friends
    - Tyranny of the majority: E.g. killing an innocent person is clearly wrong, but act utilitarianism would say it's morally justified to kill an innocent person if doing so would please a large crowd, for example
      - Reply: *rule* utilitarianism: We should follow general rules that maximise happiness, even if they don't maximise happiness in every situation. For example, there may be individual situations where killing an innocent person results in the greatest happiness, but if we lived in a society where it was acceptable to kill innocent people sometimes, it would cause general anxiety that would result in a net decrease in happiness. So, we should follow the rule 'don't kill innocent people'
    - Ignores preferences: E.g. many people would prefer not to go into Nozick's experience machine even though it would maximise their happiness. But utilitarianism implies we should force them in anyway
      - Reply: *Preference* utilitarianism (also called non-hedonistic utilitarianism): we should act to maximise people's preferences, not pleasures

        - Utilitarianism ignores the moral integrity of the individual and is a "doctrine of swine" because it reduces the value of human life to the same base pleasures felt by pigs and other animals
            - Reply: Higher and lower pleasures (i.e. *qualitative* hedonistic utilitarianism): there is a qualitative difference between lower pleasures (i.e. bodily pleasures) vs. higher pleasures (e.g. love of family, creative pursuits, meaningful accomplishment). People who have experienced both higher and lower pleasures prefer higher pleasures because they have dignity (and dignity is part of happiness)

Kant's deontological ethics

- Overview:
    - The only thing that's good without qualification (i.e. good in itself) is *good will*
    - Good will = acting for the sake of *duty*
    - Humans have a duty to follow the *moral law*
    - Moral laws are *universal* laws
    - The *categorical imperative* is two tests to see if a law/maxim/action is universal:
        - Contradiction in conception: A law like "it's OK to steal" leads to a contradiction in conception because if we made it universal then it would always be OK to steal, but then it wouldn't make sense that someone could have private property to steal in the first place (because if everyone has a right to steal your stuff, then in what sense is it yours?). So, we have a *perfect duty* not to steal
        - Contradiction in will: A law like "don't help others in need" leads to a contradiction in will because there are cases when you want (will) other people to help you. So it's hypocritical (i.e. a contradiction in will) to act so that you never help others in need whilst also expecting other people to help you. So, we have an *imperfect duty* to help others in need
- Arguments *against* Kant's deontological ethics (and replies):
    - You can just say the maxim/law you are acting on is something else (e.g. "don't steal" becomes "don't steal unless it's Wednesday and your name is Christian") and that way you don't get a contradiction in conception

        - Reply: Be honest with yourself. You have to consider the *actual* maxim you are following, not some made up one
    - Conflicts between duties: E.g. what if there was a situation where you had to either lie or steal? Which do you choose?
        - Reply: Real conflicts between duties are impossible if you formulate duties carefully
    - Ignores consequences: E.g. a maniac chasing someone with a knife asks "which way did he go?" - telling a lie (not that bad) will avoid someone being murdered (very bad), but we can't do this if we follow Kant's theory because "don't lie" is a perfect duty
        - Reply: Yep, that's just how it goes. Acting for the sake of duty is the only thing that is good without qualification, consequences are not important
    - Ignores the value of other motivations: E.g. good will = acting for the sake of duty (see above). So it's better to go see your ill friend in hospital out of duty (even if you don't want to) than it is to go see them because you actually care about them
        - Reply: It's fine to be motivated by emotions. But we shouldn't choose how to act because of them. Our actions may be in accordance with duty but to be morally praiseworthy they need to be chosen because of duty

Aristotle's virtue ethics

- Overview:
    - Unlike the other two theories, virtue ethics focuses on being a good *person* in general (agent-centred) rather than looking at what correct *actions* are (action-centred)
    - Eudamonia = the good life for a human being
    - The good life for a human being must consist of some unique function (ergon = function) for humans
    - Human beings are rational animals, and reason is their unique function (i.e. reason is their ergon)
    - So, the good life (eudaimonia) is a life full of actions chosen according to (good) reason
    - Virtues are character traits that enable us to fulfil our ergon (they are

personality traits that help us choose actions according to good reason)
    - Golden mean: the virtue is the middle point between a vice of deficiency and a vice of excess
        - E.g. Cowardly | *Courageous* | Rash
    - Virtues are developed through habit and training (virtues are a kind of practical wisdom, or phronesis)
- Arguments *against* Aristotle's virtue ethics (and replies):
    - No clear guidance: How do we know, e.g., when too much courage becomes acting rashly? In a genuine moral dilemma (e.g. stealing to feed a hungry orphan or killing one person to save 20) Aristotle's theory isn't much help in telling us what is right
        - Reply: Virtue ethics was never intended as a set of *rules*. Anyway, we can still reflect e.g. "is this action courageous?" or kind or whatever which could help decide how to act
    - Conflicts between virtues: E.g. justice is a virtue, so is mercy. So if you're a judge and someone has stolen something, you have to choose between justice (i.e. punishing the criminal) and mercy (i.e. letting the criminal go). What then?
        - Reply: Conflicts between virtues are impossible. Phronesis tells you what each virtue tells you to do *and in what amount* according to the situation. E.g. act according to justice but show appropriate mercy if there are difficult circumstances
    - Circularity: Aristotle defines a virtuous act as an act a virtuous person would do and a virtuous person as a person who does virtuous acts. This definition is circular and provides no clarity on what virtuous acts and virtuous people are
        - Reply: You can say Aristotle describes the virtuous person in terms of *eudaimonia* and so the definition is not circular
    - Difference between morally good and eudaimonia: E.g. a nurse who spends her life helping others and dies young did not achieve eudaimonia, but she was morally good
        - Reply: Aristotle was never trying to give an account of what a *morally* good life is, but a good life *in general*. However, a good life in general (eudaimonia) will also include helping others and demonstrating virtues such as generosity and kindness

Metaethics

- **Realism vs anti-realism**: This is about whether mind-independent (objective) moral properties – such as good, bad, right, wrong – exist
  - Arguments for realism:
    - Mill's "proof" of utilitarianism (naturalism): People desire happiness, so happiness is desirable. The fact that the happiness is desirable is the only possible proof we could give that happiness is *good*. All our other values (e.g. truth, freedom, dignity) constitute what makes us happy, and so not only is happiness good, it is the *only* good (because all other goods are just a part of happiness). So, good = happy.
      - Reply: Naturalistic fallacy: just because "good" and "happy" are *correlated*, it is a fallacy to conclude that they are the same thing
    - Moore's intuitionism (non-naturalism): Moral properties are non-natural properties, which we know about through rational intuition
    - Moral progress (e.g. people no longer keep slaves) suggests there are mind-independent moral properties
    - If realism is false it would lead to moral nihilism: e.g. if "goodness" and "badness" don't exist, there's no reason not to go round murdering etc.
  - Arguments for anti-realism:
    - Metaphysical queerness: Mind-independent moral properties would mean there are "goodness" and "badness" atoms or similar. If stealing is (objectively) bad, it would mean the (physical) act of stealing would have not-to-be-doneness built into it, which just sounds weird
    - Epistemic queerness: How would we *know* whether something is good/bad? It's not like knowing something is green, which you can see. To say "intuition" (Moore) doesn't really explain anything
    - Relativity: Different cultures believe different things are right/wrong. It seems unlikely that such disagreement would emerge if mind-independent moral properties existed
- **Cognitivism vs non-cognitivism**: This is about what people *mean* when they make moral statements (e.g. "murder is wrong")
  - Arguments for cognitivism:
    - Moral reasoning: E.g. 1. If stealing is wrong, then Smith is bad. 2. Stealing is wrong. 3. Therefore, Smith is bad. In contexts where people embed moral judgements like premise 1 ("if stealing is wrong…") it's

obvious they are meant as cognitive statements that can be true/false

- Arguments for non-cognitivism:
  - Ayer's verification principle: For a statement to be meaningful/ cognitive, it must be an analytic truth or empirically verifiable. Moral judgements are neither, and so moral judgements are non-cognitive
    - Hume's Fork is basically the same argument, but with different words: "Judgements of reason" must be either "relations of ideas" or "matters of fact". Moral judgements are neither, so moral judgements are not judgements of reason
  - Moral judgements motivate action: Cognitive beliefs (e.g. a belief that Paris is the capital of France) *don't* motivate action. Emotions (e.g. wanting to go to Paris, liking French architecture, etc.) *do* motivate action. Moral judgements motivate you to act, suggesting they are the latter (emotional attitudes) rather than the former (cognitive beliefs).
  - Is/ought: "Ought" statements don't logically follow from "is" statements. E.g. you have to do this switch from talking about facts (e.g. Smith murdered Jones) to values/ought (e.g. what Smith did was wrong). The reason "ought" statements don't follow from "is" statements is because "ought" statements are non-cognitive whereas factual ("is") statements are cognitive

- 5 specific metaethical theories:
  - Realist theories:
    - **Naturalism:** Moral statements are cognitivist, and are true/false in virtue of (realist) natural properties (e.g. "murder is wrong" is true because it causes pain - and pain is a natural property)
    - **Non-naturalism**: Moral statements are cognitivist, and are true/false in virtue of (realist) non-natural properties (e.g. "murder is wrong" is true because it has some non-natural property of wrongness)
  - Anti-realist theories:
    - **Error theory:** Moral statements are cognitivist, but are always false because moral properties don't exist (e.g. "murder is wrong" is false because "wrongness" doesn't exist)
    - Non-cognitivist theories:
      - **Emotivism:** Moral statements are non-cognitivist (so neither true nor false) and express emotional attitudes to things (e.g. "murder is wrong" really means something like "boo! Murder")

- **Prescriptivism:** Moral statements are non-cognitivist (so neither true nor false) and express prescriptions/instructions (e.g. "murder is wrong" really means something like "don't murder people!")

Applied Ethics

- Stealing
  - Utilitarianism: Depends on act/rule. Act: OK to steal if it increases happiness. Rule: could set up a general rule "don't steal" because it would increase societal happiness knowing that people can't just steal from you at any moment to increase their happiness
  - Kant: Never steal because if it was always OK to steal it would lead to a contradiction in conception. No exceptions.
  - Aristotle: Stealing is never in the golden mean, and so is always wrong.
- Simulated killing
  - Utilitarianism: Probably OK because the person partaking in simulated killing gets pleasure and isn't causing any harm. However, if playing violent video games makes people more violent IRL, then maybe simulated killing is not OK according to rule utilitarianism
  - Kant: Simulated killing doesn't conflict with duty
  - Aristotle: Simulated killing may encourage negative character traits (i.e. vices instead of virtues), but a virtuous person may play video games in moderation (e.g. as occasional entertainment or as competition)
- Telling lies
  - Utilitarianism: Similar to stealing: it differs according to act/rule. Act utilitarianism would say you should tell white lies (e.g. "you look nice" even when someone doesn't) because it increases happiness. Rule utilitarians might enforce a general rule not to lie.
  - Kant: Same as stealing: if everyone always told lies it would lead to a contradiction in conception. No exceptions (even if an axe murderer asks you which way the victim went)
  - Aristotle: Unclear. Aristotle says stuff that suggests he sees truth as a "final end" and so lying is always bad. But maybe he would say the virtuous person would tell lies in certain circumstances (e.g. the "you look nice" example above is just good social common sense)

- Eating animals
  - Utilitarianism: Pain and pleasure are all that matter - regardless of species (Peter Singer). Animals feel pain and if your pleasure at eating them is less than the pain of the animal being killed then you shouldn't do it. But you could say the animal wouldn't have existed if it wasn't farmed for food and so if its existence was a generally happy life then it could be better to eat animals than not. Depends on circumstances as usual with utilitarianism.
  - Kant: Animals, unlike humans, do not have rational will. So, whereas Kant says you should never treat humans as mere means to an end, you can treat animals as a means to an end.
  - Aristotle: Similar to Kant in that animals do not have a share in eudaimonia because they are incapable of reason and so it's OK to eat them. However, more recent virtue theorists (Cora Diamond) have made arguments along the lines that eating animals can be a vice if done in ways that exhibit e.g. callousness and cruelty (factory farming) but virtuous people may eat animals if they exhibit e.g. care and sympathy (ethical farming)



METAPHYSICS OF GOD

This module covers the concept of God as typically conceived by the three main monotheistic religions and whether such a concept is possible. It also covers several arguments for and against the existence of God and looks at the question of whether religious language is meaningful or meaningless.

THE CONCEPT OF GOD

OVERVIEW

The concept of God in A-level philosophy is the concept of God as generally conceived by the three main monotheistic religions – Christianity, Islam, and Judaism.

Typically, these religions agree that God has the following four divine attributes:

1. **Omnipotence**
2. **Omniscience**
3. **Supremely good (omnibenevolence)**
4. **Eternal or everlasting**

However, there are a number of arguments which claim that these characteristics are incoherent or logically inconsistent. The A-level philosophy syllabus mentions three *problems* with this concept of God:

1. **The problem of the stone**
2. **The euthyphro dilemma**
3. **Omniscience vs. Free will**

THE DIVINE ATTRIBUTES

God, as commonly understood, has a number of key characteristics. These characteristics can be described as the "divine attributes". The divine attributes are as follows:

Omnipotence

Omnipotence translates as *all powerful*.

God is imagined to be perfectly powerful – it's not possible for there to exist a being with more power than God. God is the most powerful being possible. There are no limits to God's power.

But some philosophers argue that omnipotence doesn't mean God can do literally *anything*. For example, God can't make "1+1=5" true, or create a 4-sided triangle because

these are logical contradictions. Omnipotence is thus understood by many philosophers as the claim that God can do anything that's *logically* possible. St. Thomas Aquinas argued that this isn't a *real* limitation on God's power - what is logically impossible is not anything at all, it is *meaningless*. So, even if we say God can't do what is logically impossible (i.e. something that contains a contradiction) this isn't any *real* limitation on God's power – He would still be omnipotent.

However, some philosophers argue that the concept of omnipotence is itself contradictory. For more information, see the problem of the stone (page 151).

Omniscience

Omniscience translates as *all knowing*.

This is to say God has perfect knowledge. He knows everything – or, at least, everything it is *possible* to know.

For example, some philosophers argue that God doesn't know what we humans are going to do in the future because they have free will. The claim is still that God knows everything it's *possible* to know – but that it is not possible to know the future. For further discussion of this issue, see the problem of free will vs. omniscience (page 153).

Omnibenevolence

Omnibenevolence translates as *all loving*.

It's best understood as the claim that God is perfectly *good*. God always does what's morally good – He never does anything bad or evil.

Some philosophers have suggested God's omnibenevolence means God *can't* commit evil – i.e. that it's not *possible* for God to do evil things. This would appear to contradict the claim that God is omnipotent, though, and so most philosophers argue that God *can* do evil but He *chooses* never to do so. However, the euthyphro dilemma (page 151) pushes the apparent conflict between omnipotence and omnibenevolence further: If an omnipotent God *created* morality, then wouldn't "good" just mean *whatever* God said or did?

Eternal/everlasting

There is disagreement among religious philosophers as to whether God is eternal or everlasting:

- Everlasting: God exists *within* time
- Eternal: God exists *outside of* time

If God exists within time, then He is everlasting. This is to say He was there at the beginning of time and will continue to exist forever.

The concept of an eternal God is more difficult to imagine. If God exists outside of time, then He has no beginning or end as these concepts only make sense within time.



Boethius described time as a circle. Human experience is of travelling round the circle, whereas an eternal being (centre) experiences all points on the circle simultaneously

The theologian Boethius (477-524) described God's eternal relationship with time as:

"the whole, simultaneous, and perfect possession of boundless life"

Building on Boethius' description, philosophers Eleonore Stump and Norman Kretzmann describe 4 features of an eternal being:

- Life: An eternal being has life in a psychological sense (but not a physical or biological sense)
- Limitless: An eternal being is without limit and cannot be limited and so has no beginning or end
- Atemporal: An eternal being is atemporal, existing outside of time
- Experiences all at once: Whereas humans experience time in succession, i.e. one moment at a time, an eternal being experiences all moments at once

RESPONSES TO THE CONCEPT OF GOD

The problem of the stone

The problem of the stone argues that the concept of an omnipotent (i.e. all-powerful) being is self-contradictory. It asks: If God is omnipotent, can He create a stone so heavy He can't lift it?

- If He can't then He's not powerful enough to *create* this stone
- But if He can then He's not powerful enough to *lift* the stone

It seems that either way there is something God cannot do – which would mean He's not omnipotent.

*Possible **response** to the problem of the stone:*

George Mavrodes replies to the problem of the stone by arguing that "a stone an omnipotent being can't lift" is not a possible thing – it's a contradiction. And, as discussed in the meaning of omnipotence (page 148), it's not a *real* limitation on God's power to say God can't do what's logically impossible because what's logically impossible is meaningless.

The euthyphro dilemma

The euthyphro dilemma takes its name from Plato's *Euthyphro*. It looks at whether morality is created by or is independent of God. However, it seems that both options challenge the concept of God as defined above.

Back when Plato was writing, the euthyphro dilemma was directed against the many Gods of ancient Greece. However, the euthyphro dilemma can be adapted to the modern concept of God. Applied to a moral judgement such as "torturing babies is wrong", we can ask:

1. Is torturing babies wrong because God says it's wrong?
2. Or, does God command "don't torture babies" because it is wrong?

In other words, does God *create* morality or does God just *follow* morality?

If option 2 is the case – i.e. morality is something *independent* of God – then this presents a challenge to God's *omnipotence*. The reason for this is that God's power would be limited by morality: God only *follows* the rules of morality, He can't change them. God would not be powerful enough to make "torturing babies is good" true, for example.

But if option 1 is true instead – i.e. God *created* morality – then this seems to challenge God's *omnibenevolence*. This argument can be made in two ways:

- If God creates morality then morality is *arbitrary*
- If God creates morality then "God is good" is a *tautology*

Firstly, if God creates morality, then God could have said "torturing babies is good" or whatever rule He wanted and it would be true. There would be no *real* reason why God should say some things are good because *whatever* God says is good would be good. In other words, if God creates morality, then morality is arbitrary.

A slightly different approach argues that if God created morality then "God is good" becomes a tautology: it says the same thing twice. If what is good is just whatever God says is good, then "God is good" would say nothing meaningful about God because it would be like saying "good is good", or "God says what God says" – both of which are trivially true.

So, the euthyphro dilemma drives a wedge between the claims that God is omnipotent and God is omnibenevolent. We don't want to undermine the claim that God is omnipotent (which option 2 implies), but if we claim that God creates morality (i.e. option 1) then it is difficult to make sense of the claim that God is omnibenevolent.

So, in short, the euthyphro dilemma says that either:

- God is not omnipotent, or
- "God is omnibenevolent" is meaningless

*Possible **response #1** to the euthyphro dilemma:*

We can accept that God creates morality (and so is still omnipotent) but reject the claim that this makes morality arbitrary by arguing that God decides what morality is based on *other* factors – such as love. In other words, God didn't just command "torturing babies is wrong" for no reason – He commanded it because He loves us and so doesn't like to see humans suffer.

So, God is *powerful* enough (i.e. omnipotent enough) to make "torturing babies is good" true, but He chose not to for a reason: God chose the rule "torturing babies is bad" instead because of His love for humanity, not some random and arbitrary whim. This preserves the claim that God is omnibenevolent without making the definition of "good" or "omnibenevolence" an arbitrary one.

*Possible **response #2** to the euthyphro dilemma:*

We can reject the claim that "God is good" becomes a tautology if God creates morality by arguing that "God's will" and "good" are two different *concepts* for the same thing.

For example, "water" and "H2O" refer to the same thing, but "water is H2O" is not a tautology because "water" and "H2O" are two different concepts. It's possible, say, that you could know what "water" is without knowing that "water is H2O". It would thus be a *discovery* to learn that "water is H2O".

Similarly, we can say that "good" and "God's will" are two different concepts for the same thing – we can understand them separately of each other. It is a *discovery*, then, to learn that "God's will is good" in the same way it is a discovery to learn that "water is H2O".

Omniscience vs. free will

As an omniscient being, God knows everything. But if God knows everything, then He must know what I'm going to do *before* I do it. But if God already knows what I'm going to do before I do it, then it's not possible that I can do anything else – in other words, I don't have free will.

For example:

- God knows I'm going to go to the fridge to get a beer
- If God already knows that I'm going to go to the fridge to get a beer before I do it, then it must be *true* that I go to the fridge to get a beer
- If it's true that I go to the fridge to get a beer, then it can't be false that I go to the fridge to get a beer
- In other words, I don't have a choice.

And if I don't have a choice whether or not I go to the fridge to get a beer, then I don't have free will. So, either:

- God is omniscient *but we don't have free will*
- We have free will *but God is not omniscient*

Either way, we must reject a key claim of religion.

*Possible **response #1** to the problem of omniscience vs. free will:*

We can respond that free will makes it *impossible* to know the future. And, as discussed in the meaning of omniscience (page 149), some philosophers understand "God is omniscient" as the claim that God knows everything it is *possible* to know. If free will makes it *impossible* to know the future, then we can still say that God is still omniscient because He still knows everything that's possible to know.

However, this response only works if God's relationship with time is understood as *everlasting* rather than eternal (page 149): If God is outside of time (eternal), then He must surely see what we would call the future. And so, if God is eternal, He knows the future (including what I'm going to do next).

*Possible **response #2** to the problem of omniscience vs. free will:*

From the perspective of an *eternal* God, what we call "tomorrow" is visible to Him in the same way "now" is to us – because an eternal God perceives all times at once.

Applying this to the free will dilemma, we can say God's knowledge of the future is similar to our knowledge of the *present*. For example, say you observed someone choosing a bottle of wine at a supermarket, you wouldn't take that as evidence that this person lacked free will. And similarly, the fact that God is observing what I'm doing *right now* clearly doesn't prove I lack free will – God could just be observing me *freely* choose my actions.

But an eternal God perceives all moments in time – yesterday, today, tomorrow – *simultaneously*. And so, God observes me *freely* choosing my future actions *right now*. In fact, God also observes me freely choosing my actions yesterday and tomorrow at the same time – because the concepts of "yesterday" and "tomorrow" only make sense to beings *within* time, which God is not.

So, if God is eternal, then He is constantly observing the future – and thus knows what is happening in the future. But the fact that God is observing this future does not mean we lack free will.

*Possible **response #3** to the problem of omniscience vs. free will:*

Some theologians defend Molinism: the view that God has counterfactual "middle knowledge" of all possible worlds. This means God knows what would happen *if something were the case.*

Molinists argue that God's middle knowledge played an important role in why He created our world the particular way it is: He surveyed all the possible worlds (i.e. what would happen *if He did X, Y, and Z)* and decided, based on every possible counterfactual outcome, to make *this* particular world.

So, returning to the example above: God might *counterfactually* know, for example, that "*if it's Friday and there's beer in the fridge and he wants to relax*, he will get a beer." And so, if those are the circumstances then there is some sense in which God *knows* what I'm going to do next. But God's knowledge here is not *factual* knowledge – it's *counterfactual* knowledge. So, technically, I'm free to do otherwise – even though I won't.

ARGUMENTS CONCERNING GOD'S EXISTENCE

OVERVIEW

A-level philosophy looks at four arguments relating to the existence of God:

1. **Ontological arguments**
2. **Teleological arguments**
3. **Cosmological arguments**
4. **The problem of evil**

There are various versions of each argument as well as numerous responses to each. The key points of each argument are summarised below:

	Ontological	Teleological	Cosmological	Problem of Evil
God exists?	Yes	Yes	Yes	No
Summary	God must exist by definition	The universe appears to be designed	There must be a cause of everything	If God existed, there wouldn't be evil
Versions	Anselm Descartes Malcolm (necessary existence)	Paley Hume Swinburne (temporal order)	Kalam argument Aquinas (motion, causation, and contingency) Descartes Leibniz	The logical problem of evil The evidential problem of evil
Problems	Gaunilo's island "God doesn't exist" is not a contradiction Existence is not a predicate	Evolution Hume's objections to the analogy Spatial disorder The designer might not be God	Infinite sequence The first cause might not be God "An uncaused event" is not a contradiction Fallacy of composition	Free will Soul making

ONTOLOGICAL ARGUMENTS

Ontological arguments are unique in that they use a priori reasoning. The various ontological arguments we will look at aim to deduce God's existence from the *definition* of God. And so, if these arguments work, then this means "God exists" is an analytic truth akin to "a bachelor is an unmarried man".

Anselm's ontological argument

Anselm of Canterbury (1033-1109) was the first to propose an ontological argument in his book *Proslogium*. His argument can be summarised as:

1. By definition, God is a being greater than which cannot be conceived
2. It is greater to exist in reality than to exist only in the mind
3. Therefore, God exists

In other words, imagine two beings:

- One is said to be maximally great in every way, *but does not exist.*
- The other is maximally great in every way and *does exist.*

Which of these two beings is greater?

Presumably, the second one – i.e. the being that *does* exist – because it is greater to exist in reality than in the mind. And since God is defined as the *greatest* possible being (i.e. "a being greater than which cannot be conceived"), the greatest possible being would be one that exists. And so, God exists.

Descartes' ontological argument

Descartes offered his own version of the ontological argument:

1. I have the idea of God
2. The idea of God is the idea of a supremely perfect being
3. A supremely perfect being does not lack any perfection
4. Existence is a perfection
5. Therefore, God exists

Descartes' argument here is very similar to Anselm's, except it defines God as "a supremely perfect being" rather than "a being greater than which cannot be conceived". A supremely perfect being would exist because, if it didn't, then it wouldn't be a supremely perfect being – it would lack the "perfection" of existence.

Descartes then goes on to argue that because it is possible to derive "God exists" from the definition of God in this way "God does not exist" is a self-contradiction. The idea of God not existing is, to Descartes, as contradictory as a married bachelor or a 4-sided triangle.

Responses to (standard) ontological arguments

Gaunilo's island

Gaunilo of Marmoutiers (994-1083) argued that if Anselm's argument was valid, then anything can be defined into existence. For example:

1. The perfect island is, by definition, an island greater than which cannot be conceived
2. We can coherently conceive of such an island i.e. the concept is coherent
3. It is greater to exist in reality than to exist only in the mind
4. Therefore, this island must exist

Notice that this argument has exactly the same format as Anselm's, but its conclusion is obviously false. Gaunilo argues that if Anselm's argument was valid, then we could define *anything* into existence – the perfect shoe, the perfect tree, the perfect book, etc.

We can adapt Gaunilo's objection to attack Descartes' ontological argument too. For example, we could claim that the idea of a perfect island is the idea of a *supremely perfect island*, that existence is a perfection, and so therefore the perfect island must exist. Again, though, this conclusion is silly, and so this argument format must fail somewhere.

Hume: "God does not exist" is not a contradiction

Ontological arguments reason from the *definition* of God to the conclusion that God exists. And so, if these arguments are successful, this would make "God exists" an analytic truth (or, as Hume would say, a *relation of ideas* – see page 68 for more) in the same way that "triangles have 3 sides" is an analytic truth.

Hume argues that we cannot coherently conceive of contradictions. For example, a 4-sided triangle is a contradictory idea – it cannot be conceived. You can *try* to imagine a 4-sided triangle but you'll either imagine a 4-sided shape or a 3-sided shape. You can't imagine a *4-sided triangle,* though, because the idea doesn't make sense. It isn't conceivable because contradictions are not conceivable.

But Hume argues that "God does not exist" *is* conceivable – we can coherently imagine God existing just as we can coherently imagine God not existing. And so, Hume argues,

ontological arguments must fail somewhere because "God does not exist" is not a contradiction. So, in summary, Hume's argument is:

- If ontological arguments work, then "God does not exist" is a contradiction
- A contradiction cannot be coherently conceived
- But "God does not exist" can be coherently conceived
- Therefore, "God does not exist" is not a contradiction
- Therefore, ontological arguments do not work

*Possible **response** to the claim that "God does not exist" is not a contradiction:*

Descartes would respond that even though it may seem that "God does not exist" is conceivable, closer examination reveals it is not (because it is a contradiction).

For example, we may think we can coherently conceive that "8670/90=94" is true, but close inspection reveals that such a thing is logically impossible (because the actual answer is 96.333). It's the same thing with God: "God does not exist" may initially *seem* coherently conceivable/possible, but closer inspection reveals "God does not exist" is a contradiction in the same way "8670/90=94" is a contradiction.

Kant: Existence is not a predicate

Kant argues that existence is not a predicate (i.e. a property) of things in the same way, say, "green" is a property of grass or "being over 100m tall" is a property of the Eiffel Tower. To say something exists doesn't *add anything* to our concept of it.

To illustrate this point, Kant uses the example of money. He says that the *concept* of 100 real thalers (the currency of the day) contains no more money than the *concept* of 100 possible thalers. You can't pay for stuff with the *concept* of 100 thalers because it's just a concept. Whether that concept is the concept of something real or not doesn't make a difference to the shopkeeper!

Another example: Imagine a unicorn. Then imagine a unicorn *that exists*. What's the difference between the two ideas? Nothing! Adding existence to the idea of a unicorn doesn't make the idea exist. Likewise, adding "existence" to the *concept* "God" doesn't change anything.

Another way of thinking about it is this: When someone says "God exists", they don't mean "there is a God *and He has the property of existence*". If they did, then when someone says "God does not exist", they'd mean, "there is a God and He has the property of non existence" – which doesn't make sense! Instead, what people mean when they say "God exists" is that "God exists *in the world*". However "God exists in the world" does not follow from any of the ontological arguments above and thus such arguments fail to prove God's (actual, real-world) existence.

Malcolm's ontological argument (*necessary* existence)

Kant's objection above is generally considered to be the most powerful problem for ontological arguments. So, in response, some philosophers have developed alternate ontological arguments that avoid this criticism.

Philosopher Norman Malcolm (1911-1990) accepts that Descartes and Anselm's arguments (at least as presented above) don't work. But Malcolm offers his own ontological argument. Malcolm argues that it's not *existence* that is a perfection, but the *logical impossibility of non-existence* (in other words, *necessary existence*).

Malcolm's ontological argument is as follows:

1. Either God exists or does not exist
2. God is a being greater than which cannot be conceived
3. A being greater than which cannot be conceived cannot come into existence or go out of existence
4. So, if God exists, God cannot go out of existence
5. So, if God exists, God's existence is necessary
6. And also, if God does not exist, God cannot begin to exist
7. And so, if God does not exist, God's existence is impossible
8. Therefore, God's existence is either necessary or impossible (1 + 5 + 7)
9. God's existence is impossible only if the concept of God is self-contradictory
10. The concept of God is not self-contradictory
11. Therefore, God's existence is not impossible
12. Therefore, God exists necessarily (8 + 11)

Necessary existence *is* a real predicate – it adds something to the concept – and so Malcolm's ontological argument avoids Kant's objection above.

There are a lot of premises to Malcolm's argument, so a condensed summary is as follows:

- The concept of God is the concept of a being with necessary existence
- Such a being either exists or it doesn't
- If such a being doesn't exist then it's impossible for it to ever exist
- But it's not impossible for such a being to exist
- Therefore, such a being (with necessary existence) exists

Responses to Malcolm's ontological argument

The concept of God *is* self-contradictory

We may reject Malcolm's argument by rejecting premise 10 above. There are several philosophical arguments that the concept of God *is* self-contradictory, as discussed in the concept of God section (page 151).

Fallacy of equivocation

Another potential criticism of Malcolm's argument is that the meaning of "necessary" changes between premises 5 and 8, and the conclusion (12). This is an example of the fallacy of equivocation (where the meaning of a word changes during an argument), which would make Malcolm's argument invalid.

In premises 5 and 8, Malcolm is talking about necessary existence in the sense of a *property* that something does or does not have. In the conclusion, however, Malcolm switches from talking about God having the property of necessary existence to the conclusion that *it is a necessary truth that God exists*. But this doesn't follow: the *concept* of God may have the property of necessary existence, but this does not show that it is a *necessary truth* that God exists. We can accept that *if God exists,* then God has the property of necessary existence, but still deny the conclusion that God exists necessarily.

In other words, we can imagine a being with the property of necessary existence, but this isn't the same thing as saying such a being exists necessarily.

TELEOLOGICAL ARGUMENTS

Teleological arguments are also known as arguments from design. These arguments aim to show that certain features of the world are so complex and suited for a purpose that their existence can only be explained with reference to a designer (God). Teleological arguments can be divided into two forms:

- Teleological arguments from *spatial* order argue that certain features within nature (e.g. the eye) must be designed, and that the designer is God
- Teleological arguments from *temporal* order argue that the laws of nature (e.g. gravity and electromagnetism) must be designed, and that the designer is God

Teleological arguments from spatial order

Paley's design argument

William Paley (1743-1805) wasn't the first to propose a teleological argument, but his version is perhaps the most famous.



A watch has "many parts organised to serve a purpose", which Paley says is the hallmark of design

Paley compares man-made objects, such as a watch, with certain aspects of nature, such as a pebble. If you found a pebble on a beach, say, you might assume it had just been there forever. But if you found a watch on a beach, you wouldn't assume the same thing.

The difference between the watch and the pebble, says Paley, is that a watch is composed of *many parts organised for a purpose*. Paley argues that having many parts organised for a purpose like this is the *hallmark of design*.

And aspects of nature, such as the human eye, are composed of many parts. These parts are also organised for a purpose – in the case of the eye, to see. So, like the watch, the human eye has the hallmark of design – the eye has many parts organised for a purpose.

For something to be designed, it must have had a designer. And, Paley argues, this designer is God.

Hume's design argument

Writing before Paley, David Hume also considered a version of the teleological argument (which he then goes on to reject).

Hume's teleological argument is very similar to Paley's. However, whereas Paley claims that having "many parts organised for a purpose" is a *property* of design, Hume draws an *analogy* between the "fitting of means to ends" in nature and in human design:

> *"The intricate fitting of means to ends throughout all nature is just like (though more wonderful than) the fitting of means to ends in things that have been produced by us - products of human designs, thought,wisdom, and intelligence. Since the effects resemble each other, we are led to infer by all the rules of analogy that the causes are also alike, and that the author of nature is somewhat similar to the mind of man, though he has much larger faculties to go with the grandeur of the work he has carried out. By this argument... we prove both that there is a God and that he resembles human mind and intelligence."*

So, in short, Hume's argument is:

- The "fitting of means to ends" in nature (e.g. the many parts of the eye to enable vision) resemble the "fitting of means to ends" in human design (e.g. the many parts of a watch to tell the time)
- Similar effects have similar causes
- The cause of human designs are minds
- So, by analogy, the cause of design in nature is also a mind
- And this mind is God, given the grandeur of nature's design

Responses to teleological arguments from spatial order

Hume's objections to teleological arguments

Problems with the analogy

In response to his own teleological argument (above), Hume points out various problems with drawing an analogy between the design of man-made objects and design in nature, such as:

- Man-made objects (e.g. houses, watches, phones, etc.) are very different from natural objects (e.g. eyes, legs, brains, etc.)
- We can have *observe* man-made items being designed, but we have no such experience of this in the case of nature
- Even if there are examples of design *within* the universe, it is a fallacy to conclude that this shows that the universe is *itself* designed because we cannot generalise from parts of the universe to the universe as a whole

Such differences weaken the analogy between the design of man-made items and the apparent design of nature.

Infinite time, finite matter

Hume also objects to the teleological argument on the basis that there may be better explanations of the order within nature. For example, if the following two claims are true:

- Time is infinite
- Matter is finite

Then, Hume argues, it is inevitable that matter will organise itself into combinations that appear to be designed.

It's a bit like the monkeys and typewriters thought experiment: Given an infinite amount of time, a monkey will eventually type the complete works of Shakespeare. The point here is not that the monkey would *learn* to write Shakespeare – or even that it would *understand* what it had written. The point is that, given an *infinite* amount of time, the monkey would keep bashing keys until it eventually typed *everything* (including the works of Shakespeare).



This is the nature of infinity. It's inevitable that the monkey will write something that appears to be intelligent, even though it's just hitting letters at random. And the same principle applies with the teleological argument, argues Hume: Given enough time, it is inevitable that matter will arrange itself into combinations that *appear* to be designed, even though they're not.

Causation

Hume famously argues that it takes many instances ("constant conjunction") of one event following another to know that one thing *causes* the other.

For example, experience (ever since you were a baby) tells you that if one snooker ball hits another, this will *cause* the second snooker ball to move. You know that this

relationship holds because you've seen it and similar examples happen *hundreds of times* without fail.

This is how we infer causation: you see A happen (e.g. one snooker ball touches the other snooker ball) and then you see B happen (e.g. the other snooker ball moves) and when you see B follow A *enough times without exception*, you infer that A *causes* B.

The snooker ball example is something you can observe again and again – the two things are in *constant conjunction*. But contrast the snooker ball example with something that happens just once: for example, imagine that you take a sip of tea and at the same time, your friend coughs. Do you assume that your drinking the tea *caused* your friend to cough based on this one instance? Obviously not, because you cannot infer causation from a single instance.

