'This book is ridiculously comprehensive'
Ross Cameron, University of Virginia, US

'This is an extremely useful teaching resource for academics'
Graham Stevens, University of Manchester, UK

HOW TO GET PHILOSOPHY STUDENTS TALKING

AN INSTRUCTOR'S TOOLKIT

ANDREW FISHER AND JONATHAN TALLANT

ROUTLEDGE

How to Get Philosophy Students Talking

Engaging undergraduate students and instigating debate within philosophy seminars is one of the greatest challenges faced by instructors on a daily basis. *How to Get Philosophy Students Talking: An Instructor's Toolkit* is an innovative and original resource designed for use by academics looking to help students of all abilities get the most out of their time spent in group discussions.

Each chapter features thought experiments, discussion questions and further readings on topics within the following core areas of philosophy:

- Metaphysics
- Epistemology
- Philosophy of Mind
- Philosophy of Language
- Philosophy of Religion
- Philosophy of Science
- Political Philosophy
- Normative Ethics
- Applied Ethics
- Metaethics
- Aesthetics

Group discussions and debates are a key part of undergraduate study and one of the best ways for students to learn and understand often complex philosophical theories and concepts. This book is an essential toolkit for instructors looking to get the most out of their philosophy students.

Andrew Fisher is Associate Professor in Philosophy at the University of Nottingham, UK.

Jonathan Tallant is Associate Professor in Philosophy at the University of Nottingham, UK.

How to Get Philosophy Students Talking

An instructor's toolkit

Andrew Fisher and Jonathan Tallant

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2016

by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada

by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Fisher, Andrew, 1976–
How to get philosophy students talking : an instructor's toolkit / Andrew Fisher and Jonathan Tallant. — 1 [edition].
pages cm
Includes bibliographical references and index.
1. Philosophy—Study and teaching. I. Title.
B52.F58 2015
107.1′1—dc23
2015017534

ISBN: 978-1-13-882786-8 (hbk)
ISBN: 978-1-13-882787-5 (pbk)
ISBN: 978-1-31-567064-5 (ebk)

Typeset in Times new Roman
by Swales & Willis Ltd, Exeter, Devon, UK

Contents

Introduction

Why?

It has always seemed a little strange that 'differentiated learning' is discussed in primary, secondary and further education, but not in higher education. In fact it is seen as quite a controversial subject, which has been so neglected that it is worth saying what it is and what it is not.

By 'differentiated learning' we mean the process by which academic content is delivered in a way that allows students with different abilities and learning styles to engage with it. The idea is to design the lecture, seminar, discussion group, workshop, etc. in such a way that it does not intimidate the student who is not yet at a stage to engage with it at a higher level. Yet, at the same time, it has to keep the interest of the students who have thought about the content so that they do not become bored and 'switch off'.

What it is *not* about is providing lots of different content, or individualized tasks for people. To be clear, differentiated learning isn't about 'streaming' or 'setting' at university. The fact that this is not widely discussed is strange, because it is undeniable that with the increase of students attending university and the fact that only at some institutions is philosophy an option, students enter universities with a varied skill set and variable background knowledge. Some students will not be taking philosophy because they are passionate about it. Many will come to lectures or seminars ill prepared, having not done all the required reading or not having given a second thought to the issues.

Of course this is *not* to say that students are lazy or unable to become engaged with the material. It is to say that different students have different priorities. This is particularly hard for academics to understand given the very fact that they are academics, and their own experience as students was most likely as people who were passionate about their subject – and consequently their subject was, and is, a high priority in their lives.

Why this book?

How might this relate to this book? Well, one area where we feel that *differentiated* learning – and we'll get into the details in a minute – is neglected is in

seminars. By 'seminar' we mean small-group discussions typically related to the topics of the lecture, usually involving fifteen to thirty students.

We have been teaching in universities for a number of years and in our experience there are many discussions about lectures; training courses on 'how to lecture'; 'lecturing for learning', etc. We talk about interaction in lectures, about PowerPoint vs Prezi, about how to run two-hour lectures, etc. The textbooks we use are often split into chapters which neatly fit our lecture series. Seminars are then often an afterthought, and can sometimes be taken by teaching assistants, with little guidance from the person running the module.

But this is odd. If you think, as we believe, that philosophy is best learned through discussion, then we would have supposed that if anything, lectures would be the afterthought and seminars would be the central focus. We would have expected many books to be devoted to running philosophy seminars, about exercises to run in seminars and suggested content, etc. Yet these books are not available, and that is why we wrote this book as a complement to lectures, so as to help organize effective seminars.