Applying this to teleological arguments, Hume argues that the creation of the universe was a unique event – it only happened once. And given that knowledge of causation requires multiple instances (i.e. constant conjunction), we cannot assume a causal relationship between designer and creation based on this one instance.

Spatial *dis*order

We can also object to teleological arguments by arguing that even if there are instances of order and apparent design within the universe (e.g. the eye), this apparent order is vastly outweighed by apparent *disorder* and randomness.

For example, despite searching for many decades, astronomers have never found any evidence of life on other planets. As far as we can tell, our entire galaxy – perhaps even *all* galaxies – contains no life, no "fitting of means to ends" or "parts organised for a purpose" beyond the examples we find here on Earth.

The existence of spatial disorder like this weakens the claim that the universe was designed. Why would the designer add all this randomness to his design? It would be like a watchmaker adding a bunch of useless cogs and dials to a watch he designed.

Darwin: Evolution by natural selection

Charles Darwin's (1809-1882) theory of evolution by natural selection explains how complex organisms – complete with "parts organised for a purpose", as Paley would say – can emerge from nature *without* a designer.

For example, it may seem that God *designed* giraffes to have long necks so they could reach leaves in high trees. But the long necks of giraffes can be explained without a designer, for example:

- Competition for food is tough
- An animal that cannot acquire enough food will die before it can breed and produce offspring
- An animal with a (random genetic mutation for a) neck that's 1cm longer than everyone else's will be able to access 1cm more food
- This competitive advantage makes it more likely to survive and produce offspring
- The offspring are likely to inherit the gene for a longer neck, making them more likely to survive and reproduce as well
- Longer necked-animals become more common as a result
- The environment becomes more competitive as more and more animals can reach the 1cm higher leaves
- An animal with a neck 2cm longer has the advantage in this newly competitive environment
- Repeat process over hundreds of millions of years until you have modern day giraffes
- The key idea is that – given enough time and genetic mutations – it is inevitable that animals and plants will *adapt* to their environment, thus creating the appearance of design.

This directly undermines teleological arguments from spatial order, such as Paley's and Hume's. Evolution shows that matter can arrange itself in ways that "fit means to ends" (as Hume would say) or that have "many parts organised for a purpose" (as Paley says) without a designer.

Swinburne's teleological argument from *temporal* order

Swinburne's version of the teleological argument differs from Hume's and Paley's. Swinburne distinguishes between:

- Examples of order *within* nature (spatial order)
- And the order of the laws of nature (temporal order)

Swinburne accepts that science, for example evolution, can explain the apparent design of things like the human eye (i.e. spatial order), and so neither Hume's nor Paley's teleological argument succeeds in proving God's existence. However, Swinburne argues, we can't explain the *laws* of nature (i.e. temporal order) in the same way.

For example, the law of gravity is such that it allows galaxies to form, and planets to form within these galaxies, and life to form on these planets. But if gravity had the opposite

effect – it repelled matter, say – then life would never be able to form. In fact, if gravity was even slightly stronger, planets wouldn't be able to form. How are we to explain this?

Unlike spatial order, we can't give a *scientific* explanation of why the laws of nature are as they are. Yes, science can explain and predict things using these laws – but it can't explain *why* these laws exist as they do. In the absence of a scientific explanation of why the laws of nature are the way they are, Swinburne argues, the best explanation of temporal order is a *personal* explanation.

We give personal explanations of things all the time. For example, to say "this sentence exists because *I* chose to write it" or "that building exists because *someone* built it" invokes the personal intentions of people rather than impersonal scientific laws. Swinburne argues that, by analogy, we can explain the laws of nature in a similarly personal way: someone designed them (God).

So, in short, Swinburne's argument is:

- There are two types of explanation: scientific and personal
- We give a *personal* explanation of temporal order in human actions, e.g. this sentence exists *because I chose to write it*
- By analogy, we can give a similar personal explanation of why the laws of nature are the way they are, e.g. the laws of nature are the way they are *because God designed them*
- There is no scientific explanation of why the laws of nature are the way they are
- So, this leaves only the personal explanation (i.e. that God designed these laws)
- So, God exists

Responses to Swinburne's teleological argument

Multiple universes

Hume's earlier argument regarding finite matter and infinite time (page 163) can be adapted to respond to Swinburne's teleological argument.

Instead of arguing that *time* is infinite, as Hume does, we could argue that the *number of universes* is infinite. This idea of multiple universes is popular among some physicists, as it explains various phenomena in quantum mechanics.

And so, if there is an infinite number of universes (or even just a large enough number), it's highly likely that *some* of these universes will have laws of nature that support the formation of life (i.e. temporal order). Of course, when such universes do exist, it is just sheer luck rather than design.

Another analogy: If you roll a million-sided dice enough times, you'll eventually roll a 1 – but there will be many instances where you roll other numbers, like 656 or 119,411. Likewise, if you create enough universes with randomly different scientific laws, eventually some of these universes will support life. But for every universe like this, there will be many more universes that don't support life.

In other words, there could be just as much *temporal* disorder as there is *spatial* disorder (page 164).

Is the designer God?

Both Hume and Kant have argued that even if the teleological argument succeeded in proving the existence of a *designer*, this designer would not necessarily be God (i.e. the God as defined in the Concept of God section – page 148). For example:

- God's power is supposedly infinite (omnipotence), yet the universe is seemingly not infinite
- Designers are not always creators and so the designer of the universe and the creator of the universe might be two separate people (e.g. the guy who designs a car doesn't physically build it)
- The design of the universe may be the result of many small improvements by many people
- Designers can die even if their creations live on. How do we know the designer is eternal, as God is supposed to be?

COSMOLOGICAL ARGUMENTS

Cosmological arguments start from the premise that everything is either caused by or depends on something else for its existence. For example, you depended on your parents in order to exist, and they depended on their parents, and so on. Cosmological arguments argue that this chain of reasoning can't go back forever – there must be something that isn't caused by or contingent on something else (God).

Cosmological arguments can be divided into two forms:

- Cosmological arguments from *causation* argue that everything has a cause, and so the universe itself must have a cause, and that cause is God
- Cosmological arguments from *contingency* argue that things that exist contingently cannot be explained without reference to a necessary being, and that necessary being is God

Cosmological arguments from *causation*

The Kalam argument

The Kalam argument is the simplest version of the cosmological argument in the A-level philosophy syllabus. It says:

1. Whatever begins to exist has a cause
2. The universe began to exist
3. Therefore, the universe has a cause
4. (That cause is God)

The evidence for premise 1 is pretty strong: things don't just spring into existence for no reason, they are *caused* to exist. For example, a baby is caused to exist by its parents, a painting is caused to exist by a painter, and the Earth was caused to exist by the forces of nature operating on physical matter. Premise 2 also seems pretty uncontroversial – the universe clearly exists, so it must have began to exist at some point – and the conclusion (3) follows logically from 1 and 2. Point 4 is then kind of implied: What else could possibly be powerful enough to create an *entire universe* from nothing ("ex nihilo") except God?

Aquinas' first way

St. Thomas Aquinas (1225-1274) gave five different versions of the cosmological argument, which are called Aquinas' "five ways". The syllabus lists three of the five ways, but we will deal with the third one later on (page 171) because it is a cosmological argument from *contingency* rather than *causation*.

Aquinas' first way is the argument from *motion*. The key points are:

- Some things in the world are in motion
  - E.g. a football rolling along the ground
- Whatever is in motion must have been put in motion by something else
  - E.g. someone kicked the ball
- If A is put in motion by B, then something else (C) must have put B in motion, and so on
- If this chain goes on infinitely, then there is no first mover
- If there is no first mover, then there is no subsequent movers, and so nothing would be in motion
- But things are in motion
- Therefore, there must be a first mover
- The first mover is God

Aquinas' second way

Aquinas' second way – the argument from *causation* – is basically the same as the argument from motion above, except it talks about a first cause rather than a first mover:

- Everything in the universe is subject to cause and effect
  - E.g. throwing a rock *caused* the window to smash
- C is caused by B, and B is caused by A, and so on
- If this chain of causation was infinite, there would be no first cause
- If there was no first cause, there would be no subsequent causes or effects
- But there are causes and effects in the world
- Therefore, there must have been a first cause
- The first cause is God

Descartes' cosmological argument

Descartes' version of the cosmological argument is a bit more long-winded than the others. The key points are as follows:

- If I was the cause of my own existence I would have given myself all perfections (i.e. I would have caused myself to be omnipotent, omniscient, etc.)
- I do not have all perfections
- So, I am not the cause of my own existence
- So, something else caused me to exist
- I am a thinking thing and have the idea of God
- The cause of an effect must have at least as much reality as the effect
- So, whatever caused me to exist must also be a thinking thing that has the idea of God
- Whatever caused me to exist must either be the cause of its own existence or caused by something else
- If it was caused by something else then that something else must *also* either be the cause of its own existence or caused by something else
- There cannot be an infinite chain of causes
- So, there must be something that caused its own existence
- Whatever causes its own existence is God

There's a bit more to Descartes' version than this. For example, he talks about a cause needed to keep him in existence. But, for the purposes of the exam, the points above constitute the main argument for the existence of God.

Responses to cosmological arguments from causation

Is a first cause necessary?

All of the cosmological arguments above assume something along the lines of "there can't be an infinite chain of causes". For example, the Kalam argument says the universe *began* to exist, and Aquinas says things like there must have been a *first* cause and a *prime* mover. But we can respond by challenging this claim: *why* must there be a first cause? Perhaps there is just an infinite chain of causes stretching back forever – perhaps the universe never *began* to exist and has simply existed for infinity.

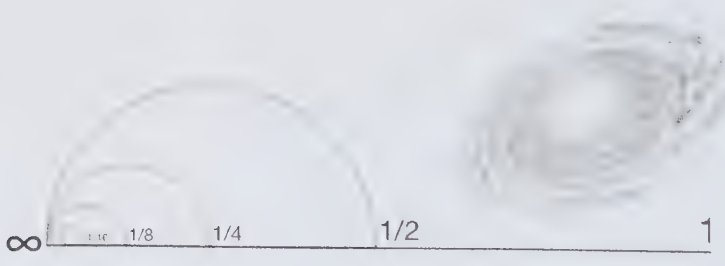
*Possible **response #1** to the claim that a first cause is not necessary:*

- An infinite chain of causes would mean an infinite amount of time has passed prior to the present moment
- If an infinite amount of time has passed, then the universe can't get any older (because infinity+1 = infinity)
- But the universe is getting older (e.g. the universe is a year older in 2021 than it was in 2020)
- Therefore an infinite amount of time has not passed
- Therefore there is not an infinite chain of causes

*Possible **response #2** to the claim that a first cause is not necessary:*

A similar response claims that if the universe is infinitely old, then an infinite sequence of successive moments has passed prior to the present moment.

But the nature of infinity is such that it can never be achieved by successively adding moments. It's thus impossible for an infinite sequence of successive moments to have passed and so there must have been a first moment in order for there to be a present moment.



Hume: "Everything has a cause" is not an analytic truth

Another key assumption (or premise) of all cosmological arguments from causation is something like "everything has a cause". But Hume points out that this is not an analytic truth. In other words, it's *logically* possible for something to exist and not have a cause of its existence.

Yes, all of our experience in the physical world tells us everything has a cause. But just because we've never *seen* something happen without a cause, doesn't mean it's never happened or that it's *impossible* – it just means we've never seen it. Even physics says something along these lines: before the big bang, there was no time and space (because time and space came into existence at the same time as the universe itself) and so no causation (because if A causes B, then B happens after A, but that doesn't make sense without time).

So, just because our experience tells us "everything has a cause" is true, we can't be certain of it in the same way we can be certain of analytic truths. There is no contradiction in the idea of an uncaused event and so it's *logically* possible that the universe exists without a cause. If it's possible for the universe to exist without a cause, then this casts doubt on arguments that assume "everything has a cause" as a premise.

Is the first cause God?

Aquinas' first and second ways and the Kalam argument only show that there is a first cause or a prime mover. But they don't show that this first cause is *God*.

Even if we accept that there was a first cause, it doesn't necessarily follow that God exists – much less the omnipotent, omniscient, and omnibenevolent being described in the concept of God section (page 148). So, even if the cosmological argument is sound, it doesn't necessarily follow that God exists.

Note: Descartes' cosmological argument is a bit different and doesn't suffer this objection. Descartes' version provides specific reasons why this first cause is God: If Descartes' idea of God is the idea of an omnipotent, omniscient, omnibenevolent being, and the cause of this idea "has as much reality" as the idea itself, then it follows that this specific being exists.

Cosmological arguments from *contingency*

Aquinas' third way

We looked at Aquinas' first and second ways - the arguments from motion and causation – above. These arguments (along with the Kalam argument and Descartes' cosmological argument) all assume some variation of the premise that "everything has a cause". Thus, these arguments are cosmological arguments from *causation*.

However, Aquinas' third way – the contingency argument – is a different kind of argument: it is a cosmological argument from *contingency*. This means that rather than talking about causes and effects (which are empirical/physical/a posteriori concepts), these arguments use the more philosophical/logical/a priori concepts of necessary and contingent existence.

Necessary and contingent existence are similar concepts to necessary and contingent *truth*, which come up in the epistemology module (page 59). We also touched on the idea of God existing necessarily in Malcolm's ontological argument (page 159). The difference between the two types of existence is this:

Contingent existence	Necessary existence
Something that *might not have existed*	Something that *must always exist* and can't go out of existence
Examples: • The tree wouldn't exist *if the gardener didn't plant the seed* • Planet Earth wouldn't exist *if the force of gravity didn't exist*	Unlike things that exist contingently, God's existence *isn't contingent* on anything else

So, using this idea of contingent existence, Aquinas argues that:

- Everything that exists contingently did not exist at some point
- If everything exists contingently, then at some point nothing existed
- If nothing existed, then nothing could begin to exist
- But since things did begin to exist, there was never nothing in existence
- Therefore, there must be something that does not exist contingently, but that exists *necessarily*
- This necessary being is God

Leibniz: Sufficient reason

Like Aquinas' third way, Leibniz's cosmological argument is another cosmological argument from *contingency*.

The key premise of Leibniz's cosmological argument is the *principle of sufficient reason*. The principle of sufficient reason says that every truth has an explanation of *why* it is the case (even if we can't know what this explanation is):

> *"No fact can be real or existing and no statement true unless it has sufficient reason why it should be thus and not otherwise."*

Leibniz then defines two different types of truth:

- Truths of reasoning (this is basically another word for *necessary* truths)
- Truths of fact (this is basically another word for *contingent* truths)

The *sufficient reason* for truths of reasoning (i.e. analytic truths) is revealed by analysis. When you analyse and understand "3+3=6", for example, you don't need a further explanation why it is true. To know what "3+3=6" *means* is to know that it is true. You don't need to appeal to anything beyond this to provide sufficient reason why "3+3=6" is true.

But it is more difficult to provide sufficient reason for truths of fact (i.e. contingent truths) because you can always provide more detail via more contingent truths. For example, you can explain the (contingent) existence of a tree by saying someone planted a seed. But the fact that someone planted a seed is also a contingent fact, and so you could then ask why the person planted the seed, or why seeds exist in the first place, or why the laws of physics are the way they are, and so on. This process of providing contingent reasons for contingent facts could go on for infinity. So, until we break this cycle of providing contingent facts, we cannot provide sufficient reason for truths of fact.

Leibniz argues that in order to escape the endless cycle of contingent facts and provide sufficient reason for truths of fact (i.e. contingent truths), we need to step *outside* the sequence of contingent facts and appeal to a *necessary* substance. This necessary substance is God, Leibniz says.

So, in short, the argument is:

- The principle of *sufficient reason:* Every true fact has a reason *why* it is true
- The sufficient reason for *necessary* truths (e.g. "3+3=6") is revealed by analysis
- But we cannot provide sufficient reason for *contingent* truths (e.g. "water boils at 100°C" or "the universe exists") without referring to other contingent truths, which also need a sufficient reason
- So, in order to provide sufficient reason for contingent truths, we need to appeal to a *necessary* substance: God

Responses to cosmological arguments from contingency

Russell: Fallacy of composition

Bertrand Russell argues that cosmological arguments fall foul of the *fallacy of composition*. The fallacy of composition is an invalid inference that because the *parts* of something have a certain property, the *entire thing* must also have this property. For example:

- Just because all the players on a football team are good, this doesn't guarantee the team is good. For example, the players might not work well together.
- Just because a sheet of paper is thin, it doesn't mean things made from sheets of paper are thin. For example, a book with enough sheets of paper can be thick.

Applying this to Leibniz's cosmological argument, we can argue that just because everything *within* the universe requires sufficient reason to explain its existence, this doesn't mean the universe *itself* requires sufficient reason to explain its existence. Instead, Russell says:

> *"the universe is just there, and that's all."*

*Possible **response** to the claim that cosmological arguments commit the fallacy of composition:*

- OK, but everything within the universe exists contingently
- And if everything within the universe didn't exist, then the universe *itself* wouldn't exist either (because that's all the universe is: the collection of things that make it up)
- So the universe itself exists contingently, not just the stuff within it
- And so the universe itself requires sufficient reason to explain its existence

Hume: "God does not exist" is not a contradiction (again)

Earlier, we looked at Hume's objection to ontological arguments that "God does not exist" is not a contradiction because it can be coherently conceived (page 157). Adapting this argument to the concept of a "necessary being", we can argue that:

1. If God exists necessarily, then "God does not exist" is a contradiction
2. A contradiction cannot be coherently conceived
3. But "God does not exist" can be coherently conceived
4. Therefore, "God does not exist" is not a contradiction
5. Therefore, God does not exist necessarily

*Possible **response** to the claim that "God does not exist" is not a contradiction:*

We can argue that the response above misunderstands what "God exists necessarily" means. It assumes "God exists necessarily" means "God exists" is *necessarily true* – i.e. that it is a necessary truth that God exists. If this is the claim, then yes, it would imply that "God does not exist" is a contradiction, which it obviously isn't (because, as Hume says, we can coherently conceive of the statement "God does not exist').

But the claim "God exists necessarily" is not the claim that "it is a necessary truth that God exists". Instead, it is a claim about *what kind of existence* God has. It is a claim about

a *property* of God. Understood this way, it is clear that premise 1 of the argument above is false: it doesn't follow that if God exists necessarily then "God does not exist" is a contradiction. In other words, we can accept that "God does not exist" is logically possible whilst also accepting that *if God exists*, then God exists necessarily.

So, in short, the argument above does not show that a necessary being is impossible. We can then argue that Aquinas' third way and Leibniz's cosmological argument provides reason to believe that such a necessary being (God) *does* exist.

Note: This is a very similar point used to object to Malcolm's ontological argument (page 160). Malcolm switches from talking about necessary existence as a property that God has, to the conclusion that it is a necessary truth that God exists.

THE PROBLEM OF EVIL

The problem of evil uses the existence of evil in the world to argue that God (as defined in the concept of God from page 148) does not exist. These arguments can be divided into two forms:

- The *logical* problem of evil says the existence of God is *logically impossible* given the existence of evil in the world
- The *evidential* problem of evil says that, while it is possible that God exists, the amount and way in which evil is distributed in our world is *pretty strong evidence* that God doesn't exist

Further, evil can be divided into two *types* of evil:

- *Moral* evil: Evil acts committed by people
  - e.g. torture, murder, genocide, etc.
- *Natural* evil: Suffering as a result of natural processes
  - e.g. earthquakes, tsunamis, volcano eruptions, etc.

One final definition: A *theodicy* is an explanation of why an omnipotent and omniscient God would permit evil (i.e. a response to the problem of evil).

The *logical* problem of evil

J.L. Mackie: The inconsistent triad

The simple version of Mackie's argument is that the following statements are logically inconsistent – i.e. it's impossible for all three to be true simultaneously:

- An omnipotent God exists
- An omnibenevolent God exists
- Evil exists

Mackie's argument is that, logically, a maximum of two of these three statements can be true – but not all three. This is sometimes referred to as the *inconsistent triad.*



Mackie argues that if God is omnibenevolent, then He *wants* to stop evil. And if God is omnipotent, then He's *powerful enough* to prevent evil. These two properties imply that evil wouldn't exist because, if God both wanted to and was powerful enough to get rid of all evil, He would just get rid of it. But, obviously, evil does exist in the world: people steal, get murdered, and so on.

So, Mackie argues that if an omnipotent and omnibenevolent being existed evil would not exist. But evil does exist. Therefore, there is no omnipotent and omnibenevolent being. Therefore, God does not exist.

Responses to the logical problem of evil (theodicies)

God couldn't exist without evil

People often make claims like "you can't appreciate the good times without experiencing some bad times". This is basically what this reply says: without evil, good couldn't exist.

*Mackie's **response** to this problem:*

Mackie rejects this response. *Why* can't we have good without evil?

For example, imagine if we lived in a world where everything was red. Presumably, we wouldn't have created a word for "red", nor would we know what it meant if someone

tried to explain it to us. But it would still be *true* that everything is red, *we just wouldn't know*.

It's a similar story with good and evil. God could have created a world in which there was no evil and *everything* was good. Like the red example, we wouldn't have created the concept of "good" because there would be no "evil" to contrast this concept against. Nevertheless, it would still be *true* that everything is good – we just wouldn't be aware of it. From God's perspective, for example, He would know that everything was good, even if humans were unaware.

We need *some* evil to maximise goodness

You could develop the "good couldn't exist without evil" response above to argue that *some* evil is necessary for certain (more valuable) types of good. For example, you couldn't be *courageous* (good) without having to overcome *fear* (evil) of pain, death, etc. Mackie sets up this argument by making the following distinction between *first order* goods and *second order* goods:

- First order good: E.g. pleasure and happiness
- Second order good: A good that requires evil/is a virtuous response to evil, e.g. courage and sympathy

The argument is that second order goods are *more good* than first order goods (a bit like the distinction between higher and lower pleasures that Mill makes in utilitarianism – page 93).

So, in order to maximise the total amount of good in the world, we need second order goods because they're worth more. But these second order goods couldn't exist without first order *evils*, and so we need some evil in order to maximise total good.

*Mackie's **response** to this problem 1:*

Mackie's first objection to this response is that the distinction between first and second order goods isn't really a proper distinction. The only reason we value second order goods (e.g. courage) is because they maximise first order goods (e.g. happiness). For example, we value the courage of someone who rescues a person from a burning building for the happiness it causes (or rather, the pain it avoids).

But if the value of second order goods is *grounded* in first order goods, then in what sense are they *more* valuable? The value of second order goods consists of the first order goods they cause and so, ultimately, first order goods are all that matter.

*Mackie's **response** to this problem 2:*

But Mackie's main objection to this response is that it only explains a need for first order evils, but not *second order* evils.

Just as second order goods maximise first order goods, second order evils maximise first order evils. For example, malevolence or cruelty would be examples of second order evils because they maximise first order evils such as pain and sadness.

So let's say we accept the argument above that first order evil is necessary for second order good to exist. How do we explain *second order* evil? God could have created a world in which people were courageous (second order good) in overcoming pain (first order evil) – *but where there was no second order evils.* Second order evils seemingly have no possible role in maximising good, so why would an omnipotent and omnibenevolent God allow them?

In other words, there is no greater good that results from allowing second order evils and so, if God existed, He would have created a world without them.

We need evil for free will

We can develop the second order evil argument above further and argue that second order evil is necessary for *free will.* And free will is inherently such a good and valuable thing that it outweighs the bad that results from people abusing free will to do evil things.

So, while allowing free will brings some suffering, the net good of having free will is greater than if we didn't. Therefore, it's logically possible that an omnipotent and omnibenevolent God would allow evil (both first order and second order) for the *greater good* of free will.

Mackie's ***response*** *to this problem:*

- An omnipotent God can create any logically possible world
- If it's logically possible to freely choose to act in a way that's good on one occasion, then it's logically possible to freely choose to act in a way that's good on every occasion
- So, an omnipotent God could create a world in which everyone *freely* chooses to act in a way that's good

In other words, there is a logically possible world with *both* free will *and* without second order evils. This, surely, would be the best of both worlds and maximise good most effectively: you would have second order goods, plus the good of free will, *but without second order evils*.

This is a logically possible world – *the logically possible world with the most good* – so why wouldn't an omnipotent and omniscient God create this specific world? Second order evils are seemingly not logically necessary, and yet they exist.

Alvin Plantinga: *God, Freedom and Evil*

Plantinga argues that we don't need a *plausible* theodicy to defeat the logical problem of evil. We don't need the *actual* explanation of why God allows evil. Instead, to defeat the logical problem of evil, all we need to show is that the existence of evil is not *logically* inconsistent with an omnipotent and omnibenevolent God.

Again, the logical problem of evil claims that the existence of evil shows that it is *logically impossible* for God to exist. And so, to disprove this claim, we only need to provide a logically possible (even if implausible) explanation of why God could allow evil.

The free will defence

Even Mackie himself admits that God's existence is not logically incompatible with *some* evil (first order evil). But his argument is that second order evil isn't necessary.

Plantinga argues, however, that it's *logically possible* (which is all we need to show to defeat the logical problem of evil) that God would allow second order evil for a greater good. His argument is as follows:

- A *morally significant action* is one that is either morally good or morally bad
- A being that is *significantly free* is one that is able to do or not do morally significant actions
- A being created by God to only do morally good actions would not be significantly free
- So, the only way God could eliminate evil (including second order evil) would be to eliminate significantly free beings
- But a world that contains significantly free beings is more good than a world that does not contain significantly free beings

In short, this argument shows that it's at least *logically possible* that God would allow second order evil for the greater good of *significant* freedom. Or, to put it another way, it's logically possible that our world is the world with the *most* good – because any reduction in evil would come at the cost of a reduction in free will.

Perhaps God could have created the world where everyone chose to only do morally good actions (as Mackie describes on page 178) – but such a world wouldn't be *significantly free*. Free will is inherently good and so significant free will could outweigh the negative of people using that significant free will to commit second order evils.

Natural evil as a form of moral evil

Plantinga's free will defence above explains why an omnipotent and omnibenevolent God would allow *moral* evil, such as malevolence and cruelty. But it doesn't explain *natural* evils.

When innocent people are killed in natural disasters, such as hurricanes and earthquakes, it doesn't seem this is the result of *free will* – it's not like we *choose* to have these natural disasters, they just happen. So, even if an omnipotent and omnibenevolent God would allow *moral* evil, why does He allow this kind of evil to exist as well?

In response to this question, Plantinga argues that it's *possible* that natural evil is the result of non-human actors such as Satan, fallen angels, demons, etc. This would make natural evil another form of *moral* evil, the existence of which would be explained by free will.

Even if this doesn't sound very *plausible*, it's at least *logically possible*. And, as mentioned earlier, Plantinga's argument is that we only need to show evil is not *logically* inconsistent with God's existence in order to defeat the logical problem of evil.

The *evidential* problem of evil

Unlike the logical problem of evil, the *evidential* problem of evil can allow that God's existence is logically possible. However, the evidential problem argues that the amount and distribution of evil in the world provides good evidence that God probably doesn't exist.

For example:

- Innocent babies born with painful congenital diseases
- The sheer number of people currently living in slavery, extreme poverty or fear
- The millions of innocent and anonymous people throughout history killed for no good reason
- Many good people have terrible luck, like getting cancer or killed in some freak accident
- Many evil people have good luck and live happy and long lives

We can reject the logical problem of evil and accept that God would allow *some* evil. But would an omnipotent and omnibenevolent God allow *so much* evil? And to people so undeserving of it?

The evidential problem of evil argues that if God did exist, there would be less evil and it would be less concentrated among those undeserving of it.

Responses to the evidential problem of evil (theodicies)

Free will (again)

Plantinga argues that the free will defence (page 179) works for the evidential problem of evil just as it does for the logical problem of evil. Sure, God *could* have made a world with less evil. But this would mean less free will. And on balance, having free will creates more good than the evil it also creates.

*Possible **response** to the free will defence against the evidential problem of evil:*

We may respond that the free will defence is not a very convincing response to the evidential problem of *natural* evil.

Plantinga's argument above was that it's logically possible that non-human actors are responsible for natural evil. If so, this is sufficient to beat the *logical* problem of evil – but it's not so convincing against the *evidential* problem of evil. Again, the evidential problem of evil claims that the amount and distribution of evil provides good evidence that God doesn't exist. And so the question becomes: How *likely* is it that Satan, demons, or other non-human entities are responsible for earthquakes, hurricanes, etc.? What physical evidence is there for this claim?

Without strong evidence for the existence of non-human entities, the proponent of the evidential problem of evil can stick to their claim that the amount and distribution of *natural* evil is strong evidence that God doesn't exist.

John Hick: Soul making

John Hick (1922-2012) responds to the evidential problem of evil by arguing that humans are unfinished beings and that part of our purpose in life is to develop personally, ethically and spiritually. Hick calls this development "soul making".

Earlier, we saw Mackie consider the argument that it would be impossible for people to display (second order) virtues such as courage without fear of (first order) evils such as pain or death (page 177). Similarly, we couldn't learn virtues such as forgiveness if people never treated us wrongly. Hick's argument here is very similar: we need certain evils in order to develop certain *virtues* in response to those evils.

Of course, God could just have *given* us these virtues right off the bat. But, Hick says, virtues acquired through hard work and discipline are "good in a richer and more valuable sense". Plus, there are some virtues, such as a genuine and authentic love of God, that cannot simply be given (because if God just made us to love him – like robots with no free will – then the love wouldn't be genuine).

This explanation goes some way towards explaining why God would allow the amount and distribution of evil we see. Hick then addresses some specific examples of evils that may not seem to fit with an omnipotent and omnibenevolent God:

- Why God allows *animals* to suffer
- Why God allows such *terrible* evils
- Why God allows such *pointless* evils

Firstly, animals: We can ask why God would allow animals to suffer when there is no benefit. After all, animals can't develop spiritually like we can – so why make them suffer? Hick's response is that God wanted to create a world in which his existence could be *doubted*. If God just proved He existed, we wouldn't be free to develop a relationship with him.

Secondly, Hick argues that it's not possible for God to just get rid of terrible evil – e.g. baby torture – and leave only ordinary evil. The reason for this is that terrible evils are only terrible in contrast to ordinary evils. So, if God did get rid of terrible evils (i.e. the 10 out of 10 evils), then the worst ordinary evils (i.e. the 9 out of 10 evils) would become *the* new terrible evils, and we could run the argument again. Further, if God kept getting rid of terrible evils, then He would have to keep reducing free will and thus the development of personal and spiritual virtues (soul making).

Thirdly, Hick considers pointless evils – e.g. anonymously dying in vain trying to save someone. Hick says such pointless evils are somewhat of a mystery. However, if every time we saw suffering we could see that it was for some higher purpose (i.e. we could see that it wasn't pointless), then we would never be able to develop virtues such as *deep sympathy*.

This all ties in with the soul making theodicy: Without seemingly unfair and pointless evils, we would never be able to develop virtues such as hope and faith – both of which require a degree of uncertainty.

*Possible **response** to the soul making argument against the evidential problem of evil:*

You could respond that Hick's soul making hypothesis is *unfalsifiable* (see page 185 for more details of what this means). In short, there is seemingly no possible distribution of evil that could count as evidence against God's existence because Hick could just claim it is necessary for soul making. This would make the soul making hypothesis meaningless.

RELIGIOUS LANGUAGE

OVERVIEW

Religious language in A-level philosophy looks at the meaning of religious statements, such as "God exists", "God answers my prayers", and "God loves us".

This topic is not about whether these statements are true or false. Instead, this topic is about whether such statements are **meaningful** or whether they are **meaningless** (in particular, whether "God exists" is a meaningful or meaningless claim).

The syllabus also mentions the distinction between cognitivist and non-cognitivist views of religious language.

COGNITIVISM AND NON-COGNITIVISM

All the arguments we've looked at so far (ontological, cosmological, teleological, problem of evil) assume a *cognitivist* view of religious language. These arguments treat "God exists" as a sort of factual statement that *literally* describes how the world is. These arguments then aim to show that this statement is either true or false.

But some philosophers argue that religious language is *non-cognitive*. This is to say that religious language is not to be taken literally as true or false in a similar way to how *moral* non-cognitivism (page 120) says moral judgements are not literally true or false.

Cognitive statements	Non-cognitive statements
Aim to literally describe how the world is	Do not aim to literally describe how the world is
Are true or false	Are neither true or false
Examples: • "Water boils at 100°c" • "Triangles have 3 sides" • "The Eiffel Tower is in Paris" • "The Eiffel Tower is in Berlin"	Examples: • "Ouch!" • "Boo!" • "Hooray!" • "Don't do that!"

Cognitivism is perhaps the common-sense view of religious language. According to cognitivism, when someone says "God exists", "God loves me", or "God answers my prayers", they are making a statement that is intended to be taken literally as true or false.

In contrast, non-cognitivist statements are neither true or false. In the context of religious language, a non-cognitivist might say religious statements like those above express someone's *attitude* to the world. For example, non-cognitivists may interpret "I believe in God" as an *expression* of a commitment to a certain set of values and a way of living.

ARGUMENTS THAT RELIGIOUS LANGUAGE IS MEANINGLESS

A.J. Ayer: Verification principle (again)

A.J. Ayer argues that religious language is *meaningless* because it fails the verification principle. The verification principle previously came up in metaethics (page 121) and says that a statement is only meaningful if it is:

- An analytic truth (e.g. "a triangle has 3 sides")
- Empirically verifiable (e.g. "water boils at 100°c")

Applying the verification principle to religious language, Ayer first argues that "God exists" is not an analytic truth because ontological arguments (see page 156) fail. Ayer further argues that religious statements are not empirically verifiable because they are not *falsifiable* (see below for more on what this means).

So, according to Ayer's verification principle, "God exists" (and religious language in general) is *meaningless* because it is neither an analytic truth nor empirically verifiable.

*Possible **response #1** to the verification principle: reject the verification principle*

The same responses to Ayer's verification principle from metaethics apply here too: firstly, you can argue that Ayer's verification principle fails its own test. Ayer's claim that "a statement is only meaningful if it is analytic or empirically verifiable" is *itself* neither an analytic truth or empirically verifiable. So, according to its own criteria, the verification principle is meaningless.

Like before, though, Ayer would respond that the verification principle is only intended as a *definition* of a meaningful statement and not a factual hypothesis about meaning. However, we can further respond by rejecting this definition of meaning. We might argue that there are examples of meaningful statements that do not fit this definition, such as Hare's "bliks" (page 189).

*Possible **response #2** to the verification principle: accept the terms of the verification principle*

Alternatively, we can accept the terms of the verification principle and argue that "God exists" is either an analytic truth or empirically verifiable:

- If ontological arguments (page 156) succeed, then "God exists" would be an analytic truth because "God exists" would be true by definition
- We may argue that "God exists" is a falsifiable claim (see below) that can be empirically or eschatologically verified (see page 187)

Falsifiability

Despite how it may sound, falsifiability is not about being true or false. Instead, falsifiability is part of what it takes for a statement to be *meaningful*.

- Falsifiable statements are meaningful and can be proved either true or false
- Unfalsifiable statements are meaningless and cannot be proved true or false

A statement is falsifiable if it is inconsistent with some possible observation. In other words, there has to be some *possible* evidence that could count against that statement otherwise the statement is meaningless.

For example, the statement "water boils at 100°c" *is* falsifiable because it could be proven wrong by some possible observation. If we heated a beaker of water, say, and it didn't boil despite a reliable thermometer showing 500°c, then this observation would count against the statement – it would *falsify* it. So, "water boils at 100°c" is a falsifiable and meaningful statement because there are possible tests that could prove it wrong.

In contrast, the statement "everything in the universe doubles in size every 10 seconds" is *unfalsifiable* because no possible observation could disprove it. You could use a ruler to measure things every 10 seconds but this couldn't disprove the claim because you could just say that the ruler doubled in size as well! There is seemingly *no possible experiment* that could disprove the statement "everything in the universe doubles in size every 10 seconds" and so it is unfalsifiable and meaningless.

In the context of religious language, we could argue that the following young Earth creationist hypothesis is an example of an unfalsifiable theory:

> *Young Earth creationist: "The earth is 10,000 years old"*
>
> *Archaeologist: "But I've found a fossil that carbon dating suggests is more than 1,000,000 years old"*
>
> *Young Earth creationist: "God put that fossil there on purpose to trick you into thinking the earth is older than it really is"*

This young Earth creationist theory cannot possibly be proven wrong. No matter how many fossils you find, or how much evidence you accumulate, it can't count as evidence against this theory because the theory can just say God planted the fossils there as a trick.

Anthony Flew: Invisible gardener

Anthony Flew (1923-2010) gives the following parable in an attempt to show that religious language – in particular the statement "God exists" – is unfalsifiable, and thus meaningless:

- Two explorers find a clearing in a jungle. Both weeds and flowers grow there
- Explorer A says the clearing is the work of a *gardener*. Explorer B disagrees
- To settle the argument, they keep watch for the gardener
- After a few days, they haven't seen a gardener, but Explorer A says it's because he is invisible
- So, they set up an electric fence and guard dogs to catch the gardener instead
- But, after a few more days, they still haven't detected him
- Explorer A then says that not only is the gardener invisible, he's also intangible, makes no sound, has no smell, etc.
- Explorer B asks: What is the difference between this claim and the claim that the gardener doesn't even exist?

In other words, Explorer A's theory is *unfalsifiable* – nothing could possibly prove this theory wrong, because every possible observation to disprove the theory can be explained away. But a theory that cannot possibly be proven wrong cannot be proven right either. And so, Explorer A's theory is unfalsifiable and *meaningless*.

Flew's parable here is an analogy for religious belief:

- Jungle clearing = the world
- Invisible gardener = God
- Flowers = Good
- Weeds = Evil

Flew is arguing that "God exists" is as meaningless as the claim that an invisible and intangible gardener manages the jungle clearing: Flew says we can't see, hear, or feel God, just as we can't see, hear, or otherwise detect the gardener.

Further, if there was a gardener, you would at least expect to see evidence of his work – for example, you would expect him to have removed the weeds from the clearing and left the flowers. The parallel with God is this: If God existed, you would expect Him to have

removed evil from the world. This last point is discussed in the problem of evil (page 175).

Flew's argument is that the statement "God exists" is meaningless in the same way that Explorer A's theory of the invisible gardener is meaningless.

ARGUMENTS THAT RELIGIOUS LANGUAGE IS MEANINGFUL

Hick: Eschatological verification

Eschatological verification: A statement that can be verified after death, or at the end of time.

John Hick agrees with Ayer that "God exists" is not empirically verifiable *in this life*. However, he argues that many religious claims are about things *beyond* the limits of human life. And, he argues, such claims *are* meaningful because it is possible to verify them after we die. For example, many theists believe in a life after death during which they will meet or otherwise experience God (which would be unambiguous verification that "God exists" is true).

To illustrate this, Hick tells a parable of a "celestial city":

- Two men are travelling on a road – it is the only road there is, so they both must travel it
- Traveller A believes the road leads to a celestial city, whereas Traveller B believes the road leads nowhere and that the journey is meaningless
- As they travel along the road, they experience both "refreshment and delight" and "hardship and danger"
- If Traveller A is correct, they will eventually arrive at the celestial city and he will be proved right
- If Traveller B is correct, they will just keep going forever, and neither will be proved right

In this parable, Traveller A is the theist and Traveller B is the atheist. The "hardship and danger" represents the problem of evil (page 175). If the theist is correct, his belief will be *verified* in the afterlife when he meets God - this is the equivalent of reaching the celestial city. However, if the atheist is correct, his belief will never be verified (and cannot be falsified) because he'll be dead and unable to verify anything – this is the equivalent of walking forever.



So, in short, Hick says "God exists" is not meaningless because it is eschatologically verifiable:

- If "God exists" is false, then it is unfalsifiable
- But if "God exists" is true, then it can be verified after we die

Basil Mitchell: Resistance fighter

Basil Mitchell (1917-2011) agrees with Flew and Ayer that in order for a statement or belief to be meaningful it must be possible for some observation to count against it (i.e. it must be falsifiable in order to be meaningful).

But, Mitchell argues, just because there are some observations that count against a certain belief, that doesn't automatically mean we have to withdraw from that belief. Mitchell gives the following example to illustrate this:

- You are in a war, your country has been occupied by an enemy
- You meet a stranger who claims to be leader of the resistance
- You trust this man
- But the stranger acts ambiguously, sometimes doing things that appear to support the enemy rather than your own side
- Yet you continue to believe the stranger is on your side despite this and trust that he has good reasons for these ambiguous actions

In this analogy, the stranger represents God and his ambiguous actions represent the problem of evil (page 175). Mitchell is arguing that we can accept that the existence of evil counts as evidence against the statement "God exists" (and so it is falsifiable) without having to withdraw from belief in this statement.

Mitchell argues that religious beliefs are not "provisional hypotheses" like scientific statements. They are not beliefs that the believer is totally detached from and will just abandon as soon as evidence against them turns up. But nor are religious beliefs "vacuous formulae" that the believer holds regardless of *any* evidence to the contrary.

Instead, reasonable religious beliefs are a kind of middle ground: they are *"significant articles of faith"*. The religious believer is *invested* in these beliefs and so doesn't withdraw from them as soon as the slightest evidence to the contrary turns up. However, this is not to say that the religious believer would believe "God exists" in spite of *overwhelming* evidence to the contrary - that would be irrational or, to use Mitchell's term, a "vacuous formulae".

So, in short, Mitchell says:

- We can accept that the existence of evil counts as evidence against God's existence
- And so, "God exists" is falsifiable
- And so, "God exists" is meaningful
- But the religious believer can accept there is some evidence against the statement "God exists" without having to withdraw belief that this statement is true

R.M. Hare: Bliks

According to Hare, religious statements are not things that can just be shown to be true or false. Instead, they are part of someone's view of the world – Hare calls these attitudes "bliks".

To illustrate what bliks are, Hare comes up with an example of a paranoid student who has a belief that university lecturers are trying to kill him:

You assure this student that university lecturers are not trying to kill him and provide tons of evidence, yet the student still believes it anyway. Imagine, for example, that to prove your point you take him to speak to a university lecturer and the lecturer acts totally normal:

> *You: See, he's fine – the university lecturer isn't trying to kill you!*
>
> *Paranoid student: But he was just pretending to be normal so as not to reveal his true plan to kill me!*

So, no amount of evidence/reassurance will convince the student that his blik is false. In other words, their blik is *unfalsifiable*.

But despite being unfalsifiable, Hare argues that bliks are still *meaningful* to the person who holds them. In the case of the paranoid student example, the blik clearly means something to the paranoid student because it has an effect on his behaviour: he won't go to lectures, and will look over his shoulder to check university lecturers aren't following him, for example.

Hare argues that religious language is the same: "God exists" may be unfalsifiable to people who have this blik, but it clearly means something to them. For example, people who believe "God exists" might pray and go to Church – it means enough to them that it affects their behaviour.

In other words, a blik is unfalsifiable *but still meaningful* to the person who holds it.

METAPHYSICS OF GOD – MODULE SUMMARY

The concept of God

- Omnipotent: The most *powerful* being possible
- Omniscient: *Knows* everything
- Omnibenevolent: Perfectly *good*
- Eternal (existing outside of time) *or* everlasting (existing within time since the beginning and will continue to exist forever)
- Problems:
  - The problem of the stone: Can God create a stone so heavy He can't lift it? Either way, there's something God can't do, so He's not omnipotent
    - Reply: A stone an omnipotent being can't lift is a *contradiction*. God can do anything that's *logically* possible
  - Euthyphro dilemma: Either God created morality, so "God is good" is a tautology, or God has to *follow* moral laws, so God is not omnipotent
    - Reply: God created morality based on His other attributes, such as love
  - Omniscience vs Free Will: If God is omniscient, He knows what I'm going to do next. But if what I'm going to do next is already known, then I'm not free
    - Replies:
      - The future is impossible to know. An *everlasting* God's omniscience only means He knows everything that is *possible* to know
      - An *eternal* God is outside of time and so is observing (and thus knows) the past, present, and future all at once

Ontological arguments: A priori arguments that aim to prove God exists from the definition of God

- Anselm: God = a being greater than which cannot be conceived, it is greater to exist in reality than in the mind, therefore God exists
- Descartes: God = supremely perfect being, existence is a perfection, therefore God exists
- Replies:
  - Gaunilo's island: The perfect *island* is an island greater than which cannot be conceived… Therefore the perfect island exists

- Hume: If ontological arguments work, then "God does not exist" is a contradiction. But we can coherently conceive of God *not* existing, so it is *not* a contradiction. So, ontological arguments must fail somewhere
- Kant: existence is not a predicate: the *concept* of money *that exists* buys no more in reality than the *concept* of imaginary money does
  - Reply: Malcolm's Ontological argument (*necessary* existence, which is a predicate):
    - Either God exists or He doesn't
    - If God doesn't exist, then it's impossible for God to exist
    - If God exists, then God exists necessarily
    - But it's not impossible for God to exist
    - Therefore, God exists necessarily

Teleological arguments: Many things in nature appear to be designed, and that designer is God

- *Spatial* order teleological arguments:
  - Paley: Things with "multiple parts organised for a purpose" = hallmark of design (compare watch vs. pebble). Things in nature contain multiple parts organised for a purpose, and thus are designed (by God)
  - Hume: The "fitting of means to ends" in nature resemble the "fitting of means to ends" in human design. Similar effects have similar causes and so, by analogy, the cause of design in nature is also a mind (God)
  - Replies:
    - Hume: Problems with the analogy:
      - Nature and man made things are very different
      - We can have observe man-made items being designed, but we have no such experience of this in the case of nature
      - Even if there are examples of design *within* the universe, we cannot generalise this to the claim that universe *itself* is designed
    - Spatial disorder: There are huge areas of the universe that are empty, or just filled with rocks etc. Doesn't seem very designed
    - Hume: Causal relationships require multiple examples to establish relationship. The creation of the universe only happened once
    - Hume: If time is infinite and matter is finite then it is inevitable that

matter will organise itself in ways that *appear* to be designed
    - Darwin: Evolution by natural selection can explain how nature can create complex things without being designed
- Swinburne's teleological argument from *temporal* order: The laws of the universe are so perfect that they appear to be designed. E.g. if gravity was stronger/weaker, planets couldn't form. The best explanation for this is a personal one: God designed them that way.
    - Replies:
        - If there are multiple universes, it is inevitable that some of these universes will have laws that sustain life
        - The designer might not be God (as defined in the concept of God)

Cosmological arguments: two types: 1. Causation: God *caused* the universe to exist, and 2. Contingency: A necessary being (God) is needed to explain contingent truths

- Cosmological arguments from *causation*:
    - Kalam argument: Everything that begins to exist has a cause, the universe began to exist, therefore the universe has a cause (this cause is God)
    - Aquinas' first way: A moves B, etc. If there was no first mover there would be no subsequent movement. So, there must be a first mover (God)
    - Aquinas' second way: A causes B (effect) etc. If there was no first cause there would be no subsequent effects, so there must be a first cause i.e. a cause that caused all subsequent effects (God)
    - Descartes: I am not the cause of my existence, I am a thinking thing with the idea of God, the cause of my existence must also be a thinking thing with the idea of God…
    - Replies:
        - Hume: "everything has a cause" is not an analytic truth
        - It's possible there's just an infinite sequence of causes
        - Even if there is a first cause, this first cause might not be God
- Cosmological arguments from *contingency*:
    - Aquinas' third way: Everything (in the universe) exists contingently. If it's possible for something not to exist, then at some point it didn't exist. If nothing existed then nothing could begin to exist. Therefore there must be something that exists necessarily

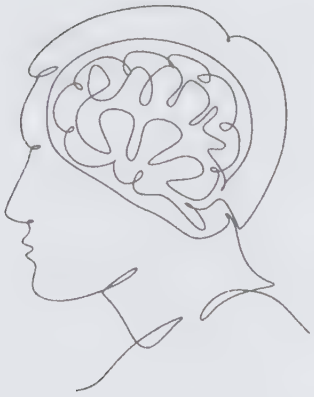
- Leibniz's principle of sufficient reason: Every truth has an explanation of *why* it is the case (even if we can't know this explanation). Contingent truths can only be explained by other contingent truths, which also need explanation with contingent truths, and so on. We need to step outside the sequence of contingent facts and appeal to a necessary substance to provide sufficient reason for contingent truths
- Replies:
  - Russell: Fallacy of composition: Just because things *within* the universe require sufficient reason to explain why they exist, doesn't mean the universe itself needs such an explanation
  - Hume: The impossibility of a necessary being: If God exists necessarily, then "God does not exist" would be a contradiction. But "God does not exist" is not a contradiction as evidenced by the fact that whatever we can conceive of as existing we can also conceive of it not existing

The problem of evil

- Logical problem of evil: Inconsistent triad: the existence of God is logically impossible with an omnipotent and omnibenevolent being
  - Replies:
    - The world is better with *some* evil (you couldn't have *second order* goods without first order evils)
    - Without the ability to be evil, we couldn't have free will
      - Reply: But free will doesn't explain *natural* evils (e.g. earthquakes)
        - Reply to reply: It's logically possible (which is all we need to show to defeat the logical problem of evil) that natural evils are a form of *moral* evil committed by non-human agents (e.g. demons), so are covered by the free will explanation above
- Evidential problem of evil: Evil is not *logically* incompatible with God, but the amount of evil and how it is distributed suggests God probably doesn't exist
  - Replies:
    - Soul making: Evil is needed to develop morally and spiritually. Pointless evils are needed to develop virtues such as faith and deep sympathy
    - Terrible evil is relative. If God got rid of the worst evils (i.e. the 10/10s), then the second worst evils (i.e. the 9/10s) would become terrible

Religious language:

- Cognitivism vs. non cognitivism:
  - Cognitivism: "God exists" is literally either true or false
  - Non-cognitivism: "God exists" is more an expression of an attitude to the world, rather than a scientific or empirical claim that is true/false
- Arguments that religious language is *meaningless*:
  - Ayer's verificationism: Statements are only meaningful if they are analytic or empirically verifiable – "God exists" is neither, so it's meaningless
  - "God exists" is unfalsifiable: It can't possibly be proved or disproved, and is therefore meaningless
    - Flew's jungle clearing analogy: "God exists" is like saying an invisible and undetectable gardener is growing both weeds and flowers in some scruffy area of the jungle
    - Replies/arguments that religious language is *meaningful*:
      - Hick's Eschatological verification: "God exists" is unfalsifiable *if false*, but verifiable *if true* (because when we die we would meet God, which would verify "God exists" is true)
      - Mitchell's Ambiguous resistance fighter analogy: The problem of evil is evidence that "God exists" is false, and so "God exists" is falsifiable. However, *some* (inconclusive) evidence against "God exists", doesn't mean you have to withdraw belief
      - Hare's Bliks: Evidence can't change bliks, but bliks are still meaningful to the person who has a blik (because e.g. it changes their behaviour). For example, people who have the blik that "God exists" might go to church and pray and so it clearly means something to them



METAPHYSICS OF MIND

This module looks at what *mental states* - such as "pain" , "happy", and "sad" – and *minds* actually are. We look at six competing theories of the mind. These six theories can be divided into two categories: theories that say the mind is a *physical* thing and theories which say it is a *non-physical* thing.

DUALIST THEORIES

OVERVIEW

Dualist theories of the mind argue that there are two kinds of thing – hence, dualism. The A-level philosophy syllabus lists two forms of dualism:

1. **Substance dualism**
2. **Property dualism**

The syllabus lists specific arguments for each version as well as several responses. There are also various arguments against each theory, and against dualism in general. The key points are as follows:

	Substance dualism	Property dualism
Summary	The mind is a completely separate *substance* to physical things	There are mental *properties* (qualia) that are neither physical nor supervenient on the physical
Arguments for	Conceivability argument Divisibility argument	Zombie argument Knowledge argument
Arguments against	Interaction problem Problem of other minds	Introspective self-knowledge Evolutionary role of qualia

SUBSTANCE DUALISM

Substance dualism says there are two completely different kinds of substance in our universe:

- Mental substances
- Physical substances

Physical substances are things like the trees, cars, houses, etc. Your body – your arms, legs, etc. – is a physical thing as well. The *brain* is part of the body, hence it is also physical.

But dualists deny that the *mind* is the same thing as the brain. Instead, dualists argue that the mind is something completely different to the brain – something *non-physical*.

Most people who believe in souls are dualists. Typically, the soul isn't thought of as a physical thing – you can't touch it, or see it, for example. Similarly, things like ghosts and angels are somewhat dualist because they're not thought to be made of physical stuff either. Like these examples, substance dualists believe that the mind is similarly non-physical – it's a separate kind of substance altogether.

Descartes' arguments for substance dualism

The syllabus lists two arguments from Descartes that argue that mind and body are separate substances: the conceivability argument and the divisibility argument.

The conceivability argument

Descartes' conceivability argument relies somewhat on the notion of *clear and distinct ideas* that we covered in epistemology (page 66). The argument can be summarised as:

- I have a clear and distinct idea of my mind as a thinking thing that is *not* extended in space
- I have a clear and distinct idea of my body as a non-thinking thing that *is* extended in space
- Anything I can conceive of clearly and distinctly is something that God could create
- So, God could create my mind as a thinking thing that is not extended in space and my body as a non-thinking thing that is extended in space
- So, it is possible for mind and body to exist independently of each other
- So, mind and body are two separate substances

In other words, Descartes is saying that it is conceivable – and therefore possible – for mind and body to exist separately. He then takes this to show that mind and body *are* separate things.

This last step might seem like a bit of a leap in logic: Just because it's *possible* for mind and body to be separate, that doesn't guarantee that's how it *actually* is. So, in defence of Descartes (before we press this objection further), he does say he has *clear and distinct ideas* of mind and body as separate. For Descartes, this means these ideas are trustworthy, complete, and *true* rather than just random thoughts that may or may not be accurate.

Responses to the conceivability argument

Mind without body is not conceivable

One way we can respond to the conceivability argument is by rejecting the idea that mind without body is even conceivable.

Behaviourism (page 223) analyses mental states as behaviours and behavioural dispositions. This means that the mind is inextricably linked with behaviour – to have a mind/mental states *means* to have certain behavioural dispositions. But you can't have behaviours or behavioural dispositions if you don't have a body.

And so, if behaviourism is correct, it's impossible to have a mind (or mental states) without a body: mind without body would be like having behavioural dispositions without having behavioural dispositions – a contradiction. And contradictions are *inconceivable*, and so mind without body is inconceivable.

In other words:

- Behaviourism says that to have mental states is to have behavioural dispositions
- To have behavioural dispositions is to be disposed to move your body in certain ways
- It is inconceivable to be disposed to move your body in certain ways if you don't have a body
- So, it is inconceivable to have mental states if you don't have a body
- So, mind without body is inconceivable

*Possible **response** to the claim that mind without body is not conceivable:*

This whole response assumes that behaviourism is the correct account of the mind. If behaviourism is wrong, then there is no contradiction in the idea of mental states without behavioural dispositions, and so we can conceive of disembodied thoughts.

What is conceivable may not be possible

Another response to the conceivability argument attacks Descartes' inference from the claim that mind without body is *conceivable* to the conclusion that mind exists without body.

The fallibility of this inference can be shown with the following example:

1. I conceive of Batman as a caped crusader
2. I conceive of Bruce Wayne as a billionaire who is not a caped crusader
3. Therefore, Batman is not Bruce Wayne

If you know the Batman story, Batman is Bruce Wayne's secret identity. So the conclusion is clearly false: Batman *is* Bruce Wayne.

And if Batman is, in fact, Bruce Wayne, then it's not possible for Batman to be someone else. No one can be someone else – that would be a contradiction – and so it's not *possible* for Batman to be anyone other than Bruce Wayne. Even though it may be *conceivable* that Batman is someone else, the reality is that it isn't possible for Batman to be anyone other than Bruce Wayne. To think otherwise is to fall for the masked man fallacy.



It's conceivable that these are two different people, but it may not be possible

The reason this argument is fallacious is because it switches from talking about *ideas* to talking about things *themselves*. But sometimes our ideas are mistaken – what we have an idea of as possible is, in fact, not possible.

We can argue that Descartes' conceivability argument makes the same mistake as the Batman example:

- Just because you have an *idea* of Batman and Bruce Wayne as separate people, this doesn't mean it's *possible* they are separate people.
- Similarly, just because you have an *idea* that the mind and body are separate things, this doesn't necessarily mean it's *possible* they are separate things

*Possible **response** to the masked man fallacy:*

Descartes would likely accept that you can't always infer from ideas to reality. But he could respond that there's an important difference between the two examples. In the Batman/Bruce Wayne example, my ideas are not *clear and distinct.*

A clear and distinct idea, if you remember from epistemology (page 66), is one that the thinker can see is true through the use of rational intuition – the thinker can *see* it is true and certain. Descartes would claim that merely reflecting on the proposition "my mind is a thinking thing that is not extended in space" reveals it is *true*, whereas this is not the case with the proposition "Bruce Wayne is a billionaire who is not a caped crusader".

What is possible tells us nothing about the actual world

But even if we accept that mind without body is possible, we can still reject Descartes' conclusion that substance dualism is true. Just because something is possible, this doesn't tell us anything about reality. For example, it's logically possible for me to jump to the moon from earth – because there is no logical contradiction in this idea. But just because it is possible for me to jump on to the moon from earth, this definitely doesn't tell us anything about the actual world! We can argue that it's a similar thing with Descartes' conceivability argument: Even if it's *possible* that my mind and body are separate substances, this doesn't show that my mind and body are separate substances *in reality*. It's equally possible that my mind and body are the same substance.



The divisibility argument

Descartes' second argument for substance dualism is the divisibility argument. The divisibility argument can be summarised as:

1. My body is a divisible substance
2. My mind is an indivisible substance
3. Therefore, my mind and body are separate substances

There is a fourth – *implied* – premise here too: *Leibniz's law of indiscernibility of identicals*. Leibniz's law says that if two things are the same thing, then they will have all the same properties. For example, water has the property of wetness and so does H2O, because water and H2O are the same thing.

The converse of Leibniz's law is also true: if two things have *different* properties, then they can't be the same thing. For example: "that guy can't be Jones because that guy has black hair and Jones has blond hair". In this example, Jones has the *property* of having blond hair but "that guy" has the property of having black hair. And so (assuming Leibniz's law), "that guy" and Jones cannot be the same person.

Descartes is making a similar argument with the mind and the physical body: the body has the property of being divisible, but the mind does not have this property. And so Descartes takes this to show that the mind and the body cannot be the same thing.

Descartes' first premise seems pretty uncontroversial. If you chopped you leg or your arm off, you would be dividing up your body, and so it definitely seems possible to divide the body. And the second premise also seems plausible: the mind does not seem divisible – at least not in the same way that physical stuff such as the body is. You can't *literally* have half a thought or half a mind, for example. It seems that you either have a complete mind or you don't – you can't split it up into pieces like the body.

Responses to the divisibility argument

The mind is divisible

One way we can respond to Descartes' divisibility argument is to reject the second premise. We can argue that the mind *is,* in fact, divisible in some way.

For example, there are forms of mental illness where the mind does literally seem divided, such as multiple personality disorder. In cases of multiple personality disorder, one person can have parts of their mind that seem to be completely separate from another. These different parts may have completely different memories, beliefs, and personality traits. The different parts may even be completely unaware of the other parts.

Another example of a divided mind could be patients who have undergone a *corpus callosotomy*. A corpus callosotomy is a surgical procedure (sometimes used to treat epilepsy) where the main connection between the left and right hemispheres of the brain – the corpus callosum – is severed. In other words, this procedure involves literally cutting the brain in half. Perhaps surprisingly, these split-brain patients often go on to live pretty normal lives. However, there are some potential side effects that appear to suggest the *mind* also becomes divided in some way. For example, studies of split-brain patients suggest that the separated right and left hemispheres can have differing personalities, tastes, and beliefs.

These examples suggest that the mind is divisible in some sense and so the second premise of Descartes' argument appears to be false.

*Possible **response** to the claim that the mind is divisible:*

However, the dualist can respond that even if the mind is in some sense divisible, it is not divisible *in the same way* as the body. The physical body is *spatially* divisible – for example, you can cut off a finger and *move* it to a different spatial location. However, the mind is not spatially divisible in this way – you can't separate two different thoughts and put them in two different locations. Instead, you could say that the examples of divided minds above are examples of minds being *functionally* divisible.

Not everything that is physical is divisible

We can also respond to the divisibility argument by arguing that even if the mind is indivisible, this doesn't necessarily prove Descartes' conclusion that the mind is a separate kind of *substance.* Instead, it's possible that the mind is just an indivisible type of physical substance.

Obviously, you can divide the physical body: the examples above of cutting off a leg or an arm show this. But if you *keep* dividing it, you might eventually reach a point where you cannot divide it any further. Eventually, for example, you'll just be left with a load of atoms, and perhaps those atoms could be divided into sub-atomic particles, but eventually you might reach a form of physical substance that is indivisible (units of energy, say, or whatever physics says – the specifics don't really matter).

The point is this: If it's possible to reach a point where physical matter becomes indivisible, then not everything that is indivisible is non-physical. And so, even if Descartes successfully shows that the mind is indivisible, this doesn't prove that the mind is non-physical. It's *possible* that the mind is the same *kind* of substance as the body (i.e. a physical substance) but just an indivisible form of that same substance.

Responses to substance dualism (in general)

The problem of other minds

The problem of other minds is the problem of how we can know other people have minds. It's very similar to the issue of solipsism we looked at in idealism (page 53).

We have experience of *our own* minds – our own thoughts, sensations, and feelings. But when it comes to *other* people, we never experience their actual minds – we only experience the behaviour of their (physical) body. This raises the question of how to know whether there is a mind "attached" to that body at all.

It's particularly difficult for substance dualism to respond to the problem of other minds because substance dualism claims that minds and bodies are completely *separate* substances. This means it's completely possible, on the dualist view, to have physical behaviour without a physical mind.

So, if substance dualism is true, it's impossible to know whether other minds exist.

*Possible **response #1** to the problem of other minds: Mill's argument from analogy*

John Stuart Mill gives the following "common sense" response to the problem of other minds:

- I have a mind
- My mind causes my behaviour
- Other people have bodies and behave similarly to me in similar situations
- By analogy, their behaviour has the same type of cause as my behaviour: a mind
- Therefore, other people have minds

However, solipsism can respond to Mill that *one* example of a relationship between mind and behaviour (i.e. my own) is not sufficient to prove the relationship holds in all cases. It would be like saying "that dog has 3 legs, therefore all dogs have 3 legs."

In other words, it's a dubious inference to go from *one instance* of a relationship (I have a mind that causes my behaviour) to the claim that this relationship holds in *all instances* (everyone has a mind that causes their behaviour).

*Possible **response #2** to the problem of other minds: Other minds are the best hypothesis*

Another way to respond is to argue that even though we can't prove the existence of other minds, the existence of other minds is the best hypothesis to explain human behaviour.

In science, it is common to accept as true the hypothesis that provides the best available explanation of a particular phenomenon at a given time. For example, scientists have believed in the existence of theoretical entities – such as some subatomic particles –

before actually observing them because positing such entities has strong explanatory power. In other words, even if you can't prove these theoretical entities exist (through observation, say), it is still reasonable to believe in them if they explain some other observation.

We can argue that minds play a similar role to theoretical entities in science. Yes, we can't directly observe other people's minds, just like we can't directly observe other theoretical entities in science. However, minds explain why humans behave in the ways that they do – in fact, they're the *best* explanation of why humans behave as they do. So, it is reasonable to believe in the existence of minds.

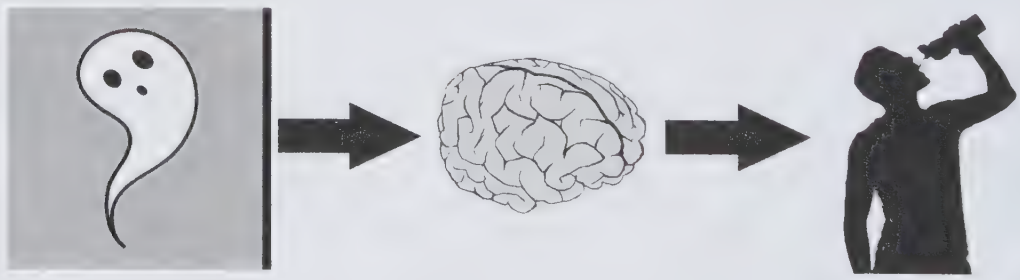
Causal interaction

Note: The following two arguments apply to interactionist (rather than epiphenomenalist – see page 209) dualism specifically, because they target the claim that mental substances cause changes in physical substances, which epiphenomenalism does not claim.

Version 1: The *conceptual* interaction problem

A general issue for substance dualism is how mental things can causally interact with physical things when they are supposed to be two completely separate substances.

Our mental states affect how we behave, for example. If I'm feeling hungry (a non-physical mental state), it might cause me to move my body (a physical thing) to the fridge to get some food. But how does the *non-physical* substance (my mental state) somehow transfer over into the *physical* world and cause things to happen there?



How does the non-physical mental state of being thirsty (left) cross over from the non-physical world into the physical world and cause changes in the physical brain and behaviour?

It's easy to explain how physical things interact with *other* physical things. When you kick a football, for example, the atoms of your leg (a physical thing) connect with the atoms of the ball (also a physical thing) and the ball moves. There's no mystery here. But it's hard

to see how a non-physical thing would interact with a physical thing in the same way. It would be like punching a ghost: your hand would go straight through!

This objection was put to Descartes himself by his student, Princess Elisabeth of Bohemia (1618-1680):

1. Physical things only move if they are pushed
2. Only something that is physical and can touch the thing that is moved can push it
3. If substance dualism is true, then the mind is not physical and so it can't touch the body
4. Therefore, either substance dualism is false, or the mind cannot move the body
5. The mind can move the body
6. Therefore, substance dualism is false

*Possible **response** to the conceptual interaction problem:*

Descartes did provide *some* answers in response to the conceptual interaction problem. For example, in *Passions of the Soul,* he says:

> *"And the activity of the soul consists entirely in the fact that simply by willing something it brings it about that the little gland to which it is closely joined moves in the manner required to produce the effect corresponding to this volition."*

The "little gland" Descartes is referring to here is now referred to as the pineal gland, and so here Descartes can be interpreted as providing a scientific account of how the mind ("soul") moves the body. This isn't exactly medically substantiated, though, and even then it still doesn't explain *how* actions of the non-physical mind (i.e. "willing") cause this physical gland to move. The exact interpretation of Descartes' writings on the interaction problem is still a matter of debate, though.

Version 2: The *empirical* interaction problem

We can also get at the causal interaction issue in an empirical way by appealing to the scientific law of conservation of energy:

- The law of conservation of energy says that: In a closed system, energy cannot be added or removed – it can only be transferred
- Our universe is such a closed system
- If substance dualism is true, it would mean energy is constantly being added into the closed system of our universe every time we think
- So, if substance dualism is true, the law of conservation of energy is false

- But there is a lot of evidence (e.g. Noether's theorem) to suggest that the law of conservation of energy is true
- So, substance dualism must be wrong

PROPERTY DUALISM

The technical definition of property dualism is the view that there are at least some mental properties that are neither reducible to nor supervenient upon physical properties. In plain English, this means that even if two things had exactly the same physical properties, they wouldn't necessarily have the same *mental* properties.

Property dualists believe that there is just one type of substance (physical), but that physical substances can have *non-physical* properties. These non-physical properties are mental properties known as *qualia*. So, property dualism says that physical brains can have non-physical properties called qualia.

Qualia

Qualia are the subjective qualities of experience – i.e. what something feels like inside. Examples of qualia include:

- The redness of my experience when I look at a ripe tomato
- The taste of beer when I have a drink
- The rough feeling I experience when I run my hand over some sandpaper

Notice how qualia are not properties of the *objects themselves* – i.e. properties of the tomato, or the beer, or the sandpaper – they are properties of my *experience* of those objects.

Knowledge of qualia is sometimes called *phenomenal knowledge*.

Supervenience

Supervenience is a relationship between two kinds of thing. If something supervenes on something else, then it is *dependent* on that thing.

Physicalism (page 222) says that everything – including the mind and mental states – is either physical or supervenes on the physical. In other words, physicalism says that two physically identical things must be the same in every way: it's *impossible* for two physically identical things to be different.

But property dualism denies this claim. According to property dualism, it's possible for two physically identical things to be *different* in some way. More specifically, property dualism says it's possible that two physically identical things could have different mental properties – *different qualia*. So, property dualists claim that qualia are neither physical nor supervene on the physical.

Interactionism and epiphenomenalism

It's worth introducing another distinction at this point between epiphenomenalist and interactionist dualism:

- *Interactionist* dualism says the mind can interact with the physical world (e.g. mental state of hunger causes you to go and get food) and the physical world can interact with the mind (e.g. getting hit in the head causes the mental state of pain). In other words, the mental and physical can interact in *both directions*.
- *Epiphenomenalist* dualism says the physical world can cause mental states (e.g. getting hit in the head causes the mental state of pain) but mental states cannot cause changes in the physical world. In other words, the physical and mental only interact in *one direction:* Physical to mental.

Epiphenomenalism is a type of *property* dualism. Some (most) property dualists are epiphenomenalists and, for the purposes of A-level philosophy, you can think of epiphenomenalism as the same thing as property dualism (and interactionism as the same thing as substance dualism).

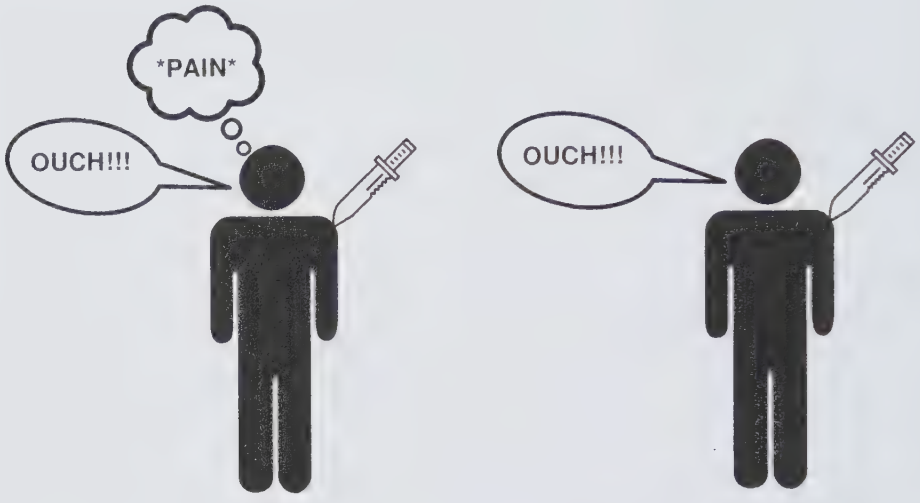
Epiphenomenal dualism avoids the causal interaction issues facing *substance* dualism (page 206) because it does not have to explain how the mental can cause changes in the physical – epiphenomenalism says that qualia are causally redundant.

David Chalmers: The zombie argument for property dualism

A philosophical zombie is a person who is physically, functionally, and behaviourally identical to an ordinary human – *except they don't have any qualia.*

A zombie will say "ouch!" when it gets stabbed and its physical brain will even light up and fire in the same way as a normal brain would – but there isn't any pain qualia associated with this. *Externally*, the zombie is physically identical in every way but *internally* there is nothing – no qualia.

Such zombies seem conceivable. We can imagine a possible world that is physically identical to this one, with the same people, but without qualia. In this world, you would behave and act in exactly the same way as in the actual world except you'd have no qualia/phenomenal experience.



The philosophical zombie (right) behaves and is physically identical to a normal human (left), except it doesn't have qualia

We can use the intuition that zombies are conceivable to form an argument for property dualism similar to Descartes' conceivability argument for *substance* dualism (page 199):

1. Philosophical zombies are conceivable
2. If philosophical zombies are conceivable then philosophical zombies are possible
3. If philosophical zombies are possible then qualia are non-physical (and do not supervene on the physical)
4. If qualia are non-physical (and do not supervene on the physical) then property dualism is true
5. Therefore, property dualism is true

Responses to the zombie argument

Zombies are not conceivable

Physicalists can respond by rejecting the first premise of the zombie argument: that philosophical zombies are conceivable.

We may argue that if we had enough physical knowledge then we would be able to understand what we currently call "qualia" in purely physical terms. In other words, the only reason zombies *seem* conceivable is because we are confused or missing some

important information. The conceivability of a physical duplicate without qualia is just an illusion – albeit a very powerful one.

The reason zombies seem conceivable is because we're labouring under a false illusion that qualia are these spooky non-physical things. Once we understand that qualia are, in fact, just physical things, then it becomes *inconceivable* to imagine a physically identical being that lacks these physical features. Imagining a philosophical zombie would be like saying "imagine something that is physically identical but that isn't physically identical" – it would be a contradiction, and contradictions aren't conceivable. It would be like trying to imagine a married bachelor or a triangle with 4 sides.

Once we understand that qualia = a physical thing, it becomes inconceivable for two physically identical beings not to have identical qualia, and so the zombie argument fails to prove property dualism.

*Possible **response** to the claim that zombies are not conceivable:*

This response assumes that a complete physical analysis of qualia exists. It assumes that, if we did have all the physical facts, we would understand that qualia is inseparable from certain physical things and thus that philosophical zombies are inconceivable.

However, this claim implies that knowing all the physical facts about seeing red, for example, would be sufficient to know what red qualia is like. This possibility is challenged by the knowledge argument for property dualism (page 213), though.

Zombies are metaphysically impossible

In addition to logical and physical possibility, we can introduce a third kind of possibility: metaphysical possibility. Using this concept, we can challenge premise 2 of the zombie argument: Just because something is *conceivable*, this doesn't mean it is *metaphysically possible*.