We have been at many meetings with academics where teaching is discussed, and a familiar 'flavour' of discussion is as follows: 'We can't get the students to say anything in the seminars!' 'They never do the reading, they are ill prepared!' 'It is the same old few who make contributions in seminars!' We suspect that you've heard these, too. Notice that in most of these discussions blame is firmly placed at the student's feet. It is the *student* who is not engaged, the *student* who hasn't done enough, who is lazy, etc. It is also worrying that the same conversations are being had at different institutions and have been taking place over a large number of years. Surely alarm bells should be ringing! Should we not be changing what *we* do? If something is broken don't we try and fix it? Don't we try to find a solution?

So if we are going to 'fix it', what does that mean? Well, let's consider the sorts of things we say as academics. Perhaps this will give us a clue as to how to proceed: 'All I set them to read last week were two papers by David Lewis! And they still didn't do it.' 'I didn't have time to set any questions for them this week, and they had nothing to say!' 'I reduced the reading down to one chapter, but some said they couldn't find it, and those that had found it only skim-read it!' What is going on? Simply put, no thought has been given to differentiating the material. After all, in seminars there will be students who have read all the set work and who have thought hard about the ideas; there will be students who attended the lectures and have a general sense of what is going on but only had a chance to prepare for the seminar the day before; students who prepared for the seminar by flicking through their notes on the bus in the morning and students who have done no preparation whatsoever. But more often than not we expect all students to be in the first category.

What should we do then? It is no good getting mad, throwing students out, using the session time to read what should have been read before, etc. We should have

reasonable expectations of the students. Notice that the word 'reasonable' is not the word 'no'. There is a famous story – maybe apocryphal – of the Baptist minister Charles Spurgeon who, when talking to a new minister who was dismayed that people weren't becoming Christians, said 'but you don't expect people to become Christians after *all* your sermons do you?' The new minister replied, 'Well, of course not!' 'Well' said Spurgeon, 'that is your problem.' Similarly if we have no expectations of students then they will meet them. So what should our expectations be?

Well, there are different answers, depending on which academic year the student is in, whether she is majoring in philosophy, etc. But across all of these areas what is absolutely key to getting anywhere is to 'differentiate' the learning. We should expect all students to get involved, to have something to contribute and to feel able to contribute. They should not feel intimidated; they should be more interested and engaged in the topic when they leave the seminar room compared to when they entered it.

Now, this sort of talk will be anathema to some because they fail to notice that 'differentiate' isn't the same as 'dumb down'. We are not – *not* – saying that the content that we expect students to know should be more basic; that we should not use primary texts, that we should never use logical notation, that we should change the marking criteria, etc. Rather we suggest that we need to put in place different ways for students to engage with this material. Differentiation is simply about accessibility to the material, not the nature of the material.

What the authors have found after many years of teaching is that if you set a paper for students to read, and this is particularly true for first-years, some will read it, some will skim-read it, and some will not read it. So what happens in the seminar as a result? There will be those students who talk while others listen, there will be students who are daunted by the whole process and there will be those who are obviously bored and probably will not turn up again. In this situation the seminar leader can find that the session turns into a mini lecture.

However, what happens if the seminar leader, recognizing there is no 'in' to the discussion, outlines a key 'thought experiment', 'story', 'argument', or 'idea'? This then gives everyone an entry point but still allows for the students who have done the reading to show their deeper understanding. This is a more successful seminar. In our experience seminars fail when the seminar leader is determined to extract from the students 'the answers' to the questions which were set previously. Then very often the leader will continue by lecturing the group for some time on what she believes to be the key features of the readings which she has set. But if our account of a 'successful' seminar is right, why not make the seminars *about* the thought experiment/story itself? Why not set out with the intention of making *that* the 'stepping off' point? This is why we wrote this book.

In this approach the content can be differentiated. How? Well, each of the thought experiments are short enough that students can read them on the bus on the way in, or in fact, they can just be read out in class. But, the 'further reading' is sophisticated and related to primary texts, and the questions we have written

should lead students to think about these issues in more depth and ask themselves how the issues might link up with different topics.

The budding academic can then read the thought experiment, prepare answers to all the questions (which if done properly is a complex piece of work) and read the suggested work. Those students who want to prepare less can do so easily but can still be engaged in the discussion. Moreover, our belief is that when they have listened and participated they are more likely to leave the seminar room with a wish to look deeper into the topics.