It seems conceivable, for example, that water could be something other than H2O because a statement like "water is H3O" is not obviously contradictory in the same way "a triangle has 4-sides" is contradictory. "Water is H2O" is not an analytic truth and so it *seems* that we can imagine water without imagining something with the chemical structure H2O. In contrast, we cannot imagine a triangle without imagining a 3-sided shape because "triangles have 3 sides" is an analytic truth. The apparent conceivability of "water is H3O" suggests such a thing is somehow possible.

However, we can reject the claim that "water is H3O" is possible by arguing that the chemical structure H2O is an essential property of what water is. Sure, you can imagine a possible world where the clear liquid in lakes and rivers is H3O, but then what you'd be imagining wouldn't be water! It would be something else.

Similarly, if qualia are essential properties of some physical things, then it's not metaphysically possible for the same physical thing to have different phenomenal properties. In other words, a physical duplicate without qualia (i.e. a philosophical zombie) is metaphysically impossible in the same way water without H2O is metaphysically impossible. So, philosophical zombies are metaphysically impossible.



*Possible **response** to the claim that zombies are metaphysically impossible:*

The property dualist can reject the analogy between water/H2O and brain/qualia. Yes, H2O is an essential property of water – you can't have water without H2O because that *just is* what water is. But it's not the same kind of relationship with phenomenal properties and physical properties:

- The physical structure (H2O) is *essential* to what water is, and so it is metaphysically impossible to have water without H2O or H2O without water
- But the physical structure (e.g. the brain and all related functional facts) *is not essential* to what qualia is, and so it is metaphysically possible to have qualia without the physical structure and the physical structure without the qualia (i.e. a philosophical zombie)

Not all properties are essential properties. For example, in our world, it is a property of water that it falls from the sky when it rains. But this property is not an *essential* property of water: if it never rained, then water would still be water.

The property dualist would argue that it's the same with qualia: the particular physical structure is not essential to what qualia is. The essential characteristic of *pain* qualia, say, is how it *feels* – not its physical structure. It is thus metaphysically possible to have pain without any particular physical structure and vice versa. So, philosophical zombies are metaphysically possible.

What is possible tells us nothing about the actual world

We can also respond to the zombies argument by rejecting premise 3: Even if zombies are possible, this still doesn't prove that qualia are *actually* non-physical. Jumping on to the moon from earth is logically and metaphysically possible, but this obviously doesn't prove people can do this in our *actual* world. So likewise, even if it's *possible* that qualia are non-physical, this doesn't show that qualia are *actually* non-physical. And so, even if zombies are possible, this doesn't necessarily show that property dualism is *true* – it just shows that property dualism is *possible*.

*Possible **response** to the claim that what is possible tells us nothing about the real world:*

The property dualist can argue that this response misunderstands property dualism and fails to defend physicalism. Physicalism claims that *everything* is either physical or supervenient upon the physical. Property dualism rejects this: it says there are at least some mental properties that are neither reducible to *nor supervenient upon* physical properties.

An analogy to illustrate what this means is as follows: Imagine God had just brought into existence the whole of the *physical* universe. Would God have to do more work to add in qualia to this purely physical universe?

- If yes, then qualia is non-physical and property dualism is true, because it means qualia depends (at least partly) on something over and above the physical facts
- But physicalism *has to say no:* the physical facts *alone* are all that's needed to fix the qualia, and so a zombie is *impossible*

If it's even *possible* there could be another world that is physically identical but with different qualia, then this shows that qualia does not supervene on the physical – because it shows that the physical facts do not necessarily "fix" the qualia. And so, if zombies are possible, then this *does* tell us something about our actual world: it tells us that qualia are neither reducible to nor supervenient on the physical. Thus, if zombies are even possible, property dualism is true.

Frank Jackson: The knowledge argument for property dualism

"Mary is confined to a black-and-white room, is educated through black-and-white books and lectures relayed on a black-and-white television. In this way she learns everything there is to know about the physical nature of the world. She knows all the physical facts about us and our environment, in a wide sense of 'physical' which includes everything in completed physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all this, including of course functional roles. If physicalism is true, she knows all there is to know."

In Frank Jackson's thought experiment here, Mary learns all the physical facts about our world from within a black and white room. She knows every possible physical fact – *including every physical fact about human experience of colour.*

Physicalism claims that everything is physical (or supervenes on the physical) and so there are only physical facts and knowledge of physical things. And so, if Mary knows every *physical* fact, physicalism must say she knows *every* fact.

But Jackson's thought experiment challenges this claim: Mary doesn't have *all* knowledge despite having *all physical knowledge* – because she doesn't know *what it is like* to see red, for example. When she leaves the black and white room and sees red for the first time, she will learn something new:

> *"It seems, however, that Mary does not know all there is to know. For when she is let out of the black-and-white room or given a color television, she will learn what it is like to see something red, say. […] Hence, physicalism is false."*

If Mary learns something new despite previously knowing all physical facts, then it follows that Mary's new knowledge is *non-physical*. So, knowledge of qualia is non-physical knowledge.

In short, Jackson's argument is this:

- Mary knows all the physical facts about colour
- But Mary does not know what it feels like to see colour
- Therefore, what it feels like to see colour is not a physical fact
- Physicalism says that all facts are physical facts
- Therefore, physicalism is false

Physicalism is taken to be false due to the supposed existence of non-physical facts. These non-physical facts are facts about non-physical properties, i.e. qualia. If such non-physical properties exist, then property dualism is true.

Responses to the knowledge argument

Mary would already know

One response to the knowledge argument is to reject the main premise of Jackson's argument: that Mary would learn something new.

We could argue that, if Mary really did know *all* the physical facts about what it's like to see red (as the thought experiment claims), then this would include knowledge of what it's like to see red. In other words, Mary would already know what it's like to see red *before* she left the black and white room.

This claim goes against an intuition most people have. We can't imagine how Mary could know *what it's like* to see red without actually having seen it herself. But how can we even imagine what it would be like to have *all physical knowledge* about something? Perhaps all physical knowledge just would include knowledge of *what it's like* to see red.

The intuition that Mary wouldn't know what it's like to see red is just that – an intuition. We can respond that this intuition is unreliable: All physical knowledge would also include knowledge of what it's like to see red.

*Possible **response** to the claim that Mary would already know what it's like to see red:*

Even though it's just an intuition that Mary would learn something new when she leaves the black and white room, the intuition is a powerful one. It's hard to make sense of how someone could know *what it is like* to see red purely from learning facts about human experience of red. It would be like describing colour to someone who has been blind from birth and expecting them to understand – it seems no amount of description could serve as a substitute for experience.

Jackson adds a further response to the claim that Mary would already know: Even if it were somehow possible that Mary could *imagine* seeing red from the physical facts alone, this still wouldn't mean Mary would *know* what it is like to see red. Being able to imagine red is not the same thing as knowing what it's like to see red – in fact, the reason Mary has to imagine what red is like is because she doesn't yet know what red is like.

The ability response

If you remember from epistemology, there are the three kinds of knowledge (page 26): ability, acquaintance, and propositional. We can use these distinctions to argue that Mary's new knowledge is physical knowledge – it's just not *propositional* physical knowledge.

We can argue that what Mary learns when she leaves the black and white room for the first time is *ability* knowledge. There is nothing spooky or non-physical about knowing how to ride a bicycle, for example – and yet knowing how to ride a bicycle is not the sort of knowledge you can learn from books or from within a black and white room.

To illustrate this, we can imagine a similar case to the Mary example but with knowing how to ride a bicycle:

- Mary learns all the physical facts about riding bicycles (and all the related causal and relational facts) from books and videos, etc. without ever actually touching a bicycle for herself
- When Mary is given a bicycle for the first time, she won't know how to ride it even though she knows all the *physical facts* about riding bicycles
- Therefore, knowing how to ride a bicycle is not a physical fact

The conclusion may be true – knowing how to ride a bike is not a physical *fact* – but it doesn't disprove physicalism. Knowing how to ride a bicycle is still a form of *physical* knowledge – it's just *ability* knowledge instead of propositional knowledge, and you can't learn ability knowledge from books.

Applied to the original Mary case, we can argue that when Mary sees red for the first time the story is similar: she gains new *abilities*. She gains the ability to imagine red, for example, and to distinguish red sensory experiences from green sensory experiences. But this new ability knowledge is just another form of *physical* knowledge.

*Possible **response** to the ability response: You can lack all abilities associated with seeing colour but still know what it is like to see colour*

However, we can imagine scenarios where a person lacks all *abilities* associated with seeing red, for example, and yet still knows what it is like to see red.

Using the examples above, we could argue that knowing what it is like to see red consists of having the *ability* to imagine what red is like and having the *ability* to distinguish red sensory experiences from green sensory experiences.

Imagine, however, a person who *lacks* these abilities for whatever reason – they have amnesia and immediately forget what red looks like, say. Now, imagine you place this person in front of a red wall. Surely, *in the moment* where this person is looking at the red wall, they *know* what it is like to see red – and yet they are unable to imagine seeing red, or any of the other abilities associated with seeing red.

If such a scenario is possible, then it shows that someone can know what it is like to see red without having any of the abilities associated with seeing red. This suggests that knowing what it is like to see red is not a form of ability knowledge.

The acquaintance response

Similar to the argument above, we can claim that although Mary learns something new, her new knowledge is still a kind of physical knowledge: knowledge by *acquaintance*.

To be *acquainted* with something – e.g. a person or a place – you need to have direct experience of that thing. For example, it wouldn't make sense to say you *know* Donald Trump – even if you know a load of *facts* about him – if you've never actually met him. Even if you learned every *physical* fact about Donald Trump, you still couldn't say you know him personally. But Donald Trump's friends and acquaintances *do* know him because they're acquainted with him. There's nothing spooky or non-physical about this knowledge, though.

You can make a (sort of) similar argument for Mary's experience of seeing the colour red: She can know all the physical facts about red – what it is, when people see it, how they

react to it, etc. – without being acquainted with redness itself. Mary is not acquainted with redness because her own brain has never had this property itself.

But when Mary sees red for the first time, the property occurs in her brain and she becomes *acquainted* with redness. Mary gains new knowledge from being acquainted with redness in this way. So, Mary does gain new knowledge – but it's just *physical* acquaintance knowledge.

*Possible **response** to the acquaintance response: Acquaintance knowledge involves propositional knowledge*

We can respond that acquaintance knowledge also involves propositional knowledge. For example, when Mary becomes acquainted with redness for the first time, she also learns propositional facts such as "seeing red is like *this*" (where "this" refers to the qualia). Whatever acquaintance knowledge Mary does or doesn't gain, it seems difficult to deny that Mary also gains propositional knowledge that "seeing red is like *this*". And this propositional knowledge is non-physical.

New knowledge, old fact

We can also respond to the knowledge argument by arguing that there is more than one way to know the same fact.

For example, "I know there is water in that glass" expresses knowledge of the same underlying physical fact that "I know there is H2O in that glass" does. But it's possible to know the former and not know the latter. Back before the chemical structure of water was discovered, for example, someone could know "there is water in the glass" but not know "there is H2O in the glass".

This example shows that the same fact can be understood via two different *concepts*. And we can argue that this is what's going on with Mary's knowledge of what it's like to see red: Before she left the black and white room, Mary only knew about redness in theoretical terms. But when she leaves and sees red she gains a new concept: the *phenomenal* concept. And it's impossible to know what it's like to see red without this concept.

We can argue that this phenomenal concept just provides a different way of understanding the same underlying fact. And so, Mary doesn't learn any new, non-physical fact. Instead, she just learns a *different way* of understanding *the same* fact. So, the knowledge argument does not prove the existence of non-physical facts or properties.

Responses to property dualism (in general)

Note: The following arguments mainly apply to epiphenomenalist dualism (rather than interactionist – see page 209) specifically, because they target the epiphenomenalist claim that mental properties (qualia) do not cause changes in physical substances.

Qualia causes introspective self-knowledge

(Epiphenomenal) property dualism avoids the conceptual and empirical interaction problems that interactionist dualism faces (page 206) because it doesn't make the claim that qualia can effect the physical world. Instead, (epiphenomenal) property dualism says that qualia have no causal powers.

However, the epiphenomenalist claim raises a different issue: It seems obvious that qualia *do* have causal powers. It seems obvious, for example, that unpleasant pain qualia is what *causes* me to pull my hand away from a hot stove.

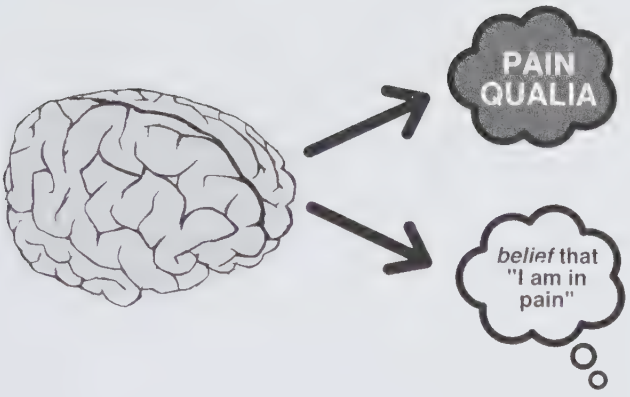
The epiphenomenalist response to this objection is that the (physical) brain causes *both* the qualia *and* the behaviour. So, in the hot stove example, the brain state causes *both* my pain qualia *and* me pulling my hand away. Here, the qualia is just an *effect* of the brain state – it doesn't actually do anything itself.

But we can push this objection further: Why do I *believe* I'm in pain, for example? The obvious answer is that I experience my own *mental* state – my own qualia – and this causes me to form the belief "I am in pain". It seems undeniable that our mental states are what cause our beliefs about our mental states. Again, though, this explanation isn't available to epiphenomenalism – because epiphenomenalism says mental states and qualia don't cause anything. The epiphenomenalist would have to say that the *brain* state is what causes my belief that I'm in pain.

But if the epiphenomenalist explanation of self-knowledge is correct, this raises a problem: My qualia becomes *irrelevant*. If my *brain* state is all that causes my beliefs about my *mental* state, then I would have the same beliefs *even if the mental state was completely different*. When I have the *brain* state of pain, for example, my pain qualia could be like my red qualia, or swap places with my pleasure qualia, or disappear entirely, and I would still form the same belief "I am in pain". So, if property dualism is true, it seems to imply that self-knowledge is impossible – but this seems wrong.

So, in short, the response is:

- If epiphenomenalism is true, qualia have no causal effects
- If qualia have no causal effects, then introspective self-knowledge is impossible
- But introspective self-knowledge is possible (e.g. I can know "I am in pain")
- So, epiphenomenalism must be false



If my brain state alone causes my belief that "I am in pain", then the qualia becomes irrelevant: you could swap pain qualia with anything and I would still form the same belief

Possible ***response*** *to the claim that qualia causes introspective self-knowledge:*

The reliabilist definition of knowledge (page 33) provides a potential way for epiphenomenalism to salvage introspective self-knowledge. If every time I have the *brain* state of pain it causes both pain qualia and a belief "I'm in pain", then that brain state counts as a reliable method for forming true beliefs about my qualia (because every time I have this belief about my qualia it turns out to be true, even though the belief isn't actually caused by the qualia itself). So, introspective self-knowledge might still be possible even if qualia have no causal powers.

Evolution

Evolution by natural selection seems to rule out epiphenomenal qualia – or, at the very least, would mean it is a massive coincidence it exists.

As briefly alluded to earlier in the context of teleological arguments (page 164), evolution says that genetic mutations occur randomly. Over millions of years, the environment selects for genes that give some benefit – either in terms of survival or reproduction. For example, having long neck genes enables a giraffe to reach food and survive. Or, to put it another way, having a long neck *causes* the giraffe not to die of starvation. The *causal* effects of long neck genes clearly explain why giraffes have long necks: they are beneficial for survival.

For every causally advantageous mutation, there are also many disadvantageous or *useless* mutations. Epiphenomenal qualia would be the latter: they would have no causal effects, and so there is no way they could confer any advantage.

It's obvious why having a *brain* state of pain, for example, is advantageous: it makes the animal get away from things that damage its body and might cause it to die prematurely. But there would be no evolutionary advantage of having the brain state of pain + *epiphenomenal pain qualia,* because the qualia doesn't do anything. The brain state *alone* would cause the animal to move away from whatever is damaging its body. Why, then, would nature select for qualia if it has no causal effects at all?

So, if minds are the product of evolution, it would suggest that epiphenomenalism is false: qualia does have some useful causal role, otherwise we wouldn't have evolved it. So, in short, the response is:

- Evolution means that only genetic traits that provide some advantage evolve
- If epiphenomenalism is true, qualia cannot provide any advantage because they have no causal powers
- So, if epiphenomenalism is true, qualia wouldn't have evolved
- But qualia have evolved
- So, epiphenomenalism must be false

*Possible **response** to the claim that evolution suggests qualia have causal effects:*

Epiphenomenalism can respond that this characterisation of evolution is too simple.

Firstly, evolution doesn't guarantee that only genetic traits that are advantageous survive. If a trait emerges that is neither advantageous or disadvantageous (as epiphenomenal qualia would be), there is no reason it would be selected against – it could just stay there silently doing nothing. So-called vestigial traits, for example, are traits that persist despite no longer having any evolutionary advantage (e.g. the human appendix).

Secondly, there are plenty of traits in nature that are simply effects of traits that *are* advantageous. Frank Jackson (the knowledge argument guy) gives one such example:

> *"Polar bears have particularly thick, warm coats. The Theory of Evolution explains this (we suppose) by pointing out that having a thick, warm coat is conducive to survival in the Arctic. But having a thick coat goes along with having a heavy coat, and having a heavy coat is not conducive to survival."*

The rules of the physical world are such that having a warm coat means also having a heavy coat. And the rules of the physical world may also be such that having sufficiently complex brain processes means also having qualia. In other words, qualia could be a by-product of brain states in the same way having a heavy coat is a by-product of having a warm coat.

The problem of other minds (again)

You can adapt the problem of other minds objection to substance dualism (page 205) to property dualism: Instead of asking how we can know other bodies have minds attached to them, we can ask how we can know other bodies have *qualia*. In other words, property dualism raises the question: How can you know you're not living in a world full of philosophical zombies (page 209)?

PHYSICALIST THEORIES

OVERVIEW

The technical definition of physicalism is the view that everything is physical or *supervenes* (see page 208) on the physical. In other words, physicalists say there is just one kind of thing – physical stuff – and that any properties or abstractions are inextricably linked with this physical stuff.

So, according to physicalism, two *physically* identical things must be *mentally* identical – it's *impossible* to have the same physical state but a different mental state.

Physicalist theories all agree that the mind is a physical thing (or supervenes on physical things). But they disagree about what kind of physical thing mental states are. The syllabus looks at four physicalist explanations of what mental states are:

1. **Behaviourism**
2. **Type identity theory**
3. **Functionalism**
4. **Eliminative materialism**

A summary of each theory and the key arguments is as follows:

	Behaviourism	**Type identity**	**Functionalism**	**Eliminativism**
Summary	Mental states are *behavioural dispositions*	Mental states are *brain states*	Mental states are *functions* within the cognitive system	Folk-psychology is a mistaken theory of the mind
Example	Pain is a disposition to say "ouch!" and rub the area that hurts	Pain is c-fibers firing in the brain	Pain is an unpleasant sensation that causes other mental states such as a desire to get away from the pain source	The concept of pain should be *eliminated* in favour of a neuroscientific alternative

	Behaviourism	Type identity	Functionalism	Eliminativism
Arguments for	Category mistake	Ockham's razor	Multiple realisability	Neuroscience is a more accurate theory
Arguments against	Asymmetry between self- and other- knowledge Super Spartans Circularity	Multiple realisability Zombies Divisibility argument Location problem	Inverted spectrum China brain Knowledge argument	Neuroscience is a less accurate theory Direct certainty of folk psychology Self-refuting

BEHAVIOURISM

Behaviourism says that mental states – such as pain, pleasure, sad, happy, etc. – are behavioural dispositions. Similarly, to have a mind is to have certain behavioural dispositions.

So, for a behaviourist, to be in pain is to wince, say "ouch!", try to get away from the source of the pain, and so on. What it means to have the mental state of pain is to display the behavioural dispositions associated with being in pain.

Behavioural dispositions

An obvious objection to behaviourism is that you can have a mental state but not display any behaviour. For example, you can be in pain but hide the behaviour – perhaps because you don't want to look like a wimp. Similarly, you can pretend to be in pain when you're not actually feeling anything, like when a player dives in football. This is why behaviourism analyses mental states in terms of behavioural *dispositions*, not just specific behaviours.

A disposition is how something will or is likely to behave in certain circumstances. For example, a wine glass has a disposition to break when dropped on a hard surface. The wine glass has this disposition even when it hasn't been dropped and is in perfect condition – hypothetically, *if you dropped the glass*, it would break.

Likewise, someone in the mental state of pain will have a disposition to say "ouch!" - even if they don't actually do so in every instance. Behavioural dispositions are thus hypothetical behaviours as well as actual behaviours. There is an infinite list of hypothetical actions, utterances and behaviours that make up being in pain. For example, if you were to ask someone who just stubbed their toe "did that hurt?", they would answer "yes". The person in pain has this disposition even if you never actually ask them the question.

Hard and soft behaviourism

The syllabus lists a distinction between two different types of behaviourism:

- Hard behaviourism: All propositions about mental states can be reduced *without loss of meaning* to propositions about behaviours
- Soft behaviourism: Propositions about mental states are propositions about behavioural dispositions

In other words, hard behaviourism says you can give a *complete* account of the mind in terms of behaviours: the mind is *nothing more* than behavioural dispositions. If you describe all the behaviours and behavioural dispositions, you have described the mind – there's nothing left over. In the language of the syllabus, mental states *reduce* to behaviours and behavioural dispositions.

In contrast, soft behaviourism allows for the possibility of a mind (and stuff like qualia). Soft behaviourism simply says that *propositions* about mental states, such as "I am in pain", should be understood as propositions about behavioural dispositions – but it doesn't go as far as to say that "pain" is nothing more than behavioural dispositions. In other words, a complete description of the behaviours and behavioural dispositions of "pain" would not quite be a complete description of what "pain" is.

Gilbert Ryle: *The Concept of Mind*

Note: Ryle is arguing for soft behaviourism.

General arguments against dualism

Gilbert Ryle's (1900-1976) arguments for behaviourism start with some of the criticisms of dualism we looked at previously – such as the problem of other minds (page 205) and the problem of causal interaction (page 206). He then gives a new argument against dualism: that if dualism were true, mental concepts would be impossible to use.

Consider this: If dualism were true and mental states such as pain referred only to some non-physical and completely *private* mental state, how could we ever talk about them? You can't literally show me what is going on in your mind, so how could I know what you mean when you say you are in pain? If dualism were true, it would be impossible for anyone to know anything about others' mental states because there would be no way of seeing or otherwise experiencing their minds and mental states.

Category mistake

Ryle argues that to think mental states are distinct from their associated behaviours is to make a *category mistake*. He gives the following example to illustrate what he means by this:

- Someone wants to know what Oxford University is
- So you show them the library, the lecture theatres, the teachers, the students, and so on
- After the tour is complete, the person says: "OK, you've shown me the lecture theatres and the teachers and so on, *but what is Oxford University?*"

To ask such a question is to make a category mistake: Oxford University isn't anything over and above the buildings and people etc. that make it up. There's no one single thing you can point to and go: "*that* is Oxford University".

Ryle argues that dualists make the same sort of category mistake when talking about mental states – they are nothing over and above the behaviours and behavioural dispositions. To illustrate this, we can imagine a similar scenario to the one above, but with the mental state of pain:

- An alien comes down and wants to know what the mental state of pain is
- So you show the alien people stubbing their toes, being tortured, wincing, saying "ouch!", and so on
- After showing the alien these examples of pain, it asks: "OK, you've shown me all these people being tortured and so on, *but what is pain?*"

In just the same way that showing someone the buildings, teachers, and so on is to show them what *Oxford University* is, Ryle is arguing that showing someone the various behavioural dispositions associated with pain is to show them what the mental state of *pain* is. There is nothing you can show someone over and above these examples of behaviours – no single thing you can point to and say "*that* is pain". It is a category mistake to assume otherwise.

Responses to behaviourism

Asymmetry between self- and other- knowledge

Behaviourism's analysis of mental states in terms of behaviours may work well for understanding *other* people's mental states, but it doesn't appear to work in the case of *self*-knowledge. The following are two key differences between self-knowledge and knowledge of other people's mental states:

- Knowledge of my own mental states is *direct*: When I stub my toe, I have direct access to the feeling of pain it produces. But if I see someone else do the same thing, I have to *infer* the fact that they're in pain from their behaviour – so my knowledge of their mental state is *indirect*.
- I cannot be mistaken about my own mental states: When I feel pain, there's no way I could be mistaken as to what I'm feeling – I can't think I'm in pain and then realise I'm not. However, if I see someone else scream "ouch!", I might mistakenly form the belief that they're in pain only to find out they were acting and not in pain at all. So, I can be mistaken about other people's mental states but not my own.

These examples suggest there's a big difference between how you experience your own mental states and other people's. But if behaviourism were true, this shouldn't be the case.

Remember, according to behaviourism, mental states are behavioural dispositions. So, to be in pain is to have certain behavioural dispositions – but it doesn't feel like that when it's happening to you! If someone were to ask how I know I'm in pain, it wouldn't make sense to answer "because I winced" or "because I said "ouch!" or give some other behavioural explanation. I just *feel* the pain directly and know I'm in pain from the unpleasant feeling. But there's no room for this "unpleasant feeling" with behaviourism because behaviourism analyses mental states in terms of behaviour.

So, in short, the argument looks like this:

- Behaviourism seems to rule out any asymmetry between self knowledge and knowledge of other people's mental states
- But there clearly is an asymmetry between self knowledge and knowledge of other people's mental states
- So, behaviourism must be false

*Possible **response** to the asymmetry between self- and other- knowledge:*

Ryle's reply to this problem is to reject that self knowledge and knowledge of other people's mental states are a different *kind* of thing. According to Ryle, we gain self-

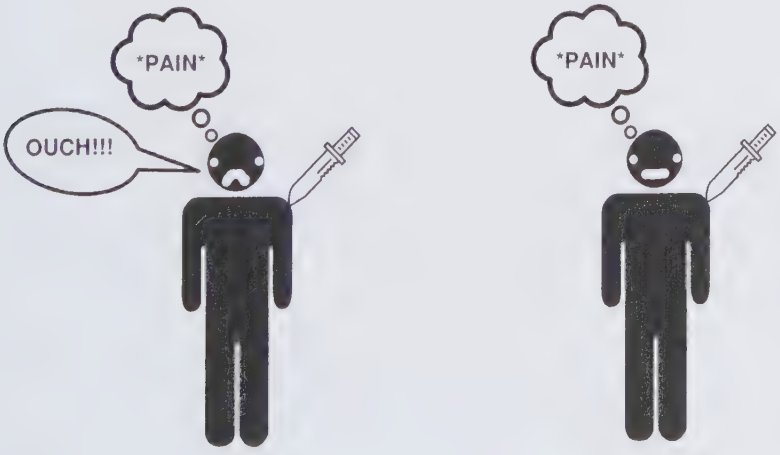
knowledge and knowledge of others in the same way: by paying attention. The apparent asymmetry is just an illusion that comes from having far more information in the case of self-knowledge vs. knowledge of others.

Ryle likens *thinking*, for example, to a kind of *internal* speech – and speech is an *external* behaviour. The fact that thinking is silent and private whereas speech is external and public is not a particularly important distinction: they're both a kind of verbal practice/*behaviour.*

Returning to the asymmetry objection, Ryle could say that there is no difference *in kind* between knowledge of your own thoughts vs. knowing someone else's: You can know what someone is thinking by paying attention to what they say, just as you can know what you're thinking by paying attention to your thoughts. Yes, you have a lot more information available in your own case – and it may even be more direct - but the mechanism is ultimately the same: paying attention.

Super Spartans

Philosopher Hilary Putnam develops the asymmetry argument further with his example of "super Spartans". Super Spartans are an imagined community of people who completely suppress any outward demonstration of pain. They don't wince, flinch, say "ouch!", or anything like that – under *any* circumstances. Super Spartans even control their heart rate and any other external indications of pain. In other words, super Spartans never ever have any dispositions toward pain behaviour whatsoever.



Both the normal human (left) and the super Spartan (right) feel pain internally, but the super Spartan shows no outward manifestations whatsoever

Nevertheless, we can imagine the super Spartans do feel pain internally. They might never show it externally, but they would still experience a subjective experience of pain if they were tortured, say.

Remember, behaviourism says pain is a disposition to behave a certain way. But here we have an example of the mental state of pain without the associated behavioural dispositions. Therefore the two things – pain and pain behaviour/dispositions – are two separate things. So, if super Spartans are possible, then behaviourism is false.

So, in short, the argument looks like this:

- Behaviourism says to be in pain is to have a disposition to behave in certain ways
- Super Spartans can be in pain but not have a disposition to behave in those ways
- So, pain is not the same thing as behavioural dispositions
- So, behaviourism is false

*Possible **response** to super Spartans:*

Behaviourists could reply that super Spartans are impossible: Without any sort of outward display of pain, it would be impossible to form the concept of pain. In other words, the concept of "pain" would be completely meaningless in a society where there was no evidence of its existence. If there was no external evidence of pain – no pain behaviour, ever – then how would Super Spartans even be able to work out what behaviours they were supposed to repress in the first place? So, Putnam's super Spartans are an incoherent concept.

Zombies (again)

The zombie argument for property dualism (page 209) can also be used to argue against behaviourism.

A zombie is basically the exact opposite of a super Spartan: where the Spartan has qualia but not behaviour, the zombie has behaviour but no qualia. The zombie might say "ouch!" when it gets stabbed, but it doesn't feel any pain internally. In other words, the zombie has all the behavioural dispositions associated with pain, but it doesn't have the mental state of pain.

If zombies are possible, then the behavioural disposition of pain is separate from the mental state of pain. And if the behavioural disposition can be separate from the mental state then behaviourism is false.

*Possible **response** to zombies:*

Behaviourism could argue that zombies are inconceivable (page 210): Mental states *are* behavioural dispositions, and so mental states without behavioural dispositions are inconceivable.

The conceivability argument (again)

We can also adapt Descartes' conceivability argument for substance dualism (page 199) to argue against behaviourism, for example:

- Mental states without behavioural dispositions are conceivable (e.g. super Spartans)
- So, mental states without behavioural dispositions are possible
- So, mental states are not behavioural dispositions
- So, behaviourism is false

*Possible **response** to the conceivability argument:*

Behaviourism would reject the conceivability of mental states without behavioural dispositions. According to behaviourism, mental states *are* behavioural dispositions, and so mental states without behavioural dispositions (or vice versa) are inconceivable. See page 200 for more.

Multiple realisability

Behaviourism identifies mental states with behavioural dispositions. But the same mental state can be realised through multiple (and even completely contradictory) behaviours depending on other mental states, *which also need to be identified with behavioural dispositions*. This process leads to an endless cycle such that behaviourism can never give a proper account of what mental states are.

Let's take the mental state of being thirsty as an example: Behaviourism would say this mental state is the same as having a disposition to drink a drink if it was in front of you. But this isn't always the case: The person might be thirsty and yet not drink the drink – *if they also have the mental state of believing that the drink is poisoned*.

So, in order to explain why the mental state of me being thirsty leads to this particular behavioural disposition (i.e. not drinking), we need to appeal to *another* mental state (my belief that the drink is poisoned). But this additional mental state (believing the drink is poisoned) *also* needs to be analysed in terms of behaviour.

The mental state of believing that a drink is poisoned would probably cause you to behave by avoiding the drink or pouring it away. But someone else might drink the drink – despite also believing that the drink is poisoned – if they also have the mental state of being suicidal. And the mental state of being suicidal can also be realised in multiple different behaviours depending on a person's other mental states, and so on and so on.



The mental state of being thirsty can be realised by multiple behaviours depending on other mental states. But those other mental states can also be realised by multiple different behaviours depending on other mental states, and so on forever.

So, in short, the objection is:

- Behaviourism says mental states should be understood as behavioural dispositions
- So, if behaviourism is correct, different mental states are different behavioural dispositions
- But each mental state can be realised by *any* behaviour (depending on other mental states)
- And those other mental states can also be realised by *any* behaviour
- This makes it impossible to understand mental states as behavioural dispositions because any mental state could be realised by any behaviour and vice versa

Circularity

We can further press the multiple realisability objection above to argue that this process of analysing mental states in terms of behaviour becomes *circular*.

As shown above, behaviourism must assume other mental states in order to give an analysis of mental states in terms of behavioural dispositions. But, if you keep going with this analysis, you will eventually end up defining those other mental states in terms of the *original* mental state you set out to define!

Overly simplified example:

- To be in *pain* is to be disposed to say "ouch"
- But you won't say "ouch" if you have a *desire* to not look like a wimp
- To have a *desire* to not look like a wimp is to be disposed to hide when you are in *pain*

TYPE IDENTITY THEORY

Type identity theory is perhaps the most obvious physicalist theory of mind. It says that mental states reduce to (physical) brain states. In other words, type identity theory says mental states are the same things as brain states.

An example often used in philosophy of mind to illustrate brain states is *c-fibers*. To say someone's c-fibers are firing is just technical shorthand for the *brain* state associated with pain. And so, a type identity theorist would say that pain is the same thing as c-fibers firing.

It may not *seem* like the mental state of pain is the same thing as c-fibers firing – one is a subjective private feeling (pain qualia) and the other is an objective physical thing (brain state) – but, according to type identity theory, the two things are one and the same. The only reason they seem different is because of how we know about them: we experience pain directly, whereas we know about the brain through observation and neuroscience.

So, type identity theory would say all of the following:

- The mind is the brain
- Mental states are brain states
- Pain is c-fibers firing

J.J.C. Smart: *Sensations and Brain Processes*

Ontological reduction / contingent identity

J.J.C. Smart (1920-2012) claims that mental states and brain states are *contingently* identical. Another way of saying this is that mental states *ontologically* reduce to brain states (which is how the A-level philosophy syllabus defines type identity theory).

Other examples of ontological reductions/contingently identical things are:

- Lightning is electrical discharge
- Water is H2O
- This table is an old suitcase

These relationships are not merely *correlations*. It's not like electrical discharge happens at the same time as lightning, or that the electrical discharge *causes* the lightning – the lightning and the electrical discharge *are the same thing*. According to type identity theory, it's the same story with pain and c-fibers – they just are the same thing.

However, ontological reductions like this are not *analytic* reductions, such as "a bachelor is an unmarried man". "A bachelor is an unmarried man" is an analytic reduction because the opposite idea (i.e. a married bachelor) is a contradiction. In contrast, even though

"lightning is electrical discharge" is true, there is no *logical* contradiction in saying "lightning is *not* electrical discharge".

So, according to type identity theory, pain *ontologically* reduces to c-fibers firing, but it does not *analytically* reduce to it. It is an empirical/a posteriori discovery that pain and c-fibers firing are the same thing.

Ockham's razor (against dualism)

Note: Ockham's Razor can be used as an argument against dualism and in favour of physicalism more generally. It's not just an argument for type identity theory.

Smart's main argument for type identity theory is Ockham's razor. Ockham's razor is a scientific/philosophical principle which says something like:

> *"Do not multiply entities beyond necessity"*

A more colloquial formulation of Ockham's razor would be something like "the simplest explanation is the best". In practice this means that if two theories make the same accurate predictions, the theory that does so using the fewest number of entities is likely to be the better and more accurate theory.

An example of Ockham's razor in action would be geocentric vs. heliocentric models of planetary orbits. Back before it was widely accepted that the earth revolves around the sun, astronomers would posit all these crazy forces and weird orbit patterns to explain why planets and stars sometimes appear to slow down, stop, and move backwards. Such geocentric theories were often complex, but they were able to make pretty accurate predictions about the movement of planets.



Geocentric model (left) and heliocentric model (riaht)

However, Copernicus' heliocentric model made the same accurate predictions as these geocentric models – *but without positing all these crazy forces and weird orbits*. So, Ockham's razor (correctly) suggests that the Earth revolves around the sun and not the other way round.

Smart has a similar idea in mind when arguing against dualism: He argues that type identity theory can predict everything that dualism can, but type identity theory does so with *one* entity (the brain) rather than *two* (mind and brain).

If there are no overwhelming arguments or proof of dualism, Smart argues, we shouldn't posit extra entities to explain the mind. We can explain just as much about mental states by referring to the brain as we can by referring to a non-physical mind. For example, when I feel pain, brain scans show that my c-fibers get activated. And when my c-fibers get activated, I feel pain. This suggests they are the same thing. We don't need to posit an additional substance here.

So, where a dualist would say the c-fibers and the pain are two separate substances, type identity theory says they are the same physical thing (but different concepts).

Responses to type identity theory

Location Problem

If my c-fibers are firing, it's presumably pretty easy to locate where this is happening. You could put me in an MRI scanner, for example, and find out the exact location of the c-fibers firing.

But my pain doesn't seem to have the same physical location – it seems like it's somewhere else. If you locate my c-fibers it doesn't seem like you've located my subjective mental sensation of pain. You can point to my c-fibers, say, or even touch them, but you can't point to or touch my *mental state* of pain.

We touched on Leibniz's law of indiscernibility of identicals in Descartes' divisibility argument earlier (page 203) and we can use it to make an argument here: My brain state of c-fibers firing has the property of *having a physical location*, but my mental state of pain *does not have the property of having a physical location*. And so, assuming Leibniz's law, my mental state of pain cannot be the same thing as my brain state of c-fibers firing, because the two things have different properties.

So, in short, the argument is something like this:

1. If pain and c-fibers firing are identical then they must share all the same properties
2. C-fibers have a precise physical location

3. Pain does not have a precise physical location
4. Therefore, pain and c-fibers firing are not identical

*Possible **response** to the location problem:*

Smart responds that the location problem is only a linguistic one: It just *sounds* weird to attribute a physical location to a mental state, because we don't do this in our ordinary language. But just because it sounds weird to talk about mental states having a physical location, that doesn't show they don't, in fact, have a physical location in space.

Further, Smart argues, this response begs the question: it assumes the truth of its conclusion in its premises. To assume the premise that mental states *do not* have a physical location is to assume they are not brain states (which do have a physical location). A proper understanding of the world would reveal premise 3 to be false: Pain *does* have a precise physical location.

Zombies (again)

The zombie argument for property dualism (see page 209) can also be used against type identity theory.

Type identity theory says pain is identical to c-fibers firing, but we can imagine a zombie with the brain state (c-fibers firing) but not the mental state (pain). So, if zombies are possible, pain and other mental states are not the same things as brain states.

*Possible **response** to zombies:*

Type identity theorists would likely argue that zombies are metaphysically impossible (page 211): Mental states *just are* brain states in the same way water *just is* H2O, and so it is metaphysically impossible to have the brain state without the mental state in the same way it is metaphysically impossible to have water without H2O.

The conceivability argument (again)

We can also adapt Descartes' conceivability argument for substance dualism (page 199) to argue against type identity theory, for example:

- Mental states without brain states are conceivable
- So, mental states without brain states is possible
- So, mental states are not brain states
- So, type identity theory is false

*Possible **response** to the conceivability argument:*

See responses to the conceivability argument starting from page 200.

The divisibility argument (again)

Descartes' conceivability argument for substance dualism (page 199) is basically an argument against type identity theory (just substitute "body" for "brain"). For example:

1. My brain is divisible
2. My mind is indivisible
3. Therefore, my mind and body are separate substances

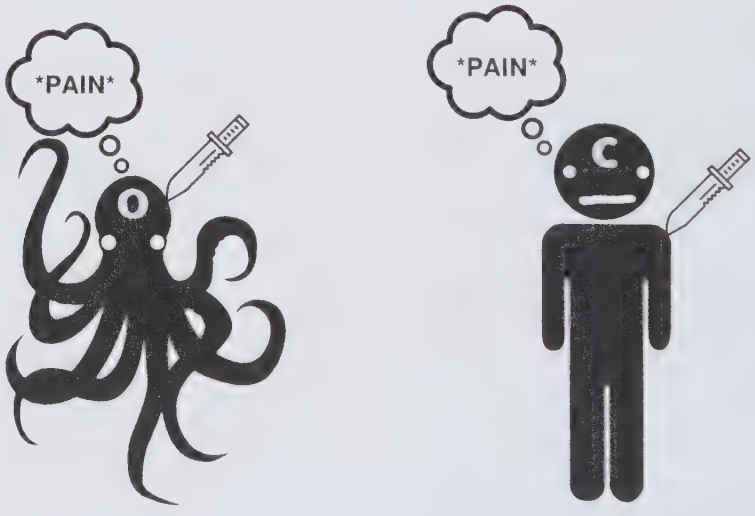
*Possible **response** to the conceivability argument:*

See responses to the divisibility argument starting from page 203.

Multiple realisability

Hilary Putnam argues that mental states cannot be reduced to brain states because mental states are *multiply realisable*. What this means is that the same mental state can come from many different brain states.

For example, an octopus has a very different brain setup to a human. Let's say, for example, that octopuses have *o-fibers* instead of *c-fibers*. However, if you stabbed an octopus and it writhed about, surely you would say that's because it's in *pain*. But if type identity theory is true, this isn't possible: an octopus *can't feel pain*, because pain is c-fibers firing, and an octopus doesn't have c-fibers.



Both the octopus and the human have the mental state of pain, but the human's brain has c-fibers, whereas the octopus' brain has o-fibers

Another example that Putnam uses is silicon-based aliens. If type identity is true, then an alien and a human can't both have the same belief that "grass is green" because my brain is made from carbon and the alien's brain is made from silicon. But this just seems wrong: we both share the same *mental* state (a *belief* that "grass is green") despite our differing physical states.

So, in short, the argument is something like this:

- If type identity theory is true, you cannot have the same mental state without having the same brain state
- An octopus and a human can have the same mental state (e.g. they can both be in pain)
- But an octopus and a human do not have the same brains or brain states
- Therefore, type identity theory is false

FUNCTIONALISM

Functionalism defines mental states as *functional* states within an organism.

For example, we might say the functional role of pain is an unpleasant sensation that creates a desire for the pain to stop and causes the organism to get away from the thing that's causing it harm. That function is what mental states, such as pain, are.

Functions should be understood within the context of the entire mind. So, the function of pain, for example, isn't simply to cause behavioural dispositions (as behaviourism claims). Part of the function of pain is to cause *other* mental states – such as a belief that you are in pain, or a desire for the pain to stop. So, where behaviourism talks only of inputs and outputs, functionalism talks of inputs, mental states, and outputs.

Hilary Putnam: *The Nature of Mental States*

Similar to the description above, Hilary Putnam (1926-2016) argues that "being in pain is a functional state of the organism." Putnam then argues that this analysis of mental states is superior to both type identity theory and behaviourism.

Arguments against type identity theory

Putnam's argument against type identity theory is the multiple realisability of mental states. Pain, for example, can be experienced by a human, an octopus or an alien – but these things all have different physical structures, which suggests that pain is not any single physical property (e.g. c-fibers firing).

We already looked at multiple realisability as an argument against type identity theory above (page 236), but it can also be used as an argument for functionalism:

Putnam doesn't use this example, but functionalists would generally agree that mental states are analogous to a knife: A knife can be made from metal, or plastic, or wood – as long as it performs its *function* (to cut things). Functionalism makes a similar claim about mental states: Rather than analysing pain in terms of its specific physical characteristics – e.g. c-fibers firing – functionalism says that mental states such as pain are *functional states* within the cognitive system. This function is not specific to any physical instantiation, and so functionalism avoids the multiple realisability objection to type identity theory and can allow that humans, octopuses, and aliens can all feel pain.

Note: In theory, even some non-physical thing could serve the function of pain, and so functionalism is compatible with substance dualism (page 199). However, most functionalists are physicalists.

Arguments against behaviourism

Putnam argues that behaviourism fails as a theory of mind because it is circular. We already looked at this multiple realisability/circularity problem as an argument against behaviourism (page 231), but we can also use it as an argument for functionalism:

Behaviourism is restricted to identifying mental states *solely* in terms of behavioural dispositions, but this causes problems when we realise that the same mental state can be realised by pretty much any behaviour (given other mental states). For example, being in pain may be analysed as a disposition to say "ouch!" - but this analysis doesn't work when you also have the mental state of not wanting to look like a wimp.

Functionalism avoids this problem because the function of mental states is more than simply to cause behaviours – *mental states can also cause other mental states*. My mental state of pain, say, could serve the function of causing other mental states such as a *belief* that I am in pain and a *desire* for the pain to stop. My mental state of pain serves this function *even if that mental state ends up having no effect on my behaviour*. Further, the behaviour my mental state causes depends on my other mental states.

Functionalism thus avoids tying mental states to *specific* behaviours, and thus doesn't end up having to give circular explanations of how a person can be in pain when their behaviour says otherwise, as behaviourism does:

> *"Since the behavior of the Machine (in this case, an organism) will depend not merely on the sensory inputs,* ***but also on the Total State (i.e., on other values, beliefs, etc.),*** *it seems hopeless to make any general statement about how an organism in such a condition must behave [which is what behaviourism does]"*

In short, behaviourism takes too narrow a view in associating specific mental states with specific behaviours, whereas functionalism acknowledges that mental states can be realised by multiple (often contradictory) behaviours. The functional role of mental states within an organism is more than simply to cause behaviours – a mental state's function is also determined by its effects on other mental states.

Responses to functionalism

Inverted spectrum

What if my experience of green was like your experience of red and vice versa?

For example, what if my qualia when I look at grass is more like your qualia when you look at ripe tomatoes. When we both look at the grass, our mental states would be *functionally* identical. They would both, for example, cause us to believe "grass is green" and cause us to say "green" when asked "what colour is the grass?"



These two mental states are functionally the same, but they are clearly different mental states

And since our mental states are functionally identical, functionalism must say they are the *same* mental state. But they're clearly not the same: my qualia is completely different from yours.

So the argument is something like:

- Functionalism says mental states are functional states
- If functionalism is true, then two functionally identical mental states are the same mental state
- My mental state when I look at green grass is functionally identical to yours but phenomenally different
- Therefore, our two mental states are not the same mental state despite being functionally identical
- Therefore, functionalism is false

Zombies (again)

The zombies argument for property dualism (page 209) can be used to make a very similar point to the inverted spectrum argument above. However, in this example, instead of two people having *different* qualia, one person has normal qualia and the zombie has *no* qualia.

The zombie's mental state would again be *functionally* identical to a normal human's mental state. For example, looking at the sea would cause a state that is functionally identical to a belief that "the sea is blue". The zombie would also say "blue" when asked "what colour is the sea?"

However, despite being functionally identical, the zombie's mental state is clearly not the *same* as a normal human's mental state: it lacks the qualia. Thus, functionalism must be false because you can have two different mental states that are nevertheless functionally identical.

The knowledge argument (again)

We can adapt the knowledge argument for property dualism (page 213) to criticise functionalism.

In the original argument, it is argued that all the physical facts would not be enough for Mary to know what it's like to see red if she'd never seen it for herself. To adapt this argument for functionalism, we can specify that Mary knows all the physical *and functional* facts about seeing colour. However, this would seemingly still not be enough for Mary to know what it's like to see red. If true, the mental state of seeing red would be something more than just a functional state. So, functionalism is false.

The China brain

Ned Block's China Brain thought experiment describes a setup that is functionally identical to a mind but is clearly not the same thing.

These are the key points:

- Imagine we have a complete functional description of human mental states
- A human body is hooked up to the entire population of China
- Every person in China is linked to other people (neurons) via two-way radios
- They communicate according to the rules set out in the complete functional description of human mental states described earlier
- Some of these people (neurons) are linked to the outputs of the body
- Imagine the Chinese population recreated the functions of the neurons

So, the same input leads to exactly the same output as a normal human, and everything in between is functionally identical.

Basically, the scenario above is designed to replicate a human brain. The population of China is roughly equal to the number of neurons in a brain and the two-way radios replicate the firings of the neurons. According to functionalism, the China brain would actually be in pain, say, given the appropriate inputs (like being stabbed). But this is obviously false.

Just because the state of the China brain is *functionally* identical to human pain, it clearly isn't *actually* in pain. So functionalism must be false: there's clearly more to mental states than their functional role.

ELIMINATIVE MATERIALISM

Eliminative materialism (also called *eliminativism*) is the view that our common sense understanding of mental states is radically mistaken: Some or even all mental states (as we typically understand them) *don't exist*.

All the other theories we've looked at (including dualism) assume mental states such as pain *exist* - they just disagree about *type* of thing they are. Eliminative materialism, however, rejects this assumption and argues that terms like "belief" and "pain" don't correspond to anything specific. These terms might be a useful and practical way of talking about mental states in everyday life, but when we look closer at what they

actually *are*, they don't reduce to anything in particular. So, whereas other theories have tried to *reduce* mental states to something else – their function, or physical instantiation, or whatever – eliminative materialism says we should *eliminate* them.

Eliminativists think that a proper analysis of mental states will look more like neuroscience, with specific descriptions of the mechanics of the brain. Although this may sound similar to type identity theory, the difference is this: Type identity theory says the concept of pain *reduces* to c-fibers firing (i.e. pain and c-fibers are the same thing), whereas eliminativism says the concept of pain doesn't reduce to anything and so should be *eliminated* (i.e. pain doesn't exist, but c-fibers do exist).

Paul Churchland: *Eliminative Materialism and the Propositional Attitudes*

Folk psychology

Paul Churchland argues against what he calls *folk psychology*. Folk psychology is the everyday psychological concepts and explanations of behaviour we use.

Examples of folk psychology include:

- He ran away because he was *scared*
- She got a drink because she was *thirsty*
- They ran out of the building because there was a *panic*
- He studied for the test because he *wanted* to get a good grade
- She rubbed her ankle because it *hurt*

Folk psychology is such an integral part of our language that it might be hard to even see what it is. It's basically just the common sense way we talk about mental states.

The theories we've looked at so far seek to analyse folk psychology terms in various ways. For example, type identity theory reduces the folk psychology concept of pain to c-fibers, functionalism reduces it to its functional role, and behaviourism reduces it to a set of behavioural dispositions. All these theories implicitly agree that pain is a thing that exists and that can be reduced to something else – but they just disagree about *what* it reduces to.

Eliminative materialism rejects this view, however. Eliminativists argue that folk psychology concepts do not reduce to anything and so should instead be *eliminated*.

Folk psychology as a scientific theory

Churchland argues that folk psychology is a *scientific* theory like any other, and so we should treat it as such. If a better theory comes along, we should abandon the folk psychology theory in favour of the better theory.

Scientific theories have laws and rules, and these laws can be used to make predictions. For example, the laws of gravity can predict where Mars will be in the sky on 17th August 2127. A scientific theory that makes accurate predictions is a better (more representative of reality) theory than one that makes inaccurate predictions.

Folk psychology has its own – loosely defined – set of laws and rules. And these folk psychology laws have reasonable predictive power. For example, you might predict "if she gets *sad* she will cry" or "if he has a *belief* that eating animals is wrong, he won't order the chicken". However, the predictions of the folk psychology laws are not 100% accurate. Sometimes a person will surprise you by behaving in a way you don't expect. For example, our folk psychology framework may *predict* that someone will be sad and cry in response to a sad movie, but then they don't. So, in terms of predictive power, folk psychology is not a perfect scientific theory.

Further, there are many mental phenomena that the laws of folk psychology can't really account for such as mental illness, sleep, and learning. Folk psychology can't really explain why people get depressed, for example, or how babies learn to speak. This counts as further evidence against the folk psychology scientific theory because it shows folk psychology is incomplete as well as inaccurate.

Another point, Churchland argues, is that good scientific theories *advance* over time. For example, Aristotle's physics was replaced by Newton's, and then Newton's physics was replaced by Einstein's. Each new scientific theory is more accurate and makes better predictions than the last. Folk psychology, though, hasn't advanced in this way – it's still the same theory as it was thousands of years ago. The *words* we use to describe folk psychology might have changed, but the general theory is the same: the Romans would have said folk psychology things like "he laughs when he's nervous" – just in their own language.

Finally, Churchland argues, the *intentionality* of folk psychology doesn't fit well with pretty much all other areas of science. Intentionality is the "aboutness" of mental states. For example, folk psychology talks about having a thought *about* something (e.g. "I am thinking *about* a frog") or a belief *about* something (e.g. "he has a belief *about* the existence of God").

But, assuming physicalism is true and mental states are physical things, it's not clear how a physical thing could be *about* anything in the way folk psychology assumes. A table or a chair doesn't have intentionality in this way – it isn't *about* anything. Ordinary objects don't have intentionality. In fact, nothing else in science has intentionality – we don't say gravity or temperature are *about* anything, for example. So, either mental states are this weird outlier to the rest of science, or folk psychology is making a mistake in its analysis.

So, in short, Churchland's arguments against the folk psychology scientific theory are:

- Folk psychology does not always make accurate predictions
- Folk psychology cannot explain aspects of mental life, such as mental illness and learning
- Folk psychology hasn't changed or improved in thousands of years
- The intentionality of folk psychology is inconsistent with other (very accurate) scientific theories

Folk psychology vs. neuroscience

Given the problems with folk psychology described above, Churchland argues we should look to *eliminate* folk psychology and replace it with a more accurate scientific theory – in particular, the theories of neuroscience.

The main argument in favour of neuroscience is the obvious connections between the mind and the brain. As mentioned in the Ockham's razor argument for type identity (page 233), a brain scan can reveal that the brain lights up and activates in particular ways whenever a person feels certain mental states. This suggests mind and brain are linked.

But where eliminative materialism differs from type identity is on the issue of *reduction*. Type identity theory says folk psychology reduces to particular brain states. In other words, type identity theory says that for every folk psychology concept (e.g. pain, belief, happy, sad, think) there is a corresponding brain state that is *identical* to that folk psychology concept. Eliminativism, however, does not expect such a neat and tidy reduction. Yes, there may be general correlations between our folk psychology concepts and certain types of physical brain states – but these correlations aren't perfect. Close analysis of our folk psychology concepts will reveal they don't reduce to anything in particular and so these folk psychology concepts are not an accurate way to think about the reality of the mind.

It's important to point out that Churchland isn't saying we should eliminate folk psychology from *everyday* language in favour of neuroscience – that wouldn't be practical. An analogy for this can be found in engineering: Engineers often use Newton's equations to make their calculations because doing so is quicker than using Einstein's equations. Even though Einstein's equations are technically more accurate, they would take forever to calculate and the difference in outcome is so small that it doesn't matter. Similarly, folk psychology is useful in everyday contexts: it's much quicker than doing millions of brain scans and neuroscientific analysis, and it's fairly accurate in its predictions. However, if we truly want to describe reality (e.g. in science or philosophy of mind) we should eliminate folk psychology in favour of the correct and more accurate neuroscientific explanations.

Responses to eliminative materialism

Folk psychology has good predictive power

We saw above that Churchland criticises folk psychology for its explanatory and predictive failures. But we can respond that folk psychology *does* explain and make fairly accurate predictions about how people behave, such as the following:

- When he feels *nervous* he talks really fast
- If she has a *belief* that eating animals is wrong, she won't order the chicken
- He shouted and stomped about because he was *angry*
- If she wins the lottery she will be *happy* and jump about cheering
- When he is in *pain* he swears loudly

In contrast, neuroscience is pretty bad at predicting behaviour – at least at present. The brain is a highly complex structure and this makes it incredibly difficult for neuroscience to model and predict even the simple behaviours in the bullet points above. It's doubtful whether a team of the best neuroscientists in the world, using the most advanced equipment available today, could more accurately predict a typical human's behaviour than folk psychology can.

*Possible **response** to the claim that folk psychology has good predictive power:*

Eliminativists can respond that even if folk psychology is more accurate than neuroscience at present, a future neuroscience will be superior to folk psychology.

Compared to theories elsewhere in science, folk psychology is far weaker at making accurate predictions and explaining phenomena. The gaps in folk psychology, such as its imperfect predictions and inability to explain mental illness, show that folk psychology is a flawed theory of the mind that is a far less accurate model of the world than most theories in, for example, physics and chemistry. Whereas folk psychology often gets predictions of behaviour wrong, a physics theory like general relativity will *always* make accurate predictions about e.g. where a planet will be in the sky. Eliminativists would argue that the correct theory of mind must be equally accurate and that the only way to achieve such accuracy will be through neuroscience. A perfected neuroscience would be able to predict everything folk psychology can – and more – but with greater accuracy.

The direct certainty of folk psychology

We can argue that eliminative materialism goes against some of the strongest and most certain intuitions we have.

For example, we saw in Descartes' cogito argument (page 65) how even from within a position of extreme scepticism, we can still be certain of the proposition "I think". Surely,

it doesn't make sense to doubt such a proposition – the very act of doubting is an act of thinking – and yet *think* is a folk psychology concept. If eliminative materialism is true, then this most basic certainty is mistaken. But that just seems too counter-intuitive: The direct certainty of propositions about our own mental states, such as "I think", should take priority over physicalist considerations, such as "eliminative materialism is true".

*Possible **response** to the direct certainty of folk psychology:*

Eliminativism can respond that this response misunderstands the theory. Eliminative materialism doesn't deny the existence of the mental phenomena altogether – there is definitely *something* going on when we talk about "beliefs", "pain", "thought", etc. However, what eliminative materialism does say is that folk psychology is the wrong account of these mental phenomena. There is definitely something that's happening when Descartes is doing his philosophy and making arguments for claims such as "I exist" – so Descartes can be certain of that – but the reality of what's going on is some neuroscientific alternative: *thoughts* don't exist, and so "I *think*" is false.

Belief that eliminative materialism is true is self-refuting

If Churchland is arguing that we should believe eliminative materialism is true, then it seems his theory is self-refuting.

Eliminative materialism claims that beliefs *don't exist* – they're a mistaken folk psychology concept. But, in arguing for eliminative materialism, Churchland is expressing his belief in the truth of this theory. After all, why would anybody argue that something is true if they didn't *believe* it was true? Arguments are expressions of belief.

So, if Churchland *believes* that eliminative materialism is true, then this disproves his own theory: Churchland has proved that beliefs exist.

*Possible **response** to the claim that articulating belief in folk psychology is self-refuting:*

This response commits the fallacy of begging the question. It assumes the very thing it's trying to prove: that beliefs exist.

Again, eliminative materialism doesn't deny the existence of mental phenomena entirely – it just denies that folk psychology is the correct account of them. Eliminative materialism can accept that there is *something* going on in the case of what we currently call "beliefs", but it just disagrees that folk psychology is the correct model of what's going on. Neuroscience, not concepts such as "belief", are the more accurate theory of why people like Churchland assert claims such as "eliminative materialism is true".

However, this response is less effective at eliminating a closely related folk psychology concept: *intentionality*.

Eliminative materialism can't explain intentionality

Even if we accept the response above that saying things like "eliminative materialism is true" doesn't prove the existence of *beliefs,* we can still argue that it proves the existence of *intentionality*. And intentionality is a folk psychology concept that is seemingly impossible to eliminate.

Again, intentionality means the "aboutness" of mental states. For example, you can have a thought *about* a frog or a belief *about* aliens. These mental states have intentionality because they are about things, but eliminative materialism argues that intentionality is at odds with the rest of our scientific explanations. For example, we don't say atoms or forces or physical objects are *about* anything, and yet we say that beliefs and thoughts and desires *are* about things. So, eliminativists argue that a complete and consistent scientific picture of the world would eliminate the latter category (things with intentionality) in favour of explanations more like the former (things without intentionality).

But eliminating intentionality in this way is seemingly impossible. To even *make sense* of statements like "eliminative materialism is true" or "folk psychology is false" or "this is a more accurate scientific theory" *we must presuppose intentionality* – we must understand what these statements are *about*. So, the argument that eliminative materialism is self-refuting re-emerges: to even be able to talk about eliminative materialism requires intentionality, which is a folk psychology concept. It's not clear how neuroscience could ever offer an alternative account of intentionality, and so folk psychology can never be fully eliminated.

METAPHYSICS OF MIND – MODULE SUMMARY

Substance Dualism: Mind and body are two separate *substances*

- Arguments for substance dualism:
  - Conceivability argument: Mind without body is conceivable, so mind without body is possible, so mind can exist separately of body, so mind and body are separate substances
    - Replies:
      - Mind without body is not conceivable (e.g. behaviourism says mental states = behaviours, and behaviours need bodies)
      - Masked man fallacy: Just because I can *conceive* of Batman and Bruce Wayne as separate people, it doesn't follow that Batman and Bruce Wayne *are* separate people. Likewise, just because I conceive of mind and body as separate, doesn't mean they actually are
  - Divisibility/indivisibility argument: Body is a divisible substance, mind is an indivisible substance, so mind and body can't be the same substance
    - Replies:
      - Mind is actually divisible (e.g. multiple personality disorder or split-brain patients)
      - Some physical things are indivisible (e.g. sub-atomic particles) and so even if mind is indivisible doesn't prove it's non-physical
- Arguments against substance dualism:
  - Causal interaction (interactionist dualism: minds interact with body and vice versa)
    - Conceptual interaction problem: Physical things can only be moved by other things, so how can mind move body if mind is non-physical?
    - Empirical interaction problem: If non-physical mind interacts with physical body, it would violate law of conservation of energy because it would add energy into a closed system
  - Problem of other minds: If minds are separate to bodies, how can we know other people even have minds when we only perceive their bodies?
  - Ockham's razor: Why explain with *two* (mind and brain) what can be explained with *one* (brain only)

Property Dualism: Qualia are properties that are neither reducible to nor supervenient on the physical brain, i.e. qualia are non-physical *properties*

- Arguments for property dualism:
  - Zombies: Possible beings that are physically identical to humans but without qualia. If it's possible to be physically identical but lack qualia, then qualia are non-physical
    - Replies:
      - Zombies are inconceivable (e.g. type identity theory or hard behaviourism)
      - Zombies are metaphysically impossible: A physical duplicate without qualia (i.e. a philosophical zombie) is metaphysically impossible in the same way water without H2O is metaphysically impossible
  - Knowledge argument: Mary learns all the physical facts about seeing colour from within a black and white room, but when she leaves she learns something new (*what it is like* to see colour), and so this new knowledge is non-physical
    - Replies:
      - Mary would already know: If you really did know all the physical facts about seeing colour, you would already know *what it is like* to see colour
      - Mary's new knowledge is just (physical) *acquaintance* knowledge, like how you can't know a person from reading books about them
      - Mary's new knowledge is just (physical) *ability* knowledge, like how you can't learn how to ride a bike from reading books about bikes
      - New knowledge, old fact: Mary just learns a new conceptual way of looking at the same old (physical) facts
- Arguments against property dualism:
  - Introspection: Epiphenomenalism says qualia don't cause anything, but it seems qualia *do* have causal powers, e.g. my pain qualia can causes me to form a belief that "I am in pain"
  - Evolution: Why would we evolve qualia if epiphenomenalism is true? Qualia couldn't confer any evolutionary advantage because they don't have causal effects

Behaviourism: Mental states (e.g. pain) are behavioural dispositions (e.g. saying "ouch" or having a disposition to answer "yes" if someone asked you "did that hurt?")

- Arguments for behaviourism:
  - Category mistake: Oxford University analogy: if you've seen the buildings and teachers etc., then you've seen the uni. Likewise, if you've seen the behaviours, then you've seen the mental state. It's a category mistake to think mental states are something over and above these behaviours.
- Arguments against behaviourism:
  - Asymmetry between self/other knowledge: There are important differences between knowledge of one's *own* mental states vs. *others'* mental states. For example, I have to *infer* that someone else is in pain from their behaviour but I can know I'm in pain *directly* even if I have no knowledge of my behaviour. If mental states are just behaviours, how can this be?
  - Super Spartans: Perfect actors who have the qualia of pain, but no pain behaviour whatsoever. The fact that they have the mental state but not the behavioural dispositions shows there is more to pain than just behaviour
  - Circularity/multiple realisability: Mental states can be realised in multiple different ways, e.g. saying "ouch!" or not saying it. How someone actually does behave depends on *other* mental states they have, which *also* need to be analysed in terms of behaviour, and so on and so on

Type Identity: Mental states (e.g. pain) ontologically reduce to brain states (e.g. c-fibers) in the same way "lightning" reduces to "electrical discharge"

- Arguments for type identity theory:
  - Ockham's razor: Why explain with two (mind and brain) what can be explained with one (brain only)
- Arguments against type identity theory:
  - Divisibility argument: My mental state is not divisible, but my brain state is divisible, so my mental state cannot be the same thing as my brain state
  - Multiple realisability: Type identity theory says pain = c-fibers firing. However, an octopus can feel pain even though it doesn't have c-fibers. So, pain cannot be the same thing as c-fibers firing
  - Zombies: Type identity theory says pain = c-fibers. However, a philosophical zombie has c-fibers but not pain (qualia). So, pain can't be the same as c-fibers firing because you can have one without the other

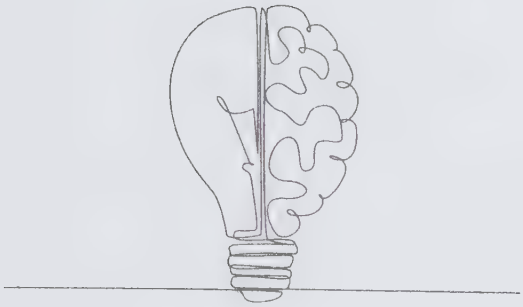
Functionalism: Mental states (e.g. pain) are functional states within the entire mind (e.g. a state that causes you to avoid bodily harm and a belief that you are in pain)

- Arguments for functionalism:
  - Multiple realisability: Functionalism correctly acknowledges that the same mental state can be realised in lots of different physical ways (e.g. human, octopus, and alien can all feel pain)
- Arguments against functionalism:
  - Inverted qualia: Functionalism says if two mental states are functionally identical, they are the same mental state. However, if my blue qualia is like your green qualia, we have a different mental state even though our mental states are functionally identical
  - Knowledge argument again: Mary has all the physical facts + *all the functional facts* about seeing red, but still learns something new when she leaves. Therefore, knowledge of the mental state of seeing red must be knowledge of something more than knowledge of a functional state
  - China brain: The China brain has a state that is functionally identical to pain, but isn't actually pain (it's just a load of people with radios imitating a brain), so there is more to mental states than their functional role

Eliminative materialism: Rather than trying to *reduce* folk psychology concepts (e.g. "pain" or "belief" or "think"), we should *eliminate* them in favour of more precise neuroscientific explanations

- Arguments for eliminative materialism (i.e. arguments against folk psychology):
  - Folk psychology hasn't progressed as a scientific theory: E.g. "pain" is the same concept the Romans had, for example, just a different name
  - Folk psychology can't explain all aspects of mental life: E.g. why some people are mentally ill and how people learn
  - Intentionality (how mental states can be "about" things) doesn't fit with the rest of science (e.g. gravity or temperature aren't *"about"* anything)
- Arguments against eliminative materialism (i.e. in defence of folk psychology):
  - Direct certainty of folk psychology: E.g. Descartes took "I think" to be the most basic certainty. The certainty of "I think" should take priority over the certainty of statements such as "eliminative materialism is true"
    - Response: Eliminativism doesn't deny mental phenomena *like* thoughts, only that folk psychology is the correct account of them

- Folk psychology has good predictive power: E.g. if someone has a *belief* that eating animals is wrong, they won't order chicken from the menu
  - Response: neuroscience will have even more predictive power than folk psychology as it progresses and we understand it better
- Self-refuting: E.g. by arguing for eliminativism, you are demonstrating a *belief* (a folk psychology concept) that eliminativism is true, thus proving that beliefs exist and that folk psychology is true
  - Response: This assumes the same thing it is trying to prove, i.e. that beliefs and other folk psychology concepts *exist*
- Intentionality: Eliminativism says intentionality is a problem because it doesn't fit with the rest of science, but intentionality is required to even make sense of claims such as "eliminative materialism is true" and "folk psychology is false"



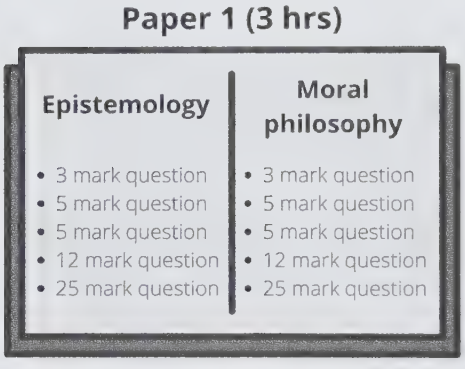
HOW TO GET AN A/A* IN THE EXAM

The exam papers contain 3 mark, 5 mark, 12 mark, and 25 mark questions. This section explains how to answer these different question formats and includes example answer plans for pretty much every 25 mark question that could come up in the exam.

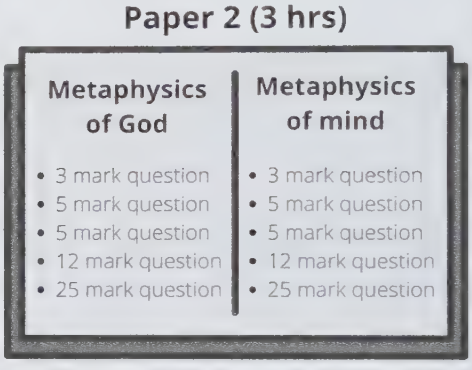
OVERVIEW

The AQA philosophy A-level assessment is sat at the end of the course and consists of two 3-hour papers.

Paper 1 covers the epistemology and moral philosophy modules and has 100 available marks. Paper 2 covers the metaphysics of God and metaphysics of mind modules and also has 100 available marks. The format of each paper is as follows:



100 available marks, 50% of overall grade



100 available marks, 50% of overall grade

The syllabus describes the two "assessment objectives" of these papers:

- Assessment objective 1 (AO1): *Knowledge and understanding*
- Assessment objective 2 (AO2): *Analysis and evaluation*

The 3, 5, and 12 mark questions are 100% AO1. This means that you only have to *explain* the idea or argument you are asked about - you don't have to assess that idea or argument or say whether it is correct or not.

It is only the 25 mark questions where AO2 comes in. In the 25 mark questions, you will have to philosophically *evaluate* ideas and *argue* towards a conclusion. There is a specific formula for how to construct these answers that is described below.

AO1: 3, 5, AND 12 MARK QUESTIONS

The 3, 5, and 12 mark questions are 100% AO1. This means you only have to *explain* an idea or argument. You don't have to evaluate whether that idea or argument is correct or incorrect (and you may get marked down for doing so).

Where it might get a bit confusing is when you are asked to *explain* an evaluative point. For example, you might be asked to explain how argument X presents a problem for theory Y. When this happens, don't fall into the trap of evaluating whether that argument works or whether the theory it concerns is correct or not – just *explain* what you are asked to explain and move on.

With all these AO1 questions, a good rule of thumb to start with is to *define* any key words in the question.

3 MARK QUESTIONS

3 mark questions will ask you to define a simple philosophical concept, such as substance dualism or analytic truth. Don't write too much – just a sentence or two explaining the idea and an example (if applicable) is sufficient to get full marks. Writing too much – especially if it is irrelevant – will actually lose you marks.

Marks	Levels of response mark scheme
3	A full and correct answer, given precisely, with little or no redundancy.
2	The substantive content of the answer is correct, but there may be some redundancy or minor imprecision.
1	Relevant, but fragmented, points.
0	Nothing written worthy of credit.

Example 3/3 responses

Note: You can also use the definitions in the glossary (page 345) as more examples of answers to 3 mark questions.

What is moral realism? (3 marks)

> *Moral realism is the metaethical position that mind-independent external moral properties and facts exist. For example, a moral realist may believe that the properties of "good" and "bad" exist independently of minds.*

What do eliminative materialists claim about mental states? (3 marks)

Eliminative materialists claim that some or all mental states that are supposed to exist according to folk-psychology, such as beliefs and intentions, do not exist and should be eliminated from a scientific description of the world.

What is philosophical scepticism? (3 marks)

Philosophical scepticism is used as an approach to test the strength of our knowledge claims and to understand the nature of knowledge and justification more clearly. The brain in a vat thought experiment is an example of philosophical scepticism.

5 MARK QUESTIONS

5 mark questions are very similar to 3 mark questions except you will be asked to explain a slightly more detailed concept or argument. Again, don't write too much – just define any key terms in the question and then explain the idea or concept.

Marks	Levels of response mark scheme
5	A full, clear and precise explanation. The student makes logical links between precisely identified points, with no redundancy.
4	A clear explanation, with logical links, but some imprecision/redundancy.
3	The substantive content of the explanation is present and there is an attempt at logical linking. But the explanation is not full and/or precise.
2	One or two relevant points made, but not precisely. The logic is unclear.
1	Fragmented points, with no logical structure.
0	Nothing written worthy of credit

Example 5/5 responses

Outline the differences between cognitivist and non-cognitivist accounts of moral language. (5 marks)

Cognitivism and non-cognitivism are accounts of what moral judgements, such as "murder is wrong", mean. On the cognitivist account, moral judgements are cognitive statements that aim to describe the world and are capable of being true or false. For example, "murder is wrong" is true if murder has the property of wrongness. On the non-cognitivist account, moral judgements express non-

cognitive attitudes that are neither true or false. For example, "murder is wrong" may express emotional disapproval of murder or a prescriptive instruction not to commit murder.

Explain how the inverted spectrum presents a problem for functionalism. (5 marks)

Functionalism analyses mental states in terms of functional roles within a system. The inverted spectrum is a scenario where two people have different qualia but otherwise identical mental states. For example, one person might be experiencing blue qualia and the other red. Despite the different qualia, both mental states are functionally identical because both people would both say "blue" when asked "what colour is the sky?", for example. This scenario presents a problem for functionalism because the two people clearly have different mental states despite being functionally identical. This shows there is more to mental states than their function.