The authors have been trialling this approach with great success and have found that the students are much more willing to contribute. They typically feel 'safer' about trying out their ideas. And we have found they are much more likely to prepare something than nothing at all, and in fact are more likely to attend seminars. Furthermore, the more engaged students aren't isolated either. They aren't now the only ones talking in a group of peers who haven't read anything and hence remain silent.

How you use the thought experiments in class is of course up to you. There are some suggestions below. One thing this differentiated approach brings into sharp relief is the question 'What is a successful seminar leader?' If the answer is to expect the student to answer all the questions that are set or if it is about all students having a good grasp of the intricacies of the topic before they leave the seminar room, then you won't be a good seminar leader. On the other hand, if you want all students to contribute, move along in their own understanding and enjoy themselves, then this approach is, we submit, not only the best, but is achievable. Of course, some academics will be unconvinced. However, the success of such an approach is not something that can be decided 'a priori'. So we suggest that you should give it a go and see what happens.

What are these?

Within the book, there are a lot of chapters: 110 in total. Each chapter contains an example. We've described them, above, as thought experiments, although we do concede that perhaps that is a little bit of a stretch. As we understand a thought experiment, formally speaking it should involve the consideration of some hypothesis within a hypothetical scenario, and whilst some of the examples that we give perhaps satisfy that description, it wouldn't be true to say this of all of them.

When we speak of 'thought experiments', then, we mean the term a little more loosely. Each chapter of the book contains a very brief introduction to a topic, perhaps just framing an issue, followed by an example – which we call the thought experiment. What all of these examples have in common is that they invite the reader to consider a situation or an example – all of which are hypothetical. We follow this with a short discussion and then a series of questions. Thus, we're not necessarily inviting the reader to test any particular hypothesis.

The examples, then, are certainly exercises in thought. And whilst perhaps not formally analogous to a robust scientific experiment, they are intended to allow

and prompt the reader to experiment with various different ideas for themselves. Further, the questions that come towards the end of each experiment are intended to help further those thoughts and facilitate wider discussion.

In the end, then, we don't much mind whether or not these should be *formally* regarded as thought experiments, or as something else. Because they facilitate experiments in thought and invite the reader to engage in consideration of a hypothetical scenario, we're prepared to call them thought experiments – we certainly couldn't think of a better name.

As we have just explained, in addition to the thought experiments themselves there is a small amount of text before and after. For the most part, these are simply framing devices, providing a little bit of background in order to set up the particular case under consideration. In a number of cases, the text will also give a cursory sketch of some of the ways in which philosophers have responded to the thought experiments (or the ideas that they represent). These are very, *very* sketchy. We do not claim even for a moment that these would serve to give a student a full and complete introduction to one of the topics. As we've flagged elsewhere in the Introduction, there are important senses in which we see these experiments and discussions as being more of a route into philosophical ideas rather than a presentation of those ideas themselves.

Equally the reader should be aware that the way we have demarcated the topics ought not to suggest that each topic stands alone. Of course, it is true that each chapter could be read independently of the others, and we hope that the reader will feel free to dip into various chapters. However, what we don't want to misrepresent is the interconnectedness of philosophy. The world, of course, isn't demarcated into chapters, modules, courses, etc. To counter this we have indicated the interlinked chapters and direct the reader to various related topics.

Towards the end of each chapter we have included readings and ten questions. The questions are intended to get the reader thinking about the problems. We have used these kinds of questions in seminars *as* the questions set for seminars, so we also think that they can be used to spark conversation and discussion. Once again, and this is important, we don't think that any reader should take the view that her ability to answer these questions is a clear indicator that she's exhausted the topic! Rather, these sections are intended to ask questions of the reader that will prompt interest and, in some cases at least, require or inspire them to start engaging in further reading of their own.

Similarly, the readings are fairly sparse. We do not suggest that it would be possible for any reader to completely grasp a topic by simply reading the four items we've listed. In each case, we've listed things that either we think are seminal, or else that we think are accessible and/or particularly interesting. (Being seminal and being particularly accessible and interesting don't always overlap, of course.) As with any suggested reading, there's plenty of scope for disagreement: we don't think that our suggested readings are definitive. For most of the chapters, we'd also recommend that the reader takes a look at the *Stanford Encyclopedia* entry on

the topic, and use the references there to compile a further reading list for themselves. We'd also recommend that the reader has a careful look through *Philosophy Compass* for articles that will give a survey or introduction to a topic that will provide far more detail than we do here.

The approach here, then, is not intended to force the student to engage with complex written papers before she's fully digested the lecture; nor is it intended to merely regurgitate lecture materials. But by focusing attention upon a conversation piece that can be supported by lecture material and reading, we hope to try to encourage students to *think* and to then externalize their thoughts and have their ideas challenged by their peers.

Why didn't you include . . .

As soon as we sat down to write this book, we knew that we'd have to leave some very important things out. The list of things we left out would easily be as big as this book, and it would be futile to start listing particular topics. However, it is worth saying a few general things: we have not included any 'Continential' philosophy – there is no Derrida or Deleuze, for example. Nor is there any 'non-Western' philosophy – there is nothing about the thinking of Confucius, or Buddhist teachings. This was intentional. The simple fact is that neither author knew enough about these areas to do justice to their complexities.

Two other notable areas which we *have* included but *haven't* given distinct areas are feminist philosophy and the history of philosophy. We needed to keep the project manageable (and, again, to be mindful of our own areas of expertise) and so decided to touch on these themes within the other subject areas. So for example, in the philosophy of language section the reader will be able to follow a 'history of thought' from Frege. And within a number of the sections there are references to feminist thinkers, with some having their own chapters, e.g. Chapter 58, on Catherine MacKinnon's discussion of discrimination.We hope that what has been included doesn't suggest that it is the most important area in philosophy, and that what has been left out suggest that those topics are of less value. This simply isn't true.

Using the thought experiments

We don't think that there's a single *right* way to use these thought experiments: there are a number of different ways that courses could include them. Here, we outline just a few. As will be apparent from even a cursory inspection, each section of the book includes ten chapters on a particular theme: for instance, philosophy of mind, metaphysics, epistemology, philosophy of religion, etc. We see these as sitting alongside a course on a similar topic for first- or perhaps second-year level (we'd expect third-year students to be a little too advanced for this kind of strategy to be optimal, though we encourage readers to decide for themselves).

Option 1

For a ten-week lecture and seminar course, use one of the thought experiments for each seminar. The lectures will provide a general introduction to the topic in question and the chapter will provide seminar preparation. Ask students to read the chapter carefully and work through the questions, providing answers to all of the questions. Use the seminar to work through the questions that you think are most pertinent to discuss with your students. Recommend further reading, perhaps drawing on the suggestions in the chapters, to help students to progress to a standard where they can write good philosophy papers.

This, we imagine, will be a fairly standard way to run sessions.

Option 2

For a ten-week lecture and seminar course, use one of the thought experiments for each seminar. The seminars should, in this case, precede the lecture. Ask students to read the chapter carefully and work through the questions, providing answers to all of the questions. Use the seminar to work through the questions you think are most pertinent to discuss with your students. Then, use the lecture as a springboard to introduce philosophical content. Recommend further reading, perhaps drawing on those suggested in the chapters, to help students to progress to a standard where they can write good philosophy papers.

We think this approach has particular benefits, especially in modules of a relatively abstract nature (e.g. metaphysics, philosophy of language, etc.) where students typically don't have strong intuitions prior to learning about the material and can find it challenging to latch on to key ideas and themes. By engaging them in specific problems prior to introducing philosophical solutions to those problems, we think that the thought experiments can be used to provide a route into material that may otherwise pass students by.

Option 3

For a ten-week higher-level course, follow the plan described in Option 1 for five weeks. Then, for future weeks, invite (small groups of) students to prepare thought experiments, using the examples from the book as a guide, and bring them to seminars. Ask them to set five questions for discussion that are based upon the thought experiment and their reading. These should then provide a platform for seminars, whilst also allowing and encouraging these more philosophically experienced students to engage in relatively independent study.

We think that this approach has particular benefits in upper-level courses where students have greater philosophical experience and maturity. By introducing the approach in the first five weeks of the course, students will become attuned to what is expected of them (and what works for them). In the last tranche of seminars, they begin to take much more responsibility and enjoy designing the key questions

and thinking about how they can encourage their peers to discuss these questions through the choice of experiment that they make.

Option 4

The fourth way is to use the thought experiments as an occasional prompt. Over the years, we've both worked with seminar groups that won't talk and those that will. This isn't new. For some groups, simply asking them to read a particular paper and think about it is enough. For others, you can set the same task and nothing happens. They don't read it. Or, if they do, the students don't think about it at all. It's often hard to see what to do with such a group. Sometimes, though, the problem is one of access: for whatever reason, the students simply aren't able to access the material that other groups can. (We're assuming that these students are trying very hard, but not quite succeeding.)