Outline Descartes' version of the ontological argument. (5 marks)

Descartes' ontological argument is an a priori argument that derives the conclusion "God exists" from the idea/concept of God:

1. *I have the idea of God*
2. *The idea of God is the idea of a supremely perfect being*
3. *A supremely perfect being does not lack any perfection*
4. *Existence is a perfection*
5. *Therefore, God exists*

12 MARK QUESTIONS

12 mark questions are still 100% AO1. The topics you will be asked about will require you to write a bit more, but you still only have to explain an idea or argument – you don't have to evaluate whether it is correct or not.

These 12 mark questions are where you're most likely to slip into evaluation mode when you shouldn't. For example, say you get asked to explain how multiple realisability is a problem for type identity theory, just explain what type identity theory is and how the multiple realisability objection attacks it – nothing more. You shouldn't give your opinion on whether type identity theory is correct or not, or whether the multiple realisability objection is a successful objection or not. If you find yourself making arguments along these lines and giving your opinion on type identity theory then you've gone too far! Often, students panic that they haven't written enough and so start adding in this kind of evaluation – don't do this! It's a waste of time and will lose you marks.

As always, a good place to start your answer is by *defining* the key terms in the question.

Marks	Levels of response mark scheme
10-12	The answer is set out in a precise, fully-integrated and logical form. The content is correct and demonstrates detailed understanding. Points are made clearly and precisely. Relevance is sustained, with very little or no redundancy. Philosophical language is used precisely throughout.
7-9	The answer is set out in a clear, integrated and logical form. The content of the answer is correct and demonstrates detailed understanding. The content is clearly relevant and points are made clearly and precisely. Any lack of clarity with respect to particular points is not sufficient to detract from the answer. Relevance is largely sustained. There may be some redundancy, though not sufficient to detract from the answer. Philosophical language is used correctly throughout.
4-6	The answer is clear and set out in a coherent form, with logical/causal links identified. The content of the answer is largely correct and most points are made clearly. Relevance is not always sustained and there is some redundancy. Philosophical language is used correctly, with any minor errors not detracting from the response.
1-3	There are some relevant points made, but no integration. Some points are clear, but there is a lack of precision – with possibly insufficient material that is relevant or too much that is irrelevant. Philosophical language is used, though not always consistently or appropriately.
0	Nothing written worthy of credit.

Example 12/12 response

Explain how Kant's deontological ethics can be applied to the question of whether we should ever tell lies. (12 marks)

> *Kant's deontological ethics is concerned with duty. Kant claims that the only thing that is good in itself is good will, which means to act according to what duty demands. Kant argues that our duties are captured by the principle of the categorical imperative, one formulation of which says to "act only*

according to that maxim by which you can at the same time will that it should become a universal law." Applying the categorical imperative to the question of whether we should tell lies, Kant would argue that we cannot wish lying to become a universal law because this would lead to a contradiction in conception. The reason for this is that the whole point of telling a lie is to convince the other person that you are telling the truth but if everyone always lied then nobody would ever believe anyone. If nobody ever believed anyone, there would be no point in telling lies and so the concept would lose all meaning. This contradiction in conception means Kant would say we have a perfect duty never to lie and so it is never morally acceptable to lie in any circumstances. Further, telling lies also fails the humanity formula of the categorical imperative, which says to "act in such a way that you always treat others never simply as a means, but always as an end". Lying to someone uses them merely as a means to an end and ignores their status as a rational being because it deprives them of the opportunity to decide how to act based on the truth. Given that telling lies fails both these tests, Kant would argue that it is never acceptable to tell a lie no matter the consequences. So, for example, even if an axe-murderer came to your door asking for the location of your friend (who is hiding in your basement), it is still not acceptable to tell a lie even if you expect that doing so would save their life. The categorical imperative shows that lying is always morally wrong and so we are duty bound never to tell lies.

AO2: 25 MARK QUESTIONS

Whereas the other questions require you to *explain* an idea or argument, the 25 mark questions require you to *argue* for a particular philosophical viewpoint. In other words, you have to take a stance on a topic and argue that it is the correct one.

There is a pretty standard template for constructing these answers – and it's very important that your essay follows this template (at least roughly). If you don't include an introduction or a conclusion, for example, it's very difficult for an examiner to give more than around 10 marks out of 25.

To help construct answers according to this template, this section contains example essay plans for 25 mark questions on every major topic (see page 266 onwards). As I mention in that section, planning and memorising your own answers to these questions is perhaps the most effective revision technique – there are only so many possible 25 mark questions you can be asked and so it makes sense to have a rough idea what you're going to write for each topic before going into the exam.

THE 25 MARK TEMPLATE

The format of your 25 mark responses should look something like the template below:

- Introduction: Define key terms and then explain which side you are arguing for (E.g. "...In this essay I will argue for theory X")
- Argument *for* theory X
    - Possible response to this argument
        - Response to this response
- Argument *against* theory X
    - Response to this argument
- Conclusion: Theory X is correct because of the arguments above

You don't have to follow this template *exactly,* but there are several important features your essay will need to have if it is to achieve the top grade boundary: an introduction, a consistent argument, a balanced argument, and a conclusion.

Introduction

The most important thing to include in your introduction is a statement of what you will be arguing for in your essay – which side you are taking. For example, if you get asked a question on cosmological arguments, you might include one of the following:

- "...In this essay I will argue that cosmological arguments succeed in proving God exists"
- "...In this essay I will argue that cosmological arguments do not succeed in proving God exists"

Or you might opt to argue for a more nuanced position, such as:

- "...In this essay I will argue that although cosmological arguments succeed in proving that there was a first cause that created the universe, they do not succeed in showing that the first cause is God"
- "...In this essay I will argue that although cosmological arguments from causation fail, cosmological arguments from contingency succeed in proving God exists"

There's no right or wrong position to take – all that matters is how good the arguments for that position that follow are.

It's also worth defining any key terms in your introduction too. Although the 25 mark questions are primarily AO2, there are 5 marks available for AO1. Reeling off quick definitions of the terms involved in the discussion is an easy way to tick the knowledge and understanding box and pick up those 5 marks.

A consistent argument

> *"The student argues with clear intent throughout and the logic of the argument is sustained."*
>
> *21-25 mark grade boundary*

A key feature of an A* grade answer is a *consistent* argument. If you said in your introduction that you were going to argue *for* substance dualism, say, but then present a load of arguments *against* substance dualism that you don't respond to, your argument isn't consistent. You say you're going to argue one thing but then argue against it – that's not a strong argument.

So, pick a side and stick to it. Make sure your overall argument supports what you said you would argue for in your introduction and that the conclusion comes down on the side you originally said you were arguing for.

A balanced argument

> *"Arguments and counter-arguments are stated in their strongest forms. Reasoned judgements are made, on an ongoing basis and overall, about the weight to be given to each argument. Crucial arguments are clearly identified against less crucial ones."*
>
> *21-25 mark grade boundary*

However, even though you are consistently arguing for a particular viewpoint, you can't just ignore the arguments against your viewpoint. If you only present one side of the discussion then your argument isn't *balanced*.

But also, when you present the arguments against your viewpoint, don't *strawman* them. A strawman is when you set up a weak or inaccurate version of the opposing viewpoint and then respond to that rather than engaging with what the other side actually says. An A* grade argument considers the *strongest* arguments of the other side and then *responds* to those.

Getting the right balance between these two things can be hard. If you ignore or strawman the other side's arguments, then your argument won't be balanced. But if you make the other side's arguments *too* strongly, you might not be able to respond to them! This is another reason why it's so important to think through and memorise answer plans for each topic (such as the ones starting from page 266) *before* going into the exam.

An example of a balanced discussion (in this example arguing against indirect realism) would be something like the following:

- Scepticism: If indirect realism is true and all we perceive is sense data, we cannot know there is a mind-independent external world
    - Locke: The involuntary nature of experience suggests that my perceptions, unlike my imaginations, are caused by something *external* to my mind
        - Response to Locke: Even if there is an external cause of perceptions, indirect realism still can't justify its claim that this mind-independent external world *resembles* our sense data (you might be a brain in a vat, say) – so scepticism still stands

Overall, the original claim in this example stands: that indirect realism leads to scepticism about the mind-independent external world. But, as you can see, this claim is defended and refined against Locke's *response* from the involuntary nature of experience.

Conclusion

"The conclusion is clear, with the arguments in support of it stated precisely, integrated coherently and robustly defended."

21-25 mark grade boundary

Your conclusion should follow from the arguments you've presented and be consistent with the viewpoint you set out to defend in your introduction. Don't simply weigh up the arguments on both sides and then be like "well *on balance* theory X is stronger" – that's not a strong conclusion. If you've followed the template above, your arguments should lead to a decisive conclusion, such as "theory X is true" or "theory X is false" or "theory 1X is wrong but theory 2X is correct".

Finally, don't spend ages on your conclusion or introduce any new ideas here. Just summarise your main arguments and why they show that your viewpoint is correct – a few sentences is enough here.

GENERAL WRITING TIPS AND EXAMPLES

Once you've got the general 25 mark template down, the difference between a high grade and a top grade will be the level of *detail* and *precision* to your essay. The following are a couple of general tips to help with this.

Level of detail

"The student demonstrates detailed and precise understanding throughout."

21-25 mark grade boundary

It's better to cover a few arguments in *detail* rather than simply listing all the arguments you know without analysing them in any real depth. You don't have to cover every possible argument for or against a theory – that's not even possible. Instead, two or three arguments (and responses) is more than enough to achieve the top grade boundary if handled correctly.

The example answer plans (starting from page 266) should give you an idea of how many arguments to include and how to structure your paragraphs clearly. Take each argument one at a time and make sure you present it clearly and without missing any steps. At the end of the paragraph, be sure to explain how this argument relates to your overall essay.

If in doubt, "point, evidence, explanation" is a good template to use for your paragraphs. For example, you might make the conflicting virtues objection to Aristotle's virtue ethics in the following way:

Point:

> *An issue for Aristotle's virtue ethics is the question of conflicting virtues. If you find yourself in a situation where one virtue demands one course of action but a different virtue demands a completely different course of action, virtue ethics does not say what to do.*

Evidence/example:

> *For example, a judge who is sentencing a thief could either choose to punish the thief and act in accordance with justice (a virtue) or let the thief go and act in accordance with mercy (also a virtue). However, the judge can't both punish the criminal and let them go – he has to choose one or the other.*

Explanation:

> *Both justice and mercy are virtues, though, and so it seems that whichever action the judge takes he will fail to act virtuously in some way. Situations with conflicting virtues such as this strengthen the criticism that virtue ethics does not provide clear guidance for action because virtue ethics seems to suggest two contradictory courses of action in some circumstances.*

Precise use of philosophical language

> *"Philosophical language is used precisely throughout."*
>
> *21-25 mark grade boundary*

Precise use of philosophical language is the difference between a very good response and an excellent response. You can give a perfectly accurate explanation of something, such as:

> *Property dualism is the claim that mental properties are non-physical*

But you can also include a more *precise* term here, such as:

> *Property dualism is the claim that there are mental properties (qualia) that are neither physical nor supervene on the physical*

Don't obsess over precision too much, though – your overall argument is far more important to get right. There is a finite amount of time in the exam and you don't want to waste it trying to remember the exact definition of something. If you know a term like "epiphenomenalism" or "perfect duty" and it is relevant to your argument then make

sure to include it in your answer – but don't try to force these words into your essay if you don't understand them or they're not relevant.

Finally, precise philosophical distinctions can tie into the overall detail of your argument. Arguing for a conclusion that makes a precise distinction is a great way to demonstrate detailed understanding. For example:

- Arguing that cosmological arguments from *contingency* work but cosmological arguments from *causation* don't
- Arguing for *soft* behaviourism but rejecting *hard* behaviourism
- Arguing against *act* utilitarianism but for *preference* utilitarianism

In general, these kinds of nuanced conclusions might be an easier way to demonstrate a detailed understanding that gets high marks. That said, don't try to force this by arguing for some weird position that doesn't make sense – you can get 25 out of 25 by arguing for a simple conclusion too. A well-made argument for a simple conclusion (e.g. "theory X is true") is better than an attempt at precision that is mistaken or makes no sense.

EXAMPLE 25 MARK ANSWER PLANS

There are only so many 25 mark questions you can be asked given the syllabus content.

These 25 mark questions could be worded slightly differently to the ones below, but there are really only **20 or so** topics that could come up in the exam. Obviously, you can't know *which* of these 20 or so topics you will get asked about, so it makes sense to know what you are going to write for all possibilities.

It's hard enough *remembering* all the course content – let alone fitting that content into the blueprint above within the 45 minutes or so you get in the exam (assuming total time is evenly divided by the number of available marks). This is why it's so important to memorise plans – if not full essays – for each of the example 25 mark questions below. By preparing and memorising answers to each of these questions, you'll already have a balanced and consistent answer that argues towards a coherent conclusion. In other words, whatever the topic, you can go into the exam with A/A* grade answers locked and loaded.

In my experience of teaching the philosophy A-level course, this is the most effective use of revision time.

Epistemology answer plans

The definition of knowledge

How should propositional knowledge be defined? (25 marks)

- Introduction: I will argue that Sosa's virtue epistemology provides the correct definition of propositional knowledge
  - Define key terms: Propositional knowledge (vs. acquaintance or ability), Sosa's definition of knowledge
- Definition of knowledge 1: Justified true belief
  - Response: Gettier case 1: Smith's belief "the man who gets the job has 10 coins in his pocket" is justified and true, but not knowledge
    - Definition of knowledge 2: Justified true belief + no false lemmas: Adding a no false lemmas condition rules out Gettier case 1 as knowledge because Smith's belief is inferred from the false lemma "Jones will get the job"
      - Response: Fake barn county: Henry's belief "there's a barn" is true, justified, and not inferred from a false lemma, but is not knowledge
- Definition of knowledge 3: Sosa's virtue epistemology: Explain how Sosa's definition avoids both the Gettier case and fake barn county examples above because they are not *apt*
- Conclusion: Sosa's virtue epistemology avoids Gettier cases and fake barn county and is the correct definition of propositional knowledge

Direct realism

Assess direct realism. (25 marks)

- Introduction: I will argue direct realism is false
  - Define key terms: Mind-independent objects, direct perception
- Argument against direct realism 1: Time lag. E.g. the sun looks different to our perception of it because it takes the light from the sun 8 minutes to reach earth
  - Response: Time lag confuses *what* we perceive (e.g. mind-independent sun) with *how* we perceive it (via light waves). Direct realism can still claim that we perceive the sun directly, but as it was moments ago
- Argument against direct realism 2: Illusion. E.g. stick looks crooked in water but is not crooked in reality
  - Response: Relational properties: the stick has the (mind-independent) relational property of *looking* crooked to a perceiver looking at it from that perspective
    - Response to response (argument against direct realism 3): Hallucination: A hallucination is like an illusion that has no bearing in the (mind-independent) external world. But unlike the other arguments you cannot reply e.g. that "my sofa has the relational property of looking like it has gremlins on it" if you are hallucinating seeing gremlins on a sofa - that would be absurd!
- Conclusion: Direct realism is false because it cannot account for hallucinations

Indirect realism

Assess indirect realism. (25 marks)

- Intro: I will argue indirect realism is the best explanation of perception, and does not lead to scepticism
  - Define key terms: Indirect realism, sense data, mind-independent objects
- Argument for indirect realism: Perceptual variation
- Argument against indirect realism: Scepticism/veil of perception: how do we know there are mind-independent objects that cause our perceptions if all we perceive is the sense data?
  - Response: Locke: involuntary nature of experience: When I imagine things, I can choose what I perceive, but when I look at things in the real world, I cannot choose what I perceive. This suggests that the cause of my perceptions is something external to me
    - Response to response: OK, this may prove there's an external something causing perception - but how do we know that this external world is anything like our perception?
      - Response to that: Russell: External world is the best hypothesis: The hypothesis that my cat really does exist and walked from the floor to the sofa when I wasn't looking is a more likely explanation for why the cat is now on the sofa than the explanation that I am having random perceptions of cats disappearing and reappearing on the sofa randomly
- Conclusion: Indirect realism is the best theory of perception because it can account for hallucinations/perceptual variations. It does not lead to scepticism because a mind-independent external world causing sense data is the best hypothesis (Russell)

Idealism

Assess idealism. (25 marks)

- Introduction: I will argue that idealism is false
  - Define key terms: Idealism
- Argument for idealism: The master argument: The idea of an unperceived mind-independent object is inconceivable
  - Response: Berkeley's master argument conflates the idea of an object with the object itself. Just because the *idea* of a mind-independent object is impossible, doesn't mean object *itself* is impossible.
- Argument against idealism: Illusion: If "to be is to be perceived", as idealism says, then we have to say that a pencil that looks crooked in water really is crooked
  - Response: Berkeley would agree that our perception of the crooked pencil is just as real as the perception of the normal pencil
    - Response to response: OK, but what about hallucination and dreaming? Are we really supposed to believe that my hallucination of gremlins or my dream about flying is equally as real as my ordinary perceptions?
- Conclusion: Berkeley's master argument fails to prove that the existence of mind-independent objects is impossible. Further, idealism fails as a theory of perception because it is unable to suitably explain hallucinations and dreams.

Innatism vs. empiricism

Do we have innate knowledge? (25 marks)

- Introduction: In this essay I will argue that we have innate knowledge of necessary truths but that experience is necessary to articulate this knowledge
    - Define key terms: Innatism, empiricism
- Plato's argument for innatism: Meno's slave has never been taught geometry, yet has knowledge of the proof that a square with an area of 2A will have sides that are equal to the diagonal of a square with an area of 1A. This knowledge can't come from experience because he has no experience of geometry
    - Response: Locke: If we had innate knowledge, it would be universal. But "children and idiots" don't know the theorems of geometry - so it can't be innate
        - Response to Locke: But knowledge of necessary truths can't come from experience, though: Experience can only tell us how things are (contingent truths) not how they must be (necessary truths). But we know e.g. that "1+1=2" is a necessary truth - and this knowledge can't come from experience
            - Locke's response to Leibniz: Knowledge of "1+1=2" requires the concepts "1", "+", "=", and "2" but there is no evidence that newborn babies have concepts beyond those experienced in the womb (e.g. warmth and pain)
                - Leibniz's response to Locke: Obviously, a newborn baby can't verbally articulate "1+1=2", but this doesn't mean the concept isn't there. We innately know these concepts and over time we learn to recognise them and make them explicit, but they were always there in the mind.
- Conclusion: We are born with innate concepts and knowledge of necessary truths, however it may require experience to articulate this knowledge



Rationalism vs. empiricism

Can intuition and deduction establish synthetic truths a priori? (25 marks)

- Introduction: I will argue that there is at least one instance of synthetic a priori knowledge and therefore that rationalism is true and empiricism is false
    - Define key terms: Analytic, synthetic, a priori, a posteriori, empiricism, rationalism
- Argument for rationalism 1: Descartes' trademark argument. "God exists" is a synthetic truth, and Descartes claims his reasoning is purely a priori. So the trademark argument proves a synthetic truth a priori.
    - Empiricist response: The concept of God is not a priori as Descartes claims but instead comes from experience (for example, there are entire cultures that lack the concept of God). This shows Descartes' trademark argument is not a priori and so the conclusion cannot be an example of a synthetic a priori truth.
- Argument for rationalism 2: Cogito ergo sum. "I exist" is a synthetic truth, and Descartes' cogito argument is purely a priori
    - Empiricist response (Hume): "I exist" does not follow a priori from "I think" – it's logically possible the thought could just be disembodied. So, "I exist" cannot be known a priori
        - Response to Hume (defence of rationalism): OK, but "this thought exists" is technically a synthetic truth, and can be proved through a priori means alone. So, if we change the argument slightly, we still have a synthetic a priori truth.
- Conclusion: Empiricism claims that there are no synthetic truths known a priori. But "this thought exists" is a synthetic truth that can be known entirely a priori. Therefore, empiricism is false.

The limits of knowledge

Do any arguments successfully defeat global scepticism? (25 marks)

- Introduction: In this essay I will argue that knowledge of the external world is possible if formed via a reliable method, and so global scepticism can be defeated
    - Define key terms: Local scepticism, global scepticism
- Sceptical argument: Descartes' three waves of doubt (illusion, dreaming, evil demon). Evil demon = global scepticism. If we are being deceived by an evil demon then we can't know anything, because everything we have experienced is a lie/illusion/deception
    - Descartes' response to scepticism: Argument for the existence of God (e.g. Trademark, ontological - take your pick). God exists and God wouldn't let me be globally deceived in such a way
        - Sceptical response to Descartes' response: Descartes' arguments do not prove God exists (challenge a premise from whichever of Descartes' arguments you picked). If we can't know God exists, then can't know God is protecting us from being deceived. If we can't know we're not being deceived, then we're back at square one, i.e. the evil demon, i.e. global scepticism
            - Reliabilist response to scepticism: Knowledge = true belief formed via a reliable method. If I'm not being deceived, then my perception is reliable. So, if I am not being deceived, then I can know "water boils at 100°c" if I perceive it (because my perception is a reliable method, and my belief is true).
- Conclusion: Even though Descartes' responses to global scepticism fail, we can still have knowledge of the external world because knowledge = true belief formed via a reliable method. As long as we are not being deceived (even if we can't know whether we are or not), then our perception is a reliable method of forming true beliefs. Therefore, we can have knowledge (i.e. a true belief formed via a reliable method) of the external world.

Moral philosophy answer plans

Utilitarianism

How convincing is utilitarianism as an account of what makes an action morally right? (25 marks)

- Introduction: I will argue that utilitarianism (both act and rule) is not a convincing account of moral action
    - Define key terms: Consequentialist, act utilitarianism, rule utilitarianism
- Argument for utilitarianism: Intuitive appeal (happiness = good seems intuitive) and fairness (treats everyone's happiness as equally valuable)
- Argument against utilitarianism 1: Difficult to calculate. E.g. how do you predict the consequences of your actions? Especially to the extent where you can predict e.g. the intensity of the pleasure vs. the duration, for example
- Argument against utilitarianism 2: Tyranny of the majority. E.g. punishing an innocent person is clearly wrong in itself. But utilitarianism says doing so would be morally acceptable if doing so made a large crowd happy
    - Response: Rule utilitarianism can avoid this problem. The rule "Don't punish innocent people" leads to greater happiness than if we didn't follow this rule
        - Response: But rule utilitarianism leads to rule fetishism. E.g. you can't steal or lie even to save a life
- Conclusion: Utilitarianism – both the act and rule versions – is not a convincing account of moral action

Kant's deontological ethics

Assess Kant's deontological ethics. (25 marks)

- Introduction: I will argue that Kant's deontological ethics is not a convincing account of what makes actions morally right or wrong
  - Define key terms: Categorical imperative, the good will
- Brief explanation of Kant's theory: "it's OK to steal" leads to a contradiction in conception
  - Argument against Kant's theory 1: Changing the maxim to avoid perfect duty (e.g. "it's OK to steal if it's Tuesday and your name is Christian")
    - Kant's response: This is not the real maxim – your name and the day of the week is not part of your choice
      - Response to response: "It's OK to steal to save your life" is an honest maxim and one that can be universalised without contradiction
- Argument against Kant's theory 2: Ignores consequences/utilitarianism. E.g. you can't tell white lies or tell a lie even if it means saving someone's life
  - Kant's response: The only thing good without qualification is the good will. Consequences such as happiness are not good without qualification (e.g. happiness derived from hurting others is bad) and so we should always act according to good will rather than try to anticipate consequences
    - Response to response: But there are clearly other values besides good will. For example, it is better to visit your friend in hospital because you genuinely care rather than because of duty/good will
- Conclusion: Kant's deontological ethics fails as an account of moral action because it does not provide clear rules and ignores consequences and other valuable motivations for action



Aristotle's virtue ethics

Is Aristotle's Virtue Theory an accurate account of moral responsibility? (25 marks)

- Intro: I will argue Aristotle's virtue theory is correct and defend it against objections
  - Define key terms: Agent centred, virtue, golden mean, eudaimonia
- Argument against Aristotle 1: No clear guidance: How do we know, e.g., when too much courage becomes acting rashly? No guidance in a moral dilemma (e.g. stealing to feed a hungry orphan or killing one person to save 20)
  - Response: Virtue theory was never intended as a set of rules. Anyway, we can still reflect e.g. "is this action courageous?" or kind or whatever which could help decide how to act
- Argument against Aristotle 2: Conflicts between virtues: E.g. justice is a virtue, so is mercy. So if you're a judge and someone has stolen something, you have to choose between justice (i.e. punishing the criminal) and mercy (i.e. letting the criminal go). What then?
  - Response: Conflicts between virtues are impossible. Practical wisdom tells you what each virtue tells you to do and in what amount. E.g. act according to justice but show appropriate mercy for circumstances for example
- Argument against Aristotle 3: Difference between eudaimonia (i.e. a good life for me) and being morally good. E.g. a nurse who spends her entire life saving lives and helping others.
  - Response: Aristotle does not say eudaimonia is all about self interest/the best life for me. Being virtuous and achieving eudaimonia involves commitment to others. E.g. being fair, just, kind, etc. are virtues, but they are not the *only* component of eudaimonia
- Conclusion: Aristotle's theory is the correct account of moral responsibility, and objections to the theory do not succeed in refuting it

Metaethics

What is the meta-ethical status of moral statements such as "murder is wrong"? (25 marks)

- Introduction: In this essay I will argue moral anti-realism is true and moral realism is false
  - Define key terms: Moral realism, moral anti-realism
- Argument for moral realism 1: Mill's proof of utilitarianism shows that moral properties exist and are natural properties (e.g. good = happiness)
  - Response: Naturalistic fallacy (Moore). Just because happiness is *correlated* with good actions, that doesn't mean goodness and happiness are the same thing
    - Argument for moral realism 2: Moore: Moral properties are non-natural properties. We know about these properties through rational intuition
      - Response: Mackie: But non-natural moral properties would be metaphysically and epistemically queer – "intuition" doesn't explain anything
- Argument *against* moral anti-realism: Moral anti-realism leads to nihilism. If anti realism is true, there is no reason not to go on a murderous rampage
  - Response 1: Just because we don't *want* moral nihilism to be true, thus doesn't show moral anti-realism is false
  - Response 2: But even if there are no mind-independent moral properties, there are still all sorts of reasons one might choose to act in a way we typically consider to be good, e.g. empathy, care, and the law
- Conclusion: Mind-independent moral properties – whether they be natural or non-natural – do not exist. Moral anti-realism is true but does not lead to nihilism or chaos.

Metaphysics of God answer plans



The concept of God



Is the concept of God coherent? (25 marks)

- Introduction: I will argue that the concept of God is coherent
  - Define key terms: Omnipotent, omniscient, omnibenevolent, eternal or everlasting
- Explanation of the concept of God: Omnipotent, omniscient, omnibenevolent, eternal/everlasting
  - Problem 1: The problem of the stone – can God create a stone so heavy He cannot lift it?
    - Response: Mavrodes: "a stone an omnipotent being cannot lift" is a contradiction, and omnipotence means God can do anything that's logically possible
  - Problem 2: Free will vs. omniscience: How can humans have free will if God already knows what we're going to do next?
    - Response: God's omniscience means God knows everything it is possible to know. It isn't possible to know the future (because free will)
  - Problem 3: The euthyphro dilemma: Either God creates morality (in which case "God is omnibenevolent" is arbitrary) or God follows morality (in which case God would not be omnipotent because He would not be powerful enough to make murder good)
    - Response: God creates morality (so He is omnipotent), but this doesn't make morality arbitrary, because God chooses the rules for other reasons (e.g. Love)
- Conclusion: God created morality does not contradict the claim that God is omnibenevolent, and the claim that humans have free will does not contradict the claim that God is omniscient. So, the concept is coherent.

Ontological arguments

Do ontological arguments prove God exists? (25 marks)

- Introduction: I will argue ontological arguments do not prove God exists
  - Define key terms: Ontological argument, a priori
- Ontological argument 1: Anselm's ontological argument: God is a being greater than which cannot be conceived, it is greater to exist in reality than the mind, therefore God exists
  - Response: Gaunilo's island: The perfect island is an island "greater than which cannot be conceived" – does that mean the perfect island must exist too?
  - Response 2: Existence is not a predicate (Kant): Adding existence to a concept does not make that concept any more real
- Ontological argument 2: Norman Malcolm's ontological argument avoids Kant's objection because *necessary existence* is a real predicate
  - Response: Invalid inference (different meanings of "necessary existence" between premises and conclusion)
  - Response 2: The concept of God is self-contradictory (e.g. problem of the stone)
- Conclusion: Ontological arguments - both Anselm's and Malcolm's - fail to prove the existence of God from the definition of God

Teleological arguments

Does the order of the universe prove that God exists? (25 marks)

- Introduction: I will argue teleological arguments do not prove God exists
  - Define key terms: Teleological argument, spatial order, temporal order
- Teleological argument 1: William Paley's watch analogy: The hallmark of design is many parts organised to serve a purpose. Aspects of nature have this hallmark
  - Response: Darwin's theory of evolution can explain how matter can organically organise itself into complex configurations to serve a purpose without needing a designer
    - Response to response (teleological argument 2): Swinburne's temporal order: But the *laws* of nature (e.g. gravity) did not evolve and yet are so perfectly suited to sustaining life that they must have been designed
      - Response: Multiple universes: If multiverse theory is true, there are a trillion universes with varying physical laws etc. so it's inevitable that at least one would have temporal order
- Argument *against* teleological argument: But even if the universe is designed, the designer might not be God
- Conclusion: Both teleological arguments from spatial and temporal order fail. Further, even if they succeeded in showing that nature is designed, it does not necessarily follow that the designer is God

Cosmological arguments

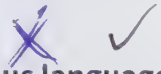
Do cosmological arguments succeed in proving the existence of God? (25 marks)

- Introduction: I will argue cosmological arguments from contingency succeed in proving God's existence
  - Define key terms: Cosmological argument (causation and contingency versions)
- Cosmological argument 1: Kalam argument or Aquinas' second way
  - Response 1: A first cause isn't necessary, there could be an infinite sequence
    - Response to response: But it is impossible for infinite amount of time to pass
  - Response 2: "Everything has a cause" is not an analytic truth (Hume). There is no contradiction in the idea of an effect without a cause and so it is logically possible that the universe doesn't have a cause of its existence.
    - Response to response (cosmological argument from contingency): Leibniz's principle of sufficient reason: Whether there is a first cause or not, the existence of the universe is contingent, and we can only provide sufficient reason for the universe's existence by appealing to a necessary being (God)
- Conclusion: Although it is logically possible for the universe to exist without a cause, this would not provide sufficient reason for why the universe exists. Only a necessary being (God) can provide sufficient reason for the existence of the universe. So, even if cosmological arguments from causation fail, cosmological arguments *from contingency* succeed in proving that God exists.

The problem of evil

Does the existence of evil mean that God does not exist? (25 marks)

- Introduction: I will argue that the problem of evil - both the logical and evidential versions - fails to disprove God's existence
  - Define key terms: Moral evil, natural evil
- Logical problem of evil: Evil is *logically incompatible* with the existence of an omnipotent and omnibenevolent God (Mackie/inconsistent triad)
  - Response: Free will theodicy: Without the ability to be evil, we couldn't have free will. And free will is so good that it outweighs the evil that people do as a result of it. Therefore, an omnipotent and omnibenevolent God could possible allow evil
    - Response to response: OK, free will explains *moral* evils. But how to explain *natural* evils (e.g. earthquakes and floods)?
      - Response to that: natural evils are another form of moral evil committed by e.g. the Devil or demons. Note: Responses to the logical problem of evil don't have to be plausible, they only have to be *possible* (because the claim of the logical problem of evil is that these three things are logically incompatible)
- Evidential problem of evil: God could logically allow some evil but the sheer amount of evil and how it is distributed suggests God probably doesn't exist
  - Response: Soul making: Our purpose in life is to develop morally and spiritually. Pointless evils are needed to develop virtues such as faith and deep sympathy
- Conclusion: Both the logical and evidential problems of evil fail to disprove God's existence.



Religious language

Is religious language meaningful? (25 marks)

- Intro: I will argue that Flew successfully shows religious language is unfalsifiable, but that this does not show religious language is meaningless
- Argument that religious language is meaningless: Flew's jungle clearing analogy shows that "God exists" is unfalsifiable
  - Response: Hick. Eschatological verification. "God exists" is unfalsifiable and meaningless if false, but verifiable and meaningful if true.
    - Response to Hick: OK, but "God exists" is empirically unfalsifiable in this life, and we only have empirical evidence to go on
      - Response to this: Hare's bliks and university lecturers example. Even if "God exists" is unfalsifiable, it clearly means something to people who have that blik (example: they go to church on Sunday because of that blik)
- Conclusion: Even if Flew successfully shows "God exists" is unfalsifiable, Flew does not succeed in showing "God exists" is meaningless. So, yes, religious language is meaningful

Metaphysics of mind answer plans

Substance dualism

Assess substance dualism. (25 marks)

- Introduction: I will argue that substance dualism is false
  - Define key terms: Substance dualism
- Substance dualism argument 1: Descartes' conceivability argument: Mind without body is conceivable, therefore possible, so mind without body can exist separately
  - Response: Masked man fallacy: Batman is not Bruce Wayne is conceivable, but that doesn't mean it's possible IRL or in any way a guide to reality
- Substance dualism argument 2: Descartes' divisibility argument: Body is divisible substance, mind is indivisible substance, therefore mind and body are separate substances
  - Response: But the mind is divisible (e.g. multiple personality disorder)
  - Response 2: Plus, even if the mind is indivisible, that doesn't necessarily mean it is a non-physical substance (e.g. subatomic particles are indivisible but physical)
- Argument against substance dualism: Conceptual interaction problem: Physical things can only be moved by other physical things, so if the mind is non-physical how does it move the body?
- Conclusion: Both the conceivability argument and the divisibility argument fail to prove substance dualism is true. Further, the conceptual interaction problem shows substance dualism is false.

Property dualism

Assess property dualism. (25 marks)

- Introduction: I will argue that property dualism is false
  - Define key terms: Qualia, supervenient, property dualism
- Property dualism argument 1: Philosophical zombies: It's conceivable and possible that there could be a physically identical human being (in another possible world) that lacks qualia
  - Response: Zombies are not possible. Physically identical but without qualia is like water without H2O – it's metaphysically impossible
- Property dualism argument 2: Knowledge argument (Mary)
  - Response: Mary's new knowledge is ability knowledge
    - Response to response: Someone could lack these abilities but still know what it's like to see colour
  - Response 2: Mary's new knowledge is acquaintance knowledge. Even if the knowledge argument succeeds in showing that Mary can't learn about qualia from the physical facts alone, this doesn't show that knowledge of qualia is non-physical. Mary's new knowledge is just (physical) acquaintance knowledge.
- Conclusion: Both the zombie argument and the Mary argument fail to prove property dualism is true.

Behaviourism

"Mental states are nothing more than behavioural dispositions." Discuss. (25 marks)

- Introduction: I will argue that behaviourism is wrong
  - Define key terms: Behaviourism, behavioural dispositions
- Argument against behaviourism : Philosophical zombies show that you can have all the behavioural dispositions but not the qualia/mental state. So, the mental state must be more than just the behavioural dispositions.
  - Response: Ryle: This makes a category mistake (Oxford Uni example). It's a mistake to think that the mental state is something separate to the behavioural dispositions. Zombies are inconceivable.
    - Response to response: But Super Spartans feel pain but have no behavioural dispositions. So it's not a category mistake: there is more to mental states than just behaviours.
- Argument against behaviourism 2: Asymmetry between self-knowledge and knowledge of others. E.g. I know I'm in pain *directly*, but I have to *infer* when others are in pain
  - Response: Ryle: This asymmetry is an illusion. We might know someone else is in pain by hearing them say "ouch!" *out loud*, whereas self-knowledge of pain is more like some inner voice saying "ouch!". But the mechanism is basically the same for both: we pay attention.
    - Response to response: Knowledge argument (qualia): No amount of knowledge of a person's behaviour and behavioural dispositions could give knowledge of that person's qualia, so there will always be an asymmetry between self- and other- knowledge
- Conclusion: This essay has discussed several examples that show mental states are not the same as behavioural dispositions, such as super Spartans, the difference between self-knowledge and other .

Type identity theory

Are mental states identical to brain states? (25 marks)

- Introduction: I will argue that type identity theory is wrong
  - Define key terms: Type identity theory, ontological reduction, c-fibers
- Argument for type identity theory: Ockham's Razor: Dualist theories explain and predict the same things using mind and brain as type identity theory does with brain alone
- Argument for type identity theory 2: Conceptual interaction problem: If the mind is a separate substance from the brain, how does the mind cause changes in the brain?
  - Response: These 2 arguments above support all physicalist theories, not just type identity
- Argument against type identity theory 1: Location problem: My brain (and brain states) have an obvious physical location, but my mind (and mental states) don't. So, assuming Leibniz's law of identicals, my mind cannot be the same thing as my brain
- Argument against type identity theory 2: Multiple realisability: An octopus, an alien, a dog, and a human can all feel pain but they don't all have c-fibers, so the mental state of pain cannot be the same thing as the brain state of c-fibers firing
- Conclusion: Mental states are not identical to brain states because they do not share all the same properties and are multiply realisable. Therefore, type identity theory is wrong



Assess the claim that mental states are functional states. (25 marks)

- Introduction: I will argue that functionalism is wrong
  - Define key terms: Functionalism
- Argument for functionalism: Multiple realisability: An octopus can feel "pain" and so can a human, but both have completely different brains. Therefore mental states (e.g. "pain") are better understood by their function rather than any specific physical realisation.
  - Response: Inverted qualia: If we had inverted qualia (e.g. my blue qualia is like your red qualia and vice versa), our mental states would be functionally identical - e.g. they would both cause us to say "the sky is a lovely shade of blue". However, despite being functionally identical, they would not be the same mental state. So, mental states are not simply functional states.
    - Response to response: Inverted qualia is impossible to prove
      - Argument against functionalism 2: China brain: The China brain is functionally identical to a human in pain, say, but is not in pain. And unlike inverted qualia you could, in theory, set this up to disprove functionalism. Therefore, mental states =/= functional states either.
- Argument against functionalism 3: Knowledge argument: Knowledge of all the physical and functional facts about e.g. seeing blue would still not be enough for Mary to know what seeing blue is like. Therefore the mental state of seeing blue must be something over and above the physical and functional facts
- Conclusion: There is more to mental states than functions (namely, qualia) and so functionalism is wrong

Eliminative materialism

Can talk of mental states be eliminated? (25 marks)

- Introduction: In this essay I will argue that eliminative materialism is false and that folk psychology cannot be fully eliminated
  - Define key terms: Folk psychology, eliminative materialism
- Argument *for* eliminative materialism: Folk psychology vs neuroscience as scientific theories. Good scientific theories progress over time, but folk psychology hasn't changed in thousands of years.
  - Response: But folk psychology has good predictive power. For example, if someone has a *belief* (folk psychology concept) that eating animals is wrong, then we can predict that they won't eat meat. Neuroscience, in contrast, can't predict behaviour very well at present
    - Response to response: neuroscience only has weak predictive power because it is a new science. As it progresses, neuroscience will be able to explain everything folk psychology can plus several things it can't such as mental health and intelligence
      - Response to that: Until this happens, folk psychology is the better scientific theory and cannot be eliminated
- Argument *against* eliminative materialism: Self-refuting. In arguing for eliminative materialism, you are demonstrating a *belief* that eliminative materialism is true. But belief is a folk psychology concept.
  - Reply: This begs the question. You can say that stating "folk psychology is false" is evidence of a *belief* that folk psychology is false, but this assumes the conclusion that beliefs exist in the first place! Further, eliminativism doesn't deny mental phenomena altogether, it just denies that folk psychology (e.g. "belief") is the accurate account
    - Reply to reply: To even make sense of the claim "folk psychology is false", we must assume intentionality - some sort of mental phenomena *about* folk psychology. But neuroscience offers no plausible account of intentionality
- Conclusion: folk psychology is currently a better scientific theory than eliminative materialism in terms of predictive power. Further, even if neuroscience advances, it is doubtful whether it could ever account for intentionality. So, folk psychology cannot be fully eliminated and so eliminative materialism is false.

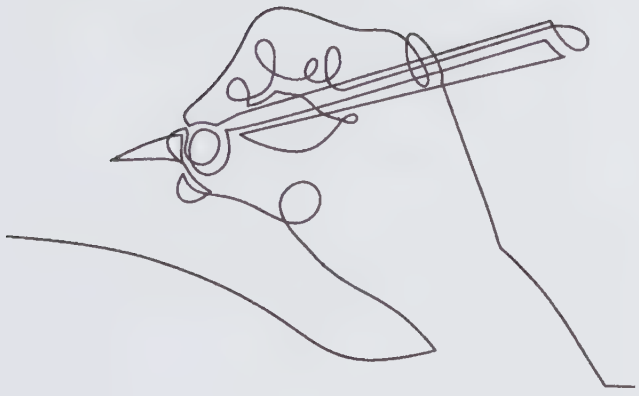
EXAM TECHNIQUE SUMMARY

AO1: *3, 5, and 12 mark questions*

- *Explain* only – if you find yourself evaluating whether an argument or theory is correct or not, you've gone too far!
- Start by defining any key terms in the question
- Use an example if appropriate
- Only talk about what the question asks – you lose marks for irrelevant points

AO2: *25 mark questions*

- These questions require you to *evaluate* and *argue* for a particular philosophical viewpoint
- Generic answer template (see from page 266 for example templates for specific topics):
    - Introduction: Define key terms and then explain which side you are arguing for (E.g. "In this essay I will argue for theory X")
    - Argument *for* theory X
        - Possible *response* to this argument
            - *Response* to this response (in defence of theory X)
    - Argument *against* theory X
        - *Response* to this argument (in defence of theory X)
    - Conclusion: Theory X is correct because of the arguments above
- Within this template, only make one argument/point per paragraph. If in doubt, follow the point, example, explanation template (page 263)
- Quality of argument > quantity of arguments. It's better to cover two or three arguments in *detail* (e.g. by considering responses and responding to them) rather than simply listing all the arguments you know without any real depth
- Use precise philosophical language where you can – but don't use words if you're not fully sure what they mean



PRACTICE QUESTIONS

This section contains 3, 5, 12, and 25 mark practice questions organised by module.

EPISTEMOLOGY

3 MARK QUESTIONS

- What is philosophical scepticism? (3 marks)
- What is a priori knowledge? (3 marks)
- What is the tripartite view of propositional knowledge? (3 marks)
- What is solipsism? (3 marks)
- What is an analytic truth? (3 marks)
- What is empiricism? (3 marks)
- What is acquaintance knowledge? (3 marks)
- What is ability knowledge? (3 marks)
- What is the no false lemmas (J+T+B+N) definition of knowledge? (3 marks)
- What is virtue epistemology? (3 marks)
- What is reliabilism? (3 marks)
- What is direct realism? (3 marks)
- What is indirect realism? (3 marks)
- What is idealism? (3 marks)
- What is innatism? (3 marks)

5 MARK QUESTIONS

- Outline the argument from illusion against direct realism. (5 marks)
- Explain Locke's distinction between simple and complex concepts. (5 marks)
- Explain Russell's claim that the existence of the external world is the best hypothesis. (5 marks)
- Explain what Plato is trying to show about knowledge in his Meno's slave argument. (5 marks)
- Explain how Berkeley's idealism differs from indirect realism. (5 marks)
- Explain how indirect realism differs from Berkley's idealism. (5 marks)

- What is Hume's Fork? (5 marks)
- How does adding a "no false lemmas" condition to the tripartite definition of knowledge avoid Gettier problems? (5 marks)
- Explain how illusion and hallucination present problems for direct realism. (5 marks)
- Explain how perceptual variation is a problem for direct realism. (5 marks)
- Explain the time lag argument against direct realism. (5 marks)
- Explain Locke's distinction between primary secondary qualities. (5 marks)
- Explain how Berkeley's idealism could lead to solipsism. (5 marks)
- What is the intuition and deduction thesis? (5 marks)
- Briefly outline Descartes' cogito argument. (5 marks)

12 MARK QUESTIONS

- Explain Berkeley's attack on Locke's distinction between primary and secondary qualities. (12 marks)
- Explain Locke's arguments against innatism. (12 marks)
- Outline Descartes' evil demon argument and explain what he then says about his knowledge of his own existence. (12 marks)
- Outline Berkley's idealism and explain how it could lead to solipsism. (12 marks)
- Briefly outline the tripartite view of knowledge and explain how a case of a lucky true belief (a Gettier-style problem) can be used to argue against this. (12 marks)
- Briefly explain direct realism and explain how the argument from hallucination can be used to argue against this view. (12 marks)
- Explain how one of Gettier's original counter examples attacks the tripartite view of knowledge. (12 marks)
- Explain the argument from Berkeley that we cannot know the nature of mind-independent objects. (12 marks)
- Outline and explain the key differences between indirect realism and idealism. (12 marks)
- Explain how Russell's argument that the external world is the best hypothesis supports indirect realism and avoids scepticism. (12 marks)

- Outline indirect realism and explain Locke's responses to the claim that indirect realism leads to scepticism. (12 marks)
- Explain how reliabilism could provide a response to philosophical scepticism. (12 marks)
- Briefly outline each of Descartes' three waves of doubt. (12 marks)

25 MARK QUESTIONS

- How should propositional knowledge be defined? (25 marks)
- Is innate knowledge possible? (25 marks)
- Can intuition and deduction establish synthetic truths a priori? (25 marks)
- Assess direct realism. (25 marks)
- Assess indirect realism. (25 marks)
- Assess idealism. (25 marks)
- Do any arguments successfully defeat global scepticism? (25 marks)

MORAL PHILOSOPHY

3 MARK QUESTIONS

- What is moral realism? (3 marks)
- What does it say to say that an ethical theory is agent centred? (3 marks)
- What is moral naturalism? (3 marks)
- What does it mean to say an ethical theory is consequentialist? (3 marks)
- What is Bentham's utility calculus? (3 marks)
- What is preference utilitarianism? (3 marks)
- What is rule utilitarianism? (3 marks)
- What is eudaimonia? (3 marks)
- What is Aristotle's doctrine of the mean? (3 marks)
- What is moral non-naturalism? (3 marks)
- What is emotivism? (3 marks)
- What is error theory? (3 marks)
- What does it mean for a theory of moral language to be cognitivist? (3 marks)
- What is a non-cognitivist view of meta ethics? (3 marks)
- What is moral anti-realism? (3 marks)

5 MARK QUESTIONS

- Explain how an act utilitarian would make a moral decision (5 marks)
- Explain what error theory claims about the status of ethical language (5 marks)
- What is the difference between moral naturalism and moral non-naturalism? (5 marks)
- Explain the issue of circularity involved in Aristotle's definition of virtuous acts (5 marks)
- Explain the criticism that utilitarianism could lead to the tyranny of the majority (5 marks)

- Explain why emotivism is a non-cognitivist theory of ethical language. (5 marks)
- According to Hare's prescriptivism, what does "x is good" mean? (5 marks)
- Explain the analogy drawn between virtues and skills within Aristotelian ethics. (5 marks)
- Explain the different between non-cognitivism and non cognitivism. (5 marks)
- Explain why emotivism is a non-cognitivist theory of ethical language. (5 marks)
- Briefly explain what Moore means by the "naturalistic fallacy". (5 marks)
- Outline Hume's is/ought gap. (5 marks)
- Explain what Kant means by the good will. (5 marks)
- Briefly explain Mill's distinction between higher and lower pleasures. (5 marks)
- Explain the difference between act utilitarianism and rule utilitarianism. (5 marks)
- Explain what Nozick's experience machine is intended to show about act utilitarianism. (5 marks)
- What is a contradiction in conception? (5 marks)
- How is a hypothetical imperative different from a categorical imperative? (5 marks)
- Briefly outline Kant's second formulation of the categorical imperative. (5 marks)
- Briefly explain the criticism that Kant's deontological ethics ignores the value of certain motives. (5 marks)
- Explain how A.J. Ayer's verification principle presents a problem for moral realism. (5 marks)
- Explain why utilitarianism is a naturalist metaethical position. (5 marks)

12 MARK QUESTIONS

- Explain the similarities and differences between what emotivists and prescriptivists say about ethical language (12 marks)
- Explain Moore's open question argument (12 marks)
- Explain Aristotle's distinction between voluntary, involuntary, and non-voluntary action (12 marks)
- Explain how Kant's deontological ethics can be applied to the question of whether we should ever tell lies. (12 marks)

- How might a utilitarian attempt to justify eating animals? (12 marks)
- How might utilitarianism be applied to the question of whether it is morally acceptable to steal? (12 marks)
- How might Aristotle's virtue theory be applied to the issue of simulated killing? (12 marks)
- Outline how Mackie's arguments from queerness support error theory. (12 marks)
- Explain how act utilitarianism could lead to a tyranny of the majority and how rule utilitarianism might avoid this outcome. (12 marks)

25 MARK QUESTIONS

- How convincing is utilitarianism as an account of what makes an action morally right? (25 marks)
- Assess Kant's deontological ethics. (25 marks)
- Is Aristotle's Virtue Theory an accurate account of moral responsibility? (25 marks)
- What is the meta-ethical status of moral statements such as “murder is wrong”? (25 marks)

METAPHYSICS OF GOD

3 MARK QUESTIONS

- What does it mean to claim that God is everlasting? (3 marks)
- What does it mean to claim that God is eternal? (3 marks)
- What does Hare mean by the term "blik"? (3 marks)
- What is a theodicy? (3 marks)
- What does omnipotent mean? (3 marks)
- What does omniscient mean? (3 marks)
- What does omnibenevolent mean? (3 marks)
- What is an ontological argument? (3 marks)
- What is a cosmological argument? (3 marks)
- What is a teleological argument? (3 marks)
- What is Ayer's verification principle? (3 marks)
- What does it mean for a claim to be unfalsifiable? (3 marks)

5 MARK QUESTIONS

- Outline the paradox of the stone. (5 marks)
- Outline Descartes' version of the ontological argument. (5 marks)
- Outline the Kalam cosmological argument. (5 marks)
- Explain the evidential problem of evil. (5 marks)
- Outline Aquinas' Third Way. (5 marks)
- Outline Paley's watch analogy. (5 marks)
- Outline Mackie's inconsistent triad argument. (5 marks)
- Explain the difference between the claims "God is eternal" and "God is everlasting". (5 marks)
- What is the logical problem of evil? (5 marks)

- Briefly explain the difference between cognitivism and non cognitivism about religious language. (5 marks)
- Explain the difference between natural evil and moral evil. (5 marks)
- Briefly explain Swinburne's teleological argument. (5 marks)
- Briefly explain Leibniz's cosmological argument from the principle of sufficient reason. (5 marks)

12 MARK QUESTIONS

- Explain the free will response to the problem of evil. (12 marks)
- Outline the verification principle and explain Hick's claim that religious statements are verifiable eschatologically. (12 marks)
- Explain Mitchell's view of religious language and how he uses his "Partisan" story to illustrate this. (12 marks)
- Outline the problem of evil and explain Hick's soul-making response to it. (12 marks)
- Briefly define "omniscience" and then explain the argument that free will is impossible if God is omniscient. (12 marks)
- Compare and contrast Paley's and Swinburne's versions of the design argument. (12 marks)
- Compare and contrast cosmological arguments from causation with cosmological arguments from contingency. (12 marks)
- Explain how Gaunilo's island provides a response to Anselm's ontological argument. (12 marks)
- Outline Flew's jungle clearing analogy and explain what this is intended to show about religious language. (12 marks)
- Explain how the Euthyphro dilemma presents a challenge to the concept of an omnipotent, omniscient, and omnibenevolent God. (12 marks)
- Compare and contrast St Anselm's ontological argument with Norman Malcolm's ontological argument. (12 marks)
- Explain how Kant's claim that existence is not a predicate presents a problem for ontological arguments. (12 marks)
- Explain how Hume's objection to the causal principle presents a problem for cosmological arguments from causation. (12 marks)

- Explain how the impossibility of an infinite series supports cosmological arguments. (12 marks)
- Explain how Hare's "bliks" can be used to respond to Flew's jungle clearing analogy. (12 marks)

25 MARK QUESTIONS

- Is the concept of God coherent? (25 marks)
- Does the order of the universe prove that God exists? (25 marks)
- Do cosmological arguments succeed in proving the existence of God? (25 marks)
- Do ontological arguments prove God exists? (25 marks)
- Does the existence of evil mean that God does not exist? (25 marks)
- Is religious language meaningful? (25 marks)

METAPHYSICS OF MIND

3 MARK QUESTIONS

- What claim do logical/analytical behaviourists make regarding statements about mental states? (3 marks)
- What is a philosophical zombie? (3 marks)
- What do eliminative materialists claim about mental states? (3 marks)
- What is property dualism? (3 marks)
- What is substance dualism? (3 marks)
- What is type identity theory? (3 marks)
- What do functionalists believe mental states are? (3 marks)
- What is a super spartan? (3 marks)
- What are qualia? (3 marks)
- What is folk psychology? (3 marks)

5 MARK QUESTIONS

- Explain the argument that it would self-refuting to articulate eliminative materialism as theory. (5 marks)
- Outline Jackson's knowledge argument (Mary the super scientist). (5 marks)
- Explain how eliminative materialism differs from mind-brain type identity theory. (5 marks)
- Outline Descartes' conceivability argument for substance dualism. (5 marks)
- Explain how Block's China brain thought experiment can be used to argue against functionalism. (5 marks)
- Explain how substance dualism differs from property dualism. (5 marks)
- Outline the philosophical zombies argument for property dualism. (5 marks)
- Explain how super Spartans could be used to argue against behaviourism. (5 marks)

- What is the difference between a philosophical zombie and a super spartan? (5 marks)
- Outline Descartes' divisibility argument. (5 marks)
- Outline the objection that dualism makes a category mistake. (5 marks)
- Explain how the inverted spectrum presents a problem for functionalism. (5 marks)
- Explain how the asymmetry between self- and other- knowledge presents a problem for behaviourism. (5 marks)
- Briefly explain how behaviourism could be criticised for using circular reasoning. (5 marks)
- What does it mean for a mental state to be 'multiply realisable'? (5 marks)
- What is the difference between soft and hard behaviourism? (5 marks)
- Explain Gilbert Ryle's category mistake objection to dualism. (5 marks)

12 MARK QUESTIONS

- Explain the similarities and differences between functionalism and mind-brain type identity theory. (12 marks)
- Explain how the asymmetry between self-knowledge and knowledge of other people's mental states might cause an issue for logical/analytic behaviourism. (12 marks)
- Outline mind-brain type identity theory and explain how the issue of multiple realisability challenges this view. (12 marks)
- Outline substance dualism and explain how the problem of causal interaction challenges this view. (12 marks)
- Briefly outline the knowledge argument and explain how this could be used to argue against functionalism. (12 marks)
- Outline eliminative materialism and explain how the predictive power of folk psychology challenges this view. (12 marks)
- Outline the ability hypothesis and the acquaintance hypothesis responses to the knowledge argument. (12 marks)
- Explain Descartes' conceivability argument and the criticism that it relies on the masked man fallacy. (12 marks)

- Outline mind-brain type identity theory and explain how Descartes' divisibility argument presents a challenge to this view. (12 marks)

25 MARK QUESTIONS

- Assess substance dualism. (25 marks)
- "Some physical substances can have non-physical properties." Discuss. (25 marks)
- "Mental states are nothing more than behavioural dispositions." Discuss. (25 marks)
- Are mental states identical to brain states? (25 marks)
- Assess the claim that mental states are functional states. (25 marks)
- Can folk psychological accounts of mental states be eliminated? (25 marks)



REVISION EXERCISES

This section contains multiple choice quizzes to test your knowledge, tables to help organise the various theories and the arguments for and against them, as well as a last minute revision checklist.

QUIZZES

EPISTEMOLOGY

1. Which of the following is <u>not</u> a condition of the reliabilist definition of knowledge?

A) True
B) Belief
C) Justified
D) Formed via a reliable method

2. Which of the following do rationalists claim?

A) There is no synthetic a posteriori knowledge
B) There is some synthetic a priori knowledge
C) There is no synthetic a priori knowledge
D) All knowledge is a priori

3. Which of the following is an analytic truth?

A) A bachelor is an unmarried man
B) Grass is green
C) A triangle is a 5-sided shape
D) Water boils at 100°c

4. Which of the following is an objection to indirect realism?

A) Illusion
B) Perceptual variation
C) Hallucination
D) Scepticism

5. Which of the following philosophers was an empiricist?

A) Leibniz
B) Locke
C) Plato
D) Descartes

6. Which definition of knowledge can allow that animals possess knowledge?

A) Reliabilism
B) Virtue epistemology
C) Infallibilism
D) No false lemmas

7. According to Hume's Fork, which are the two judgements of reason?

A) Analytic and synthetic
B) A priori knowledge and a posteriori knowledge
C) Simple and complex ideas
D) Relations of ideas and matters of fact

8. What is the conclusion of Descartes' trademark argument?

A) I exist
B) God exists
C) The external world exists
D) Cogito ergo sum

9. According to idealism, what are the immediate objects of perception?

A) Ideas
B) Sense data
C) Mind-independent objects
D) The external world

10. What does a Gettier case show?

A) Justification is not a necessary condition of knowledge
B) The tripartite definition of knowledge is not necessary
C) The tripartite definition of knowledge is not sufficient
D) Virtue epistemology is the correct definition of knowledge

11. Which of the following is <u>not</u> an example of global scepticism?

A) Being unsure of your memory of an event
B) Brain in a vat
C) Descartes' evil demon
D) The simulation hypothesis

12. Which of the following does Leibniz claim is an example of innate knowledge?

A) Contingent truths
B) Necessary truths
C) Analytic truths
D) Synthetic truths

13. How does Russell respond to the challenge of global scepticism?

A) The involuntary nature of experience means the external world exists
B) The external world is the best hypothesis
C) Reliabilism
D) God wouldn't create him with a tendency to constantly form false beliefs

14. According to Locke, which of the following is a secondary quality?

A) Size
B) Shape
C) Motion
D) Colour

15. Which of the following is <u>not</u> one of Descartes' three waves of doubt?

A) Illusion
B) Hallucination
C) Dreaming
D) Deception via evil demon

Answers

1C, 2B, 3A, 4D, 5B, 6A, 7D, 8B, 9A, 10C, 11A, 12B, 13B, 14D, 15B

MORAL PHILOSOPHY

1. Which of the following is a hypothetical imperative?

A) Don't steal
B) Act only according to that maxim you can will become a universal law
C) The humanity formula
D) Leave now if you want to get to the station on time

2. Which of the following theories is consequentialist?

A) Virtue ethics
B) Kant's deontological ethics
C) Utilitarianism
D) Emotivism

3. The experience machine is an argument that supports which form of utilitarianism?

A) Act utilitarianism
B) Rule utilitarianism
C) Preference utilitarianism
D) Quantitative hedonism

4. Which of the following metaethical theories is both cognitivist and anti-realist?

A) Non-naturalism
B) Emotivism
C) Error theory
D) Prescriptivism

5. Which of the following is an objection to Aristotle's virtue ethics?

A) Ignores consequences
B) Tyranny of the majority
C) Circularity
D) Argument from queerness

6. Moral argument and reasoning is an argument for which metaethical position?

A) Emotivism
B) Prescriptivism
C) Cognitivism
D) Non-cognitivism

7. According to prescriptivism, "stealing is wrong" means...

A) Stealing has the property of wrongness
B) "Boo! Stealing!"
C) "Don't steal!"
D) "I don't like stealing"

8. Which of the following is the best translation of "eudaimonia"?

A) Happiness
B) Someone who acts morally
C) Human flourishing
D) The function of human beings

9. Which philosopher created the felicific calculus as a way of quantifying happiness?

A) Mill
B) Nozick
C) Bentham
D) Singer

10. Why does Kant say we have a perfect duty never to tell lies?

A) It makes the person unhappy when they find out they've been lied to
B) Telling lies leads to a contradiction in will when universalised
C) Telling lies leads to a contradiction in conception when universalised
D) Telling lies will lead you to develop a bad character

11. According to Aristotle, what is the ergon of human beings?

A) To be moral
B) To develop virtuous character
C) To use reason
D) To act virtuously

12. Which of the following is an objection to Mill's proof of utilitarianism?

A) Tyranny of the majority
B) Rule utilitarianism
C) The naturalistic fallacy
D) Difficult to calculate

13. According to Kant, what is the good will?

A) To act for the sake of maximising happiness
B) To act in accordance with duty
C) To act for the sake of duty
D) To act virtuously

14. Which metaethical theory did G.E. Moore argue for?

A) Ethical naturalism
B) Ethical non-naturalism
C) Error theory
D) Emotivism

15. A maxim leads to a contradiction in conception if...

A) It would be self-contradictory for everyone to follow it
B) You cannot rationally will that everyone follow it
C) It treats someone as a means to an end
D) It contradicts another maxim you have agreed to follow

Answers

1D, 2C, 3C, 4C, 5C, 6C, 7C, 8C, 9C, 10C, 11C, 12C, 13C, 14B, 15A

METAPHYSICS OF GOD

1. Which of the following argues that "God exists" is an analytic truth?

A) Ontological argument
B) Teleological argument
C) Cosmological argument
D) The problem of evil

2. Which of the following is a cosmological argument from contingency?

A) Aquinas' first way
B) Aquinas' second way
C) Aquinas' third way
D) The kalam argument

3. Gaunilo's island is a response to which argument for the existence of God?

A) Ontological argument
B) Teleological argument
C) Cosmological argument
D) The problem of evil

4. Spatial disorder is a response to which argument for the existence of God?

A) Ontological argument
B) Teleological argument
C) Cosmological argument
D) The problem of evil

5. Which of the following is <u>not</u> one of the divine attributes?

A) Omnipotent
B) Omnibenevolent
C) Omniscient
D) Omnivore

6. Which of the following philosophers argues that "God exists" is meaningless?

A) Flew
B) Hick
C) Mitchell
D) Hare

7. Which philosopher makes a teleological argument from temporal order?

A) Hume
B) Paley
C) Aquinas
D) Swinburne

8. In his ontological argument, how does Anselm define God?

A) A supremely perfect being
B) A being greater than which cannot be conceived
C) A being that does not lack any perfections
D) A being that exists necessarily

9. According to Paley, what is the hallmark of design?

A) The fitting of means to ends
B) Something with many parts organised for a purpose
C) Something that cannot occur naturally
D) Temporal order

10. Which of the following is <u>not</u> part of the inconsistent triad?

A) An omnipotent God exists
B) An omniscient God exists
C) An omnibenevolent God exists
D) Evil exists

11. What does "God is eternal" mean?

A) God has existed from the beginning of time and will exist forever
B) God cannot come into existence or go out of existence
C) God exists outside of time
D) God cannot be destroyed

12. Which of the following is an example of moral evil?

A) An earthquake kills hundreds of people
B) A tornado destroys an art gallery
C) A virus kills hundreds of people
D) A person kills thousands of people

13. Which question best captures the Euthyphro dilemma?

A) Can God create a stone so heavy He can't lift it?
B) Does God create moral laws or does God just follow moral laws?
C) How can humans have free will if God is omniscient?
D) Why does an omnipotent and omnibenevolent God allow evil?

14. What is a theodicy?

A) A meaningless claim
B) A statement that cannot be proven false
C) An argument for the existence of God
D) An explanation of why God allows evil

15. Which philosopher describes religious beliefs as bliks?

A) Hare
B) Hick
C) Mitchell
D) Ayer

Answers

1A, 2C, 3A, 4B, 5D, 6A, 7D, 8B, 9B, 10B, 11C, 12D, 13B, 14D, 15A

METAPHYSICS OF MIND

1. Which of the following is an example of qualia?

A) The colour red
B) Wavelengths of light
C) The redness of a ripe tomato
D) The redness of my experience of a ripe tomato

2. The category mistake argument supports which theory of the mind?

A) Property dualism
B) Eliminative materialism
C) Type identity theory
D) Behaviourism

3. What is a philosophical zombie?

A) A physically identical human but with no qualia
B) A human that has qualia but with no outward behaviour
C) A being similar to a human but with a different brain structure
D) A human with different qualia to normal humans

4. The location problem is an argument against which theory of the mind?

A) Substance dualism
B) Eliminative materialism
C) Type identity theory
D) Functionalism

5. Which of the following is <u>not</u> an ontological reduction?

A) Lightning is electrical discharge
B) Water is H2O
C) A bachelor is an unmarried man
D) Pain is c-fibers firing

6. The China brain thought experiment is an argument against which theory of the mind?

A) Functionalism
B) Eliminative materialism
C) Type identity theory
D) Behaviourism

7. The knowledge argument (Mary) is an argument for which theory of the mind?

A) Substance dualism
B) Property dualism
C) Type identity theory
D) Behaviourism

8. Which of the following does <u>not</u> have intentional content?

A) A belief about unicorns
B) A thought about a frog
C) A desire to drink beer
D) C-fibers firing

9. Which of the following presents a challenge to epiphenomenal dualism?

A) Evolution
B) The knowledge argument
C) Circularity
D) Multiple realisability

10. Which of the following best describes "folk psychology"?

A) The scientific study of the mind
B) A study of the human brain
C) The common sense way we talk about the mind
D) The reduction of mental states to brain states

11. Which of the following best characterises the multiple realisability objection to type identity theory?

A) Octopuses can't feel pain
B) The same physical state can realise different mental states
C) Different physical states can realise the same mental state
D) Two people can have the same mental state but different qualia

12. What is a super Spartan ?

A) A being that does not feel pain
B) A possible being that is like an ordinary human but stronger
C) A being that does not feel pain but behaves as though it does
D) A being that feels pain but does not show it

13. What is the inverted spectrum?

A) Where people disagree what colour something is
B) Where people agree what colour something is but have different qualia
C) Where people disagree what colour something is and have different qualia
D) The view that everyone sees colour differently

14. Epiphenomenal dualism claims that...

A) The mental can cause changes to the physical but not vice versa
B) The physical can cause changes to the mental but not vice versa
C) The mental and physical both interact with each other
D) The mental and physical do not interact at all

15. The masked man fallacy is a response to which argument?

A) Descartes' conceivability argument
B) Descartes' divisibility argument
C) The knowledge argument
D) The inverted spectrum

Answers

1D, 2D, 3A, 4C, 5C, 6A, 7B, 8D, 9A, 10C, 11C, 12D, 13B, 14B, 15A

TABLES

The following tables break the key points of each major topic into a simple arguments for and arguments against format. Filling these out will help organise this information and is a good way to test your memory of the topics and arguments.

EPISTEMOLOGY

	Definition	Arguments for	Arguments against
JTB (tripartite definition)			
JTB + No false lemmas			
Reliabilism			
Virtue epistemology			
Infallibilism			

MORAL PHILOSOPHY

	What the theory is	Arguments for	Arguments against
Utilitarianism			
Kant's deontological ethics			
Virtue ethics			

METAPHYSICS OF GOD

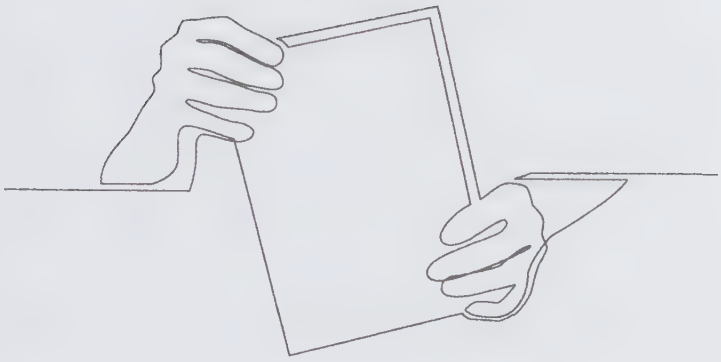
	Versions	Responses/arguments against
Ontological arguments		
Teleological arguments		
Cosmological arguments		
The problem of evil		

METAPHYSICS OF MIND

	What the theory is	Arguments for	Arguments against
Substance dualism			
Property dualism			
Behaviourism			
Type identity theory			
Functionalism			
Eliminative materialism			

LAST MINUTE REVISION CHECKLIST

- Read through the module summaries:
  - Epistemology (page 79)
  - Moral philosophy (page 138)
  - Metaphysics of God (page 191)
  - Metaphysics of mind (page 248)
- Look over the course syllabus (page 327), make a list of topics you're still unsure about, and then read through the more detailed notes on those topics
- Read the exam technique summary (page 290) and understand the difference between AO1 (3, 5, and 12 mark questions) and AO2 (25 mark questions)
- Familiarise yourself with the 25 mark question response template (page 260)
- Plan answers for each of the example 25 mark questions (page 291) according to the template above. You can find example templates for specific questions here:
  - Epistemology (page 267)
  - Moral philosophy (page 274)
  - Metaphysics of God (page 278)
  - Metaphysics of mind (page 284)



SYLLABUS

Below is the AQA syllabus for A-level philosophy for exams from 2019 onwards (course code 7172).

EPISTEMOLOGY

What is knowledge?

- The distinction between acquaintance knowledge, ability knowledge and propositional knowledge.
- The nature of definition (including Linda Zagzebski) and how propositional knowledge may be analysed/defined.

The tripartite view

Propositional knowledge is defined as justified true belief: S knows that p if and only if:

1. S is justified in believing that p,
2. p is true and
3. S believes that p (individually necessary and jointly sufficient conditions)

Issues with the tripartite view including:

- the conditions are not individually necessary
- the conditions are not sufficient – cases of lucky true beliefs (including Edmund Gettier's original two counter examples):
    - responses: alternative post-Gettier analyses/definitions of knowledge including:
    - strengthen the justification condition (ie infallibilism)
    - add a 'no false lemmas' condition (J+T+B+N)
    - replace 'justified' with 'reliably formed' (R+T+B) (ie reliabilism)
    - replace 'justified' with an account of epistemic virtue (V+T+B).

Perception as a source of knowledge

Direct realism

The immediate objects of perception are mind-independent objects and their properties

Issues including:

- the argument from illusion
- the argument from perceptual variation
- the argument from hallucination
- the time-lag argument

and responses to these issues.

Indirect realism

The immediate objects of perception are mind-dependent objects (sense-data) that are caused by and represent mind-independent objects.

- John Locke's primary/secondary quality distinction.

Issues including:

- the argument that it leads to scepticism about the existence of mind-independent objects. Responses including:
  - Locke's argument from the involuntary nature of our experience
  - the argument from the coherence of various kinds of experience, as developed by Locke and Catharine Trotter Cockburn (attrib)
  - Bertrand Russell's response that the external world is the 'best hypothesis'
- the argument from George Berkeley that we cannot know the nature of mind-independent objects because mind-dependent ideas cannot be like mind-independent objects.

Berkeley's Idealism

The immediate objects of perception (ie ordinary objects such as tables, chairs, etc) are mind dependent objects.

- Arguments for idealism including Berkeley's attack on the primary/secondary quality distinction and his 'Master' argument.

Issues including:

- arguments from illusion and hallucination
- idealism leads to solipsism
- problems with the role played by God in Berkeley's Idealism (including how can Berkeley claim that our ideas exist within God's mind given that he believes that God cannot feel pain or have sensations?)

and responses to these issues.

Reason as a source of knowledge

Innatism

Arguments from Plato (ie the 'slave boy' argument) and Gottfried Leibniz (ie his argument based on necessary truths).

Empiricist responses including:

- Locke's arguments against innatism

- the mind as a 'tabula rasa' (the nature of impressions and ideas, simple and complex concepts)

and issues with these responses.

The intuition and deduction thesis

- The meaning of 'intuition' and 'deduction' and the distinction between them.
- René Descartes' notion of 'clear and distinct ideas'.
- His cogito as an example of an a priori intuition.
- His arguments for the existence of God and his proof of the external world as examples of a priori deductions.

Empiricist responses including:

- responses to Descartes' cogito
- responses to Descartes' arguments for the existence of God and his proof of the external world (including how Hume's Fork might be applied to these arguments)

and issues with these responses.

The limits of knowledge

- Particular nature of philosophical scepticism and the distinction between philosophical scepticism and normal incredulity.
- The role/function of philosophical scepticism within epistemology
- The distinction between local and global scepticism and the (possible) global application of philosophical scepticism
- Descartes' sceptical arguments (the three 'waves of doubt')
- Responses to scepticism: the application of the following as responses to the challenge of scepticism:
    - Descartes' own response
    - empiricist responses (Locke, Berkeley and Russell)
    - reliabilism.

MORAL PHILOSOPHY

Normative ethical theories

- The meaning of good, bad, right, wrong within each of the three approaches specified below
- Similarities and differences across the three approaches specified below

Utilitarianism

- The question of what is meant by 'utility' and 'maximising utility', including:
- Jeremy Bentham's quantitative hedonistic utilitarianism (his utility calculus)
- John Stuart Mill's qualitative hedonistic utilitarianism (higher and lower pleasures) and his 'proof' of the greatest happiness principle
- non-hedonistic utilitarianism (including preference utilitarianism)
- act utilitarianism and rule utilitarianism.

Issues, including:

- whether pleasure is the only good (Nozick's experience machine)
- fairness and individual liberty/rights (including the risk of the 'tyranny of the majority')
- problems with calculation (including which beings to include)
- issues around partiality
- whether utilitarianism ignores both the moral integrity and the intentions of the individual.

Kantian deontological ethics

- Immanuel Kant's account of what is meant by a 'good will'.
- The distinction between acting in accordance with duty and acting out of duty.
- The distinction between hypothetical imperatives and categorical imperatives.
- The first formulation of the categorical imperative (including the distinction between a contradiction in conception and a contradiction in will).
- The second formulation of the categorical imperative.

Issues, including:

- clashing/competing duties
- not all universalisable maxims are distinctly moral; not all non-universalisable maxims are immoral

- the view that consequences of actions determine their moral value
- Kant ignores the value of certain motives, eg love, friendship, kindness
- morality is a system of hypothetical, rather than categorical, imperatives (Philippa Foot).

Aristotelian virtue ethics

- 'The good' for human beings: the meaning of Eudaimonia as the 'final end' and the relationship between Eudaimonia and pleasure.
- The function argument and the relationship between virtues and function.
- Aristotle's account of virtues and vices: virtues as character traits/dispositions; the role of education/habituation in the development of a moral character; the skill analogy; the importance of feelings; the doctrine of the mean and its application to particular virtues.
- Moral responsibility: voluntary, involuntary and non-voluntary actions.
- The relationship between virtues, actions and reasons and the role of practical reasoning/ practical wisdom.

Issues including:

- whether Aristotelian virtue ethics can give sufficiently clear guidance about how to act
- clashing/competing virtues
- the possibility of circularity involved in defining virtuous acts and virtuous persons in terms of each other
- whether a trait must contribute to Eudaimonia in order to be a virtue; the relationship between
- the good for the individual and moral good.

Applied ethics

Students must be able to apply the content of Normative ethical theories and metaethics to the following issues:

- stealing
- simulated killing (within computer games, plays, films etc)
- eating animals
- telling lies.

Meta-ethics

The origins of moral principles: reason, emotion/attitudes, or society.

The distinction between cognitivism and non-cognitivism about ethical language.

Moral realism

There are mind-independent moral properties/facts.

- Moral naturalism (cognitivist) – including naturalist forms of utilitarianism (including Bentham) and of virtue ethics.
- Moral non-naturalism (cognitivist) – including intuitionism and Moore's 'open question argument' against all reductive metaethical theories and the Naturalistic Fallacy.

Issues that may arise for the theories above, including:

- Hume's Fork and A J Ayer's verification principle
- Hume's argument that moral judgements are not beliefs since beliefs alone could not motivate us
- Hume's is-ought gap
- John Mackie's argument from relativity and his arguments from queerness.

Moral anti-realism

There are no mind-independent moral properties/facts.

- Error Theory (cognitivist) - Mackie
- Emotivism (non-cognitivist) – Ayer
- Prescriptivism (non-cognitivist) – Richard Hare

Issues that may arise for the theories above, including:

- whether anti-realism can account for how we use moral language, including moral reasoning, persuading, disagreeing etc.
- the problem of accounting for moral progress
- whether anti-realism becomes moral nihilism.

METAPHYSICS OF GOD

The concept and nature of 'God'

God's attributes:

- God as omniscient, omnipotent, supremely good (omnibenevolent), and the meaning(s) of these divine attributes
- competing views on such a being's relationship to time, including God being timeless (eternal) and God being within time (everlasting).
- arguments for the incoherence of the concept of God including:
  - the paradox of the stone
  - the Euthyphro dilemma
  - the compatibility, or otherwise, of the existence of an omniscient God and free human beings.

Arguments relating to the existence of God

For the arguments below, students should pay particular attention to nuances in the logical form of the arguments (deductive, inductive etc), the strengths of the conclusions (God does exist, God must exist etc) and the nature of God assumed or defended by the argument.

Ontological arguments

- St Anselm's ontological argument.
- Descartes' ontological argument.
- Norman Malcolm's ontological argument.

Issues that may arise for the arguments above, including:

- Gaunilo's 'perfect island' objection
- Empiricist objections to a priori arguments for existence
- Kant's objection based on existence not being a predicate.

Teleological/design arguments

- The design argument from analogy (as presented by Hume).
- William Paley's design argument: argument from spatial order/purpose.
- Richard Swinburne's design argument: argument from temporal order/regularity.

Issues that may arise for the arguments above, including:

- Hume's objections to the design argument from analogy
- the problem of spatial disorder (as posed by Hume and Paley)
- the design argument fails as it is an argument from a unique case (Hume)
- whether God is the best or only explanation.

Cosmological arguments

- The Kalām argument (an argument from temporal causation).
- Aquinas' 1st Way (argument from motion), 2nd Way (argument from atemporal causation) and 3rd way (an argument from contingency).
- Descartes' argument based on his continuing existence (an argument from causation).
- Leibniz's argument from the principle of sufficient reason (an argument from contingency).

Issues that may arise for the arguments above, including:

- the possibility of an infinite series
- Hume's objection to the 'causal principle'
- the argument commits the fallacy of composition (Russell)
- the impossibility of a necessary being (Hume and Russell).

The Problem of Evil

Whether God's attributes can be reconciled with the existence of evil.

- The nature of moral evil and natural evil.
- The logical and evidential forms of the problem of evil.

Responses to these issues and issues arising from these responses, including:

- the Free Will Defence (including Alvin Plantinga)
- soul-making (including John Hick).

Religious language

- The distinction between cognitivism and non-cognitivism about religious language.
- The empiricist/logical positivist challenges to the status of metaphysical (here, religious) language: the verification principle and verification/falsification (Ayer).
    - Hick's response to Ayer (eschatological verification) and issues arising from that response.
- Further responses: the 'University Debate'
    - Anthony Flew on falsification (Wisdom's 'Gardener')
    - Basil Mitchell's response to Flew (the Partisan)
    - Hare's response to Flew (bliks and the lunatic)

and issues arising from those responses.

METAPHYSICS OF MIND

What do we mean by 'mind'?

Features of mental states:

- All or at least some mental states have phenomenal properties
- Some, but not all, philosophers use the term 'qualia' to refer to these properties, where 'qualia' are defined as 'intrinsic and non-intentional phenomenal properties that are introspectively accessible'
- All or at least some mental states have intentional properties (ie intentionality).

Dualist theories

Substance dualism

Minds exist and are not identical to bodies or to parts of bodies.

The indivisibility argument for substance dualism (Descartes).

Responses, including:

- the mental is divisible in some sense
- not everything thought of as physical is divisible.

The conceivability argument for substance dualism (expressed without reference to God) (Descartes).

Responses including:

- mind without body is not conceivable
- what is conceivable may not be metaphysically possible
- what is metaphysically possible tells us nothing about the actual world.

Property dualism

There are at least some mental properties that are neither reducible to nor supervenient upon physical properties.

The 'philosophical zombies' argument for property dualism (David Chalmers).

Responses including:

- a 'philosophical zombie'/a 'zombie' world is not conceivable
- what is conceivable may not be metaphysically possible
- what is metaphysically possible tells us nothing about the actual world.

The 'knowledge/Mary' argument for property dualism (Frank Jackson).

Responses including:

- Mary does not gain new propositional knowledge but does gain ability knowledge (the 'ability knowledge' response).
- Mary does not gain new propositional knowledge but does gain acquaintance knowledge (the 'acquaintance knowledge' response).
- Mary gains new propositional knowledge, but this is knowledge of physical facts that she already knew in a different way (the 'New Knowledge / Old Fact' response).

Issues

Issues facing dualism, including:

- The problem of other minds
- Responses including:
  - the argument from analogy
  - the existence of other minds is the best hypothesis.
- Dualism makes a "category mistake" (Gilbert Ryle)

Issues facing interactionist dualism, including:

- the conceptual interaction problem (as articulated by Elisabeth, Princess of Bohemia)
- the empirical interaction problem.
- Issues facing epiphenomenalist dualism, including:
  - the challenge posed by introspective self-knowledge
  - the challenge posed by the phenomenology of our mental life (ie as involving causal connections, both psychological and psycho-physical)
  - the challenge posed by natural selection/evolution.

Physicalist theories

Physicalism

Everything is physical or supervenes upon the physical (this includes properties, events, objects and any substance(s) that exist).

Philosophical behaviourism:

- 'Hard' behaviourism: all propositions about mental states can be reduced without loss of meaning to propositions that exclusively use the language of physics to talk about bodily states/movements (including Carl Hempel).
- 'Soft' behaviourism: propositions about mental states are propositions about behavioural dispositions (ie propositions that use ordinary language) (including Gilbert Ryle).

Issues including:

- dualist arguments applied to philosophical behaviourism
- the distinctness of mental states from behaviour (including Hilary Putnam's 'Super-Spartans' and perfect actors)
- issues defining mental states satisfactorily due to (a) circularity and (b) the multiple realisability of mental states in behaviour
- the asymmetry between self-knowledge and knowledge of other people's mental states.

Mind-brain type identity theory

All mental states are identical to brain states ('ontological' reduction) although 'mental state' and 'brain state' are not synonymous (so not an 'analytic' reduction).

Issues including:

- dualist arguments applied to mind-brain type identity theory
- issues with providing the type identities (the multiple realisability of mental states).

Eliminative materialism

Some or all common-sense ("folk-psychological") mental states/properties do not exist and our common-sense understanding is radically mistaken (as defended by Patricia Churchland and Paul Churchland).

Issues including:

- our certainty about the existence of our mental states takes priority over other considerations

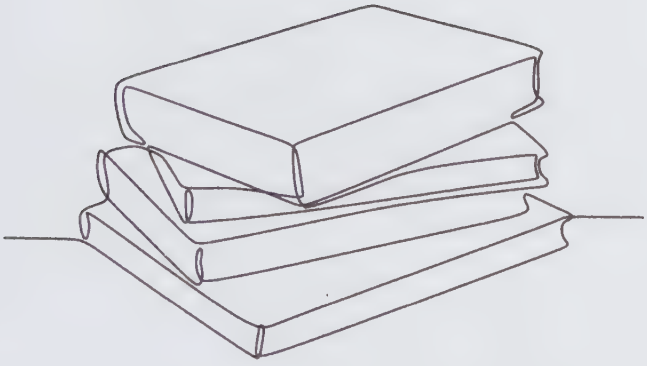
- folk-psychology has good predictive and explanatory power (and so is the best hypothesis)
- the articulation of eliminative materialism as a theory is self-refuting.

Functionalism

Functionalism: all mental states can be characterised in terms of functional roles which can be multiply realised.

Issues, including:

- the possibility of a functional duplicate with different qualia (inverted qualia)
- the possibility of a functional duplicate with no mentality/qualia (Ned Block's China thought experiment)
- the 'knowledge'/Mary argument can be applied to functional facts (no amount of facts about function suffices to explain qualia).



BIBLIOGRAPHY/READING LIST

Below is the reading list in the AQA syllabus for A-level philosophy for exams from 2019 onwards (course code 7172). The course notes in earlier sections of this book are based on these texts. You can find links to the full texts at *philosophyalevel.com/useful-links/*

EPISTEMOLOGY

Berkeley, George (1713), Three Dialogues Between Hylas and Philonous

Descartes, René (1641), Meditations on First Philosophy, 1, 2, 3, 5, 6

Gettier, Edmund (1963), 'Is Justified True Belief Knowledge?' Analysis, 23(6): 121–123

Hume, David (1748), An Enquiry Concerning Human Understanding, Section 2 and Section 4 (part 1)

Leibniz, Gottfried Wilhelm (1705), New Essays on Human Understanding, Preface and Book 1

Locke, John (1690), An Essay Concerning Human Understanding, Book 1 (esp. Chapter 2), Book 2 (esp. Chapters 1, 2, 8 and 14), Book 4 (esp. Chapter 11)

Plato, Meno, from 81e

Russell, Bertrand (1912), The Problems of Philosophy, Chapters 1, 2

Trotter Cockburn, Catharine (1732), (attrib) 'A Letter from an anonymous writer to the author of the Minute Philosopher' Appendix to G Berkeley Theory of Vision Vindicated and Explained

Zagzebski, Linda (1999), 'What is Knowledge?' in John Greco & Ernest Sosa (eds.), The Blackwell Guide to Epistemology 92 –116

MORAL PHILOSOPHY

Annas, Julia (2006), 'Virtue Ethics', in David Copp (ed.), The Oxford Handbook of Ethical Theory, Oxford: Oxford University Press, pp. 515–536

Aristotle, Nicomachean Ethics: Books 1 (1–5, 7–10, 13), 2 (1–7), 3 (1–5), 5, 6 (1, 2, 5, 8, 9, 12, 13), 7 (12–13), 10 (1–8)

Ayer, Alfred J (1973/1991), The Central Questions of Philosophy, London, Penguin, 22–29 and

Ayer, AJ (1946), Language, Truth and Logic, 2nd Edition, New York, Dover, (esp. Chapters 1 and 6)

Bentham, Jeremy (1789), Introduction to the Principles of Morals and Legislation, Oxford, Clarendon Press. Chapter 1 (The Principle of Utility) and Chapter 4 (Measuring Pleasure and Pain)

Diamond, Cora (1978), 'Eating Meat and Eating People' Philosophy 53: 465–479

Foot, Philippa (1972), 'Morality as a system of hypothetical imperatives.' Philosophical Review, vol 81, issue 3, 305–316.

Hare, Richard M (1952) The Language of Morals, Chapters 1, 5, 7, 10.2

Hume, David (1739–40), Treatise of Human Nature, Book III, Part 1

Kant, Immanuel (1785), Foundations of the Metaphysics of Morals, Chapters 1 and 2

Mackie, John L (1977), Ethics: Inventing Right and Wrong, Penguin, Chapter 1, Sections 8 and 9

Mill, John Stuart (1863), Utilitarianism, Chapters 1, 2, 4, 5

Moore, George E (1903), Principia Ethica, Cambridge University Press, Sections 6–14

Smart, Jack J C & Williams, Bernard (1973), Utilitarianism: For and Against, Chapter 2 (Act utilitarianism and rule utilitarianism) and Chapter 3 (Hedonistic and non-hedonistic utilitarianism)

METAPHYSICS OF GOD

Anselm, Proslogium, Chapters II–IV and Gaunilo, from the appendix to Anselm's Proslogium

Aquinas, Thomas, Summa Theologica, Part 1, Question 25, Article 3 and Question 2, Article 3

Ayer, Alfred J (1973/1991), The Central Questions of Philosophy, London, Penguin, 22–29 and

Ayer, AJ (1946), Language, Truth and Logic, 2nd Edition, New York, Dover, (Chapters 1 and 6)

Descartes, René (1641), Meditations on First Philosophy, 3 and 5

Flew, Antony, Richard M Hare and Basil Mitchell (1955), 'Theology and Falsification' in New Essays in Philosophical Theology, edited by Antony Flew and Alasdair MacIntyre, London, SMC Press

Hick, John (1966/1978), Evil and the God of Love, New York, Harper and Row (revised edition). Chapters 13–17 (esp chapter 13)

Hume, David (1779), Dialogues Concerning Natural Religion, Parts II, V, VIII and IX

Hume, David (1748), Enquiry Concerning Human Understanding, Part XI

Leibniz, Gottfried (1714), Monadology, sections 32–39.

Malcolm, Norman (1960), 'Anselm's ontological arguments', The Philosophical Review, 69, 41–62

Midgley, Mary (1984), Wickedness. Routledge, Chapters 1 and 5

Paley, William (1802/2008), Natural Theology, OUP, Chapters 1, 2 and 5

Plantinga, Alvin (1975), God, Freedom and Evil: Essays in Philosophy, George Allen & Unwin, 29–34 and 59–64

Plato, Euthyphro

Stump, Eleanore & Kretzmann, Norman (1981), Eternity. Journal of Philosophy 78 (8):429–458

Swinburne, Richard G (1968), 'The Argument from Design,' Philosophy, 43 (165), 199–212

METAPHYSICS OF MIND

Avramides Anita (2001). Other Minds. Routledge, Chapter 2.

Block, Ned (1980) 'Troubles with functionalism,' in Readings in Philosophy of Psychology, Vol 1. Harvard University Press, 275–278 – section 1–2.

Chalmers, David (1996), The Conscious Mind: In Search of a Fundamental Theory. Oxford University Press, Chapter 1.

Chalmers, David (2003). 'Consciousness and its place in nature' in Blackwell Guide to the Philosophy of Mind, Sections 2 and 3. Blackwell

Churchland, Patricia Smith (2002). Brain-wise. Introduction (Sections 1, 3, 4) and Chapter 4, section 2. The MIT Press.

Churchland, Paul (1981), 'Eliminative Materialism and Propositional Attitudes', Journal of Philosophy 78, 67–90 (section 2, Why folk psychology might (really) be false)

Descartes, René (1641), Meditations on First Philosophy, 6

Hempel Carl (1949/2000), 'The Logical Analysis of Psychology.' in Richard Jeffrey (Ed) Carl G Hempel: Selected Philosophical Essays, Cambridge University Press

Jackson, Frank (1982), 'Epiphenomenal Qualia', Philosophical quarterly, 32, 127–136

Putnam, Hilary (1975/1986), 'The Nature of Mental States' In Mind, Language and Reality. Philosophical Papers, Volume 2. Cambridge: Cambridge University Press.

Ryle, Gilbert (1949/2000), The Concept of Mind, London, Penguin Classics, Chapters 1, 2, 5 and 6

Shapiro Lisa (Ed) (2007), The correspondence between Princess Elisabeth of Bohemia and Rene Descartes, edited and translated. University of Chicago Press. Letters from May 1643.

Smart, Jack J C (1959), 'Sensations and brain processes', The Philosophical Review, 68 (2), 141–156



GLOSSARY

Definitions of key terms that crop up in the course and in philosophy in general.

A posteriori: Knowledge that can only be acquired from experience of the external world. For example, there is no way you could work out what temperature water boils at purely by thinking. You'd have to have experience of the external world to know the answer. *The opposite word: a priori*

A priori: Knowledge that can be acquired *without* experience of the external world, through thought alone. For example, you can work out what 900 divided by 7 is purely by thinking – you don't need experience of the external world to verify it. *The opposite word: a posteriori*

Analytic reduction: If X analytically reduces to Y, then the *meaning* of X is the same as the *meaning* of Y. For example, "triangle" analytically reduces to "3-sided shape" because the meaning of "triangle" is "3 sided shape".

Analytic truth: A proposition that is true in virtue of the meaning of the words. For example, "A bachelor is an unmarried man" or "triangles have 3 sides" or "1+1=2". Denying an analytic truth results in a logical contradiction. For example, the idea of a married bachelor or a 4-side triangle is a contradiction – it doesn't make sense. *The opposite word: synthetic truth*

Antecedent: The first part of a conditional statement such as "if A then B". In this example, the antecedent is "if A...". The second part of the conditional statement (i.e. "...then B") is known as the consequent. *The opposite word: consequent*

Arête: An ancient Greek word used by Aristotle to describe a property or virtue that enables something to achieve its ergon (function). For example, the arête of sharpness enables a knife to achieve its function (to cut things).

(Hard) Behaviourism: The view that all propositions about mental states can be reduced *without loss of meaning* to propositions about behaviours. For example, the word "pain" means nothing more than a set of behavioural dispositions.

(Soft) Behaviourism: The view that all propositions about mental states are propositions about behavioural dispositions. For example, the statement "he was in pain" is a statement about behavioural dispositions, but there is slightly more to the meaning of "pain" than these behavioural dispositions.

Blik: An unfalsifiable belief that is held in the face of conflicting evidence, but that is nevertheless meaningful.

Categorical imperative: A statement about what you should do that is not conditional on anything. According to Kant, the categorical imperative we should all follow is to "act only according to that maxim whereby you can at the same time will that it should become a universal law without contradiction".

Cognitive statement: Cognitive statements aim to literally describe how the world is and are either true or false. For example, "water boils at 100°c" , "triangles have 3 sides", "Paris is the capital of France", and "London is the capital of France" are all cognitive statements because they are capable of being true or false. *The opposite word: non-cognitive statement*

Consequent: The consequent is the second part of a conditional statement such as "if A then B". In this example, the consequent is "...then B". The first part of the conditional statement (i.e. "if A...") is known as the antecedent. *The opposite word: antecedent*

Consequentialist: An ethical theory that values as actions as good or bad according to their *consequences*. For example, utilitarianism is a consequentialist ethical theory.

Contingent truth: Something that is true but that *might not have been* true. For example, "Paris is the capital of France" is a contingent truth because they could have made Lyon the capital instead. Even "water boils at 100°c" is a contingent truth because the laws of physics could have been different. *The opposite word: necessary truth*

Contradiction: Two claims contradict one another if they cannot both be true simultaneously. For example, "today is Monday" contradicts the claim "today is *not* Monday".

Contradiction in conception: A maxim leads to a contradiction in conception if it would somehow be self-contradictory for everyone to follow it. For example, if everyone followed the maxim "to steal" then it wouldn't even be possible to steal because there would be no private property.

Contradiction in will: A maxim leads to a contradiction in will if you cannot rationally will that everyone follow it. For example, you cannot rationally will the maxim "never help people in need" if you yourself expect other people to help you when you're in need.

Deduction: A method of deriving true propositions from other true propositions (using reason). It is an a priori method of gaining knowledge.

Deductive argument: An argument where the premises are intended to logically guarantee the conclusion (i.e. an argument that is intended to be logically valid). For example, the logical problem of evil is supposed to logically guarantee the conclusion "God does not exist" whereas an *inductive* argument such as the evidential problem of evil is only supposed to provide evidence to supports its conclusion. *The opposite word: inductive argument*

Deontological ethics: Ethical theories that focus on *duty*, or what must be done. For example, Kant says we have a duty to follow the categorical imperative.

Direct realism: The view that a mind-independent external world exists and that we perceive it directly.

Disjunction introduction: The logical principle that if the statement "P" is true, then the statement "P *or Q*" must also be true.

(Ordinary) Doubt: Being unsure whether something in your everyday life is true. For example, you might doubt whether you locked the door when you left the house and so doubt whether your belief "I locked the door" is true. *Related: scepticism.*

(Philosophical) Doubt: Being unsure whether *anything* you believe is true. For example, you might doubt your perceptions and your own thoughts (and thus any beliefs formed from them) because of the possibility of global sceptical scenarios. *Related: scepticism.*

(Substance) Dualism: The view that minds exist and are a separate substance to bodies. In other words, there are two different kinds of substance: physical bodies and non-physical minds.

(Property) Dualism: The view that physical substances can have non-physical mental properties. These mental properties are neither reducible to nor supervenient upon physical properties.

Eliminative materialism/Eliminativism: The view that our common sense understanding of mental states is radically mistaken and that some or even all of these mental states don't exist and should be eliminated in favour of more accurate neuroscientific alternatives.

Emotivism: The metaethical view that moral judgements express emotional attitudes. For example, according to emotivism, 'murder is wrong' means something like "boo! Murder!".

Empiricism: Empiricism can mean different things depending on what it is contrasted against:

- Contra rationalism, empiricism says all knowledge of synthetic truths is acquired a posteriori.
- Contra innatism, empiricism says there is no such thing as innate knowledge and that all (propositional) knowledge is acquired after we are born.

Epiphenomenalism: A form of dualism that says that the mental and physical interact in only one direction: From physical to mental. For example, getting hit in the (physical) head causes the mental state of pain, but my mental state of pain don't cause anything itself. Instead, it's just the (physical) brain state that causes me to rub my (physical) head and say "ouch!".

Ergon: An ancient Greek word used by Aristotle to describe the function or characteristic activity of a thing. For example, the ergon of a knife is to cut things

Error theory: The metaethical view that moral judgements express cognitive statements but that moral properties don't exist and so all moral judgements are false. For example, "murder is wrong" is false because "wrongness" doesn't exist as a property.

Eudaimonia: An ancient Greek word used by Aristotle to describe the good life for a human being in the broadest sense. It is sometimes translated to "human flourishing". Eudaimonia is an objective property of a person's whole life and is the final end for humans.

Eternal: God is said to be eternal if He exists *outside of* time.

Everlasting: God is said to be everlasting if He exists *within* time but is without beginning or end.

Fallacy: The use of invalid or otherwise faulty reasoning.

Falsifiable: A belief or proposition that is incompatible with some *possible* observation. For example, "water boils at 100°c" is falsifiable because it incompatible with the observation of heating some water to 999°c and it not boiling. In contrast, an *unfalsifiable* proposition is compatible with *every* possible observation, and is thus said to be meaningless.

Felicific calculus: Bentham's formula for calculating net pleasure and thus for deciding the correct course of action according to act utilitarianism.

Final end: Something that is valuable in itself. You do not need to give further reason why it is valuable – it is valuable for its own sake.

Functionalism: The view that mental states can be characterised by their functional role within a cognitive system. For example, a functionalist might say that pain is an unpleasant mental state that creates a desire for it to stop and that causes the organism to get away from the thing that's causing it harm. The functional role of pain are multiply realisable: an octopus, an alien, and a human can all feel pain.

Gettier case: An example of a justified true belief that is not knowledge. Gettier cases thus show that the tripartite definition of knowledge does not provide sufficient conditions for knowledge.

Good will: To be motivated by and choose your actions for the sake of duty, according to Kant.

Hume's Fork: David Hume's claim that there are only two judgements of reason: *relations of ideas* (i.e. analytic truths known a priori) and *matters of fact* (i.e. synthetic truths known a posteriori).

Hypothetical imperative: A statement about what you should do *if you want to achieve a certain outcome*. For example, "you should leave now *if you want to get to the station in time*".

Idealism: The view that there is no such thing as a mind-independent external world and that what we perceive are mind-dependent ideas.

Indirect realism: The view that a mind-independent external world exists, but that we perceive it indirectly via sense data. This sense data is caused by and represents the mind-independent external world.

Inductive argument: An argument where the premises support the conclusion, but don't logically guarantee it. For example, the evidential problem of evil is supposed to provide good evidence for the conclusion "God does not exist". However, unlike a deductive argument, an inductive argument does not logically guarantee its conclusion. *The opposite word: deductive argument*

(Leibniz's law of) Indiscernibility of identicals: The principle that if A and B are the same thing, then A and B must have all the same properties. In other words, if two things have different properties, then they can't be the same thing. For example, "that can't be Smith's car because Smith's car is silver and that car is red".

Infallibilism: The definition of knowledge as true belief that is certain. According to infallibilism, anything that can be doubted is not knowledge. For example, on the infallibilist definition, I can't know "grass is green" because I might be being deceived by an evil demon.

Innatism: The view that there is some propositional knowledge that we are born with. Plato, for example, argues that humans have innate knowledge of geometry and attempts to show this with his slave boy example. *The opposite word: empiricism*

Intentionality: The "aboutness" of mental states. For example, you can have a thought about a frog, a belief about the moon, or a desire to drink beer (in this case, the mental state is directed at the beer – it is about beer).

Interactionism: A form of dualism that says that the mental and physical can interact in both directions. For example, getting hit in my (physical) head causes the mental state of pain, and my mental state of pain causes me to rub my (physical) head and say "ouch!".

(Rational) Intuition: The ability to know something is true just by thinking about it. It is an a priori method of gaining knowledge.

Invalid argument: An argument where it's possible for the premises to be true but the conclusion be false. For example, even though 1, 2, and 3 below are true, the following argument is invalid:

1. If Socrates is a man, then Socrates is a mortal
2. Socrates is a mortal
3. Therefore, Socrates is a man

Even though both premises (1 and 2) are true, and the conclusion (3) is also true, this argument is not valid because it's possible for the premises to be true and the conclusion to be false. For example, if Socrates was the name of my cat, then premise 2 (Socrates is a mortal) would still be true but the conclusion (Socrates is a man) would be false. *The opposite word: valid argument*

(False) Lemma: A (usually minor) premise in an argument. A *false* lemma is thus a false premise in an argument.

Maxim: A rule or law that you follow. For example, "not to steal".

Moral realism: The metaethical view that that mind-independent moral properties (e.g. good, bad, right, and wrong) and facts exist. *The opposite word: moral anti-realism*

Moral anti-realism: The metaethical view that that mind-independent moral properties (e.g. good, bad, right, and wrong) and facts *do not* exist. *The opposite word: moral realism*

Multiple realisability: Something is multiply realisable if there are many different things it can be. The claim that mental states are multiply realisable presents a challenge for behaviourism and type identity theory:

- Mental states can be realised through many different behaviours, which is a problem for the behaviourist claim that mental states are behavioural dispositions. For example, someone in pain may say "ouch!", but they may not if they don't want to look like a wimp.
- Against type identity theory, mental states can be realised by many different physical things. For example, a human, an octopus, can all feel pain but they each have physically different brains.

(Moral) Naturalism: The metaethical view that moral judgements express cognitive statements and that these cognitive statements are true or false in virtue of natural properties. For example, "murder is wrong" is true because "wrongness" is some natural property of the act of murder (e.g. it causes pain).

Naturalistic fallacy: G.E. Moore's phrase for the alleged philosophical mistake of reducing moral properties (e.g. good) to natural properties (e.g. pleasure).

Necessary condition: A condition that something must meet to be part of a concept. In other words, if anything that does not have this condition it will not be part of that concept. For example, "unmarried" is a necessary condition of "bachelor" because you have to be unmarried to fit the concept of a bachelor. *Related: sufficient conditions*

Necessary truth: Something that *must* be true (in all possible worlds). For example, "1+1=2" and "it is impossible for both a and not a to be true" are necessary truths

because there is no possible world in which they are false. *The opposite word: contingent truth*

(Moral) Nihilism: The view that moral values do not exist and thus there is no reason to justify or condemn behaving in one way over another.

Non-cognitive statement: Non-cognitive statements *do not* aim to describe how the world is and so are not capable of being either true or false. For example, "ouch!", "boo!", "hooray!", "don't do that!", and "shut the door please" are all examples of non-cognitive statements because they are not capable of being true or false. *The opposite word: cognitive statement*

(Moral) Non-naturalism: The metaethical view that moral judgements express cognitive statements and that these cognitive statements are true or false in virtue of non-natural properties. For example, "murder is wrong" is true because the act of murder has the non-natural property of "wrongness". Such moral properties are basic in that they cannot be reduced to anything simpler.

Ockham's razor: The principle that if two theories have the same explanatory and predictive power, the theory that invokes the fewest entities is usually the better one. In other words, the simplest explanation is the best (all else being equal).

Omnibenevolent: Perfectly good, i.e. nothing can be more good.

Omnipotent: Perfectly powerful, i.e. nothing can be more powerful.

Omniscient: All-knowing, i.e. nothing more can be known by an omniscient being.

Ontological reduction: If X ontologically reduces to Y, then X and Y are the same thing but "X" does not automatically mean the same thing as "Y". For example, although water ontologically reduces to H2O, it's possible that someone could not know this: they might know the proposition "there is water in the lake" and yet not know "there is H2O in that lake" because the meaning of these two propositions is different.

Phronesis: An ancient Greek word used by Aristotle to describe practical wisdom. Someone who has phronesis possesses virtues that provide a general understanding of good that enables them to think through, understand, and act virtuously. For example, they would be funny or angry or sad when the situation is right but never at an inappropriate time or in an inappropriate way.

Physicalism: The view that everything is physical or supervenes on the physical (e.g. properties and events supervene on the physical).

Prescriptivism: The metaethical view that moral judgements express prescriptive instructions that primarily aim to guide behaviour. For example, according to prescriptivism, "murder is wrong" means something like "don't murder people!".

Qualia: The subjective qualities of experience – i.e. what something feels like inside your mind. Qualia are not properties of objects, but properties of minds/experience. Examples of qualia include the redness of my experience when I look at a ripe tomato and the taste I experience when I drink beer.

Rationalism: The view that there are some synthetic truths that can be known purely through a priori means. For example, Descartes' cogito argument attempts to prove the synthetic truth "I exist" using thought alone and without reference to the external world. *The opposite word: empiricism*

Reliabilism: The definition of knowledge as true belief informed via a reliable method. For example, if I have good eyesight and see a car in the distance, my true belief "that's a car" is knowledge according to reliabilism because it is informed via a reliable method (my eyesight).

(Global) Scepticism: A belief that it is impossible to know *anything* because of the possibility of global sceptical scenarios. For example, "I might be dreaming, in which case I can't know today is Monday", or "I might be a brain in a vat, so I can't know that I'm currently reading a book". Global sceptical scenarios undermine our ordinary justifications and thus cast doubt on all our ordinary knowledge.

(Local) Scepticism: A belief that it is impossible to know anything *within a particular area.* For example, you might be a theological sceptic in which case you believe it is impossible to know anything about God's existence or character, or you might be a moral sceptic and believe it is impossible to acquire knowledge of moral truth.

Solipsism: The view that your own mind is the only thing that exists.

Sound argument: A valid argument with true premises (and thus a true conclusion).

Sufficient conditions: Conditions that, if all are met, guarantee that something is part of a concept. For example. "unmarried" and "man" are sufficient conditions of "bachelor" because everything that is an unmarried man is a bachelor. Being an unmarried man is sufficient to be a bachelor – you don't need to meet any other conditions. *Related: necessary condition*

Spatial order: Examples of organisation/order *within* the universe. For example, an eye is a complex organisation of matter that enables vision, and a watch is a complex organisation of matter that enables people to tell the time. *Related: temporal order*

Supervenience: X supervenes on Y if a change in Y is necessary for a change in X. This is mainly relevant in the discussion of property dualism: Property dualism says mental properties do not supervene on physical properties and so it is possible for two physically identical things to have different mental properties.

Synthetic truth: A proposition that is true in virtue of how the world is. For example, "grass is green" or "water boils at 100°c". Unlike an analytic truth, denying a synthetic truth does not lead to a logical contradiction. For example, even though "grass is green" is true, there is no *logical* contradiction in the idea of red grass. The idea of red grass makes sense, even if it is not how the world actually is. *The opposite word: analytic truth*

Tabula rasa: Latin for "blank slate". It is how Locke describes the mind at birth.

Tautology: A statement that says the same thing twice and thus is trivially true. For example, "red is red".

Temporal order: Examples of organisation/order of the *laws* of the universe. For example, the laws of gravity and electromagnetism are examples of temporal order because if they were different life would not be able to form within the universe. *Related: spatial order*

Tripartite definition of knowledge: The definition of knowledge as justified true belief.

Type identity theory: The view that mental states are identical to (ontologically reduce to) brain states. For example, type identity theorist might say that the mental state of pain is identical to c-fibers firing.

(Act) Utilitarianism: The ethical theory that happiness is good and that morally correct actions are those that maximise happiness and minimise pain.

(Rule) Utilitarianism: The ethical theory that happiness is good and that we should follow *general rules* that maximise happiness and minimise pain (even though there may be specific instances where following these rules does not maximise happiness).

(Preference) Utilitarianism: The ethical theory that morally correct actions are those that maximise people's preferences.

Valid argument: An argument where *if* the premises are true, the conclusion *must also be true*. For example, it is logically impossible for 1 and 2 below to be true but 3 be false:

1. If Socrates is a man, then Socrates is a mortal
2. Socrates is a man
3. Therefore, Socrates is a mortal

Note: A valid argument does not necessarily mean the conclusion is true. Even though the *logic* of the argument may be valid, the premises (and thus the conclusion) may be false.

Value judgement: A statement about whether something is good, bad, right, or wrong.

Verification principle: A.J. Ayer's claim that a statement can only be meaningful if it is either an analytic truth or empirically verifiable.

Virtue epistemology: Definitions of knowledge that invoke intellectual virtues. Intellectual virtues are things like caring about the truth and thinking rationally. Zagzebski's definition of knowledge, for example, says knowledge is belief that arises from an act of intellectual virtue.

Virtue ethics: Ethical theories that focus on what a good *person* is rather than what good actions are. For example, Aristotle's ethics talks about virtues and vices of character and how this relates to being a good person and achieving eudaimonia.

(Philosophical) Zombie: A being in another possible world that is physically identical to an ordinary human in every way but lacks qualia. For example, the zombie version of you would have exactly the same brain states and behaviours as you but would not have the corresponding phenomenal experience/qualia.

Printed in Great Britain
by Amazon

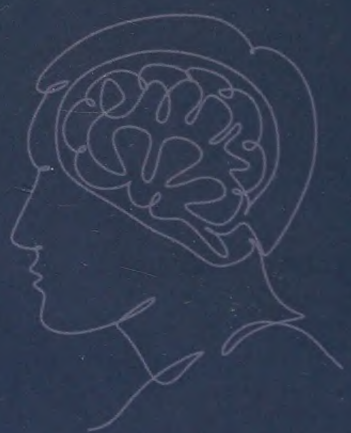




How to get an A in A-level philosophy is the course textbook written with the student in mind. Focused on the AQA philosophy syllabus from 2017 onwards (7172), this book covers all four modules in a simple, concise, and focused manner. Using plenty of examples, each idea is broken down into what that idea is, the arguments for that idea, and the arguments against. Later sections explain (with examples) how to communicate these ideas and arguments to achieve maximum marks in the exam.

Inside you will find:

- Explanation of topics for all four course modules:
  - Epistemology
  - Moral philosophy
  - Metaphysics of God
  - Metaphysics of mind
- Bullet point summaries at the end of each module
- Exam blueprint for each question type with examples
- Example 25 mark answer plans for every major topic
- 3, 5, 12, and 25 mark practice questions
- Revision exercises including multiple choice quizzes
- Glossary of key terms





Christian B. Feest is a former sixth form lecturer in philosophy. He has taught students from all over the world and is the creator of ***philosophyalevel.com*** *– a resource used by thousands of philosophy students every day.*

PHILOSOPHY
alevel.com

ISBN 9798719524658