In such situations, we imagine that the thought experiments could be very useful. Photocopy the chapter and take it along to the seminar. Ask the students to read the thought experiment. What do they think of it? Do they understand it? Is it silly? How would they set about answering the questions set?

On this latter task: we find that asking the students to work in small groups, within the seminar, can be productive; there can be reservations about speaking in front of a whole group that disappear when just working through a discussion piece with two or three like-minded friends.

Option 5

Last, but by no means least: we think that this book can be used outside the university context. We both love philosophy. We would both love to see reading groups, discussion groups, U3A and philosophy classes in schools use this book. If that's you, and you're reading this, then we think that there are lots of ways you could use it – though most of them mirror the kinds of thing that we've outlined in the first four options. Use the chapters as a starting point for a discussion (nothing gets people engaged with philosophy quite as much as talking about it). From there, you can springboard in any direction you like (though, again, we think that the *Stanford Encylopedia* articles are a great reference point), with people raring to go and becoming passionate about the topic.

Conclusion

So there we have it. One hundred and ten chapters, full of ideas and plans for seminars across eleven different courses. We have found some of the strategies we describe above to be very successful in getting people thinking and talking about philosophy. We hope that you do, too.

Part I

EPISTEMOLOGY

1 The ‘a priori’

Some knowledge looks like it comes to us through our senses. For instance, I know that there is a computer on my desk. That I know this is due (in significant part) to my *seeing* the computer on the desk. Similarly, I know that the surface of the desk is smooth. That I know this is due to my *feeling* the surface of the desk. This is all well and good.

But there is another kind of knowledge that – at least in some sense – seems to come to us via other means. This knowledge – ‘a priori’ knowledge – seems to be such that we can gain it via reflection on particular concepts or ideas and is – again, at least in some sense – independent of our senses.

Thought experiment

> Ethel is reasonably adept at mathematics. She learned to count in the usual fashion, in part through beginning to learn to count objects in the world around her, and then, further on, by being taught mathematics in school. At 18, she’s now capable of performing all sorts of sums very competently. However, she’s never (as a matter of fact) previously worked out (or been told) the result of the sum: 476 + 971.
>
> Needing to work out some financial matters pertaining to her university accommodation, Ethel now needs to work out the result. She does so completely in her head, without consulting a calculator, phone or other source of information. She comes to *know* that the answer is 1447. She did not see this answer written down. None of her experiences tell her that this is the right answer. Ethel has learned something completely new by working out what the answer is.

There is, we might think, something distinctive about this kind of knowledge that Ethel has acquired. Unlike a case where we need to experience some new sensory input, this knowledge has been derived by Ethel seemingly without any kind of novel input. She just *worked it out*. Other things that we can come to know in this kind of way include the result of logical deductions and (more contentiously) certain truths of metaphysics, where it seems we often try to work out the answer to some tricky puzzle simply through reflection.

Nonetheless, there are some concerns about the very notion of ‘a priori’ knowledge. First, it is tough to know exactly what we mean by it. For instance, it’s tempting to say that it’s knowledge that’s been derived independently of experience – but, as we can see from the above, that isn’t quite right. Ethel’s experiences of having learned mathematics are clearly playing some part in enabling her to acquire new knowledge. Second, it can often appear difficult to explain quite what process is underpinning our acquisition of ‘a priori’ knowledge. Suppose, just for a moment, we say that ‘a priori’ knowledge is described as knowledge that is acquired independently of the senses; how, then, *is it acquired?* Imagine that both

476 and 971 in the above stand for amounts of money that have been paid into her bank account. In that case, in performing that sum, Ethel is working out how much money is in her bank account: how can mere thinking give Ethel access to her bank account? This looks very strange! Just by thinking about things I cannot access the contents of my fridge: how is it, then, that Ethel's reflections manage to accurately gauge the content of her bank account?

For these and other reasons, some philosophers can be pessimistic about the existence of 'a priori' knowledge. But, that said, there does seem to be something going on in the case that is different from the kinds of ordinary sensory knowledge that we acquire – such as my knowing that there is a computer on the table. For that reason, it still seems tempting to say that there is 'a priori' knowledge, and that we must find some way in which to deal with these challenges.

Questions

1. Write down as many differences as you can between knowledge acquired through perception (where I come to know that there is a computer on my desk) and the case involving Ethel.
2. Using your answer to (1), can you come up with a working account of how you might understand the notion of 'a priori' knowledge?
3. Setting aside your answer to (2): one definition we started with was that 'a priori' knowledge is knowledge that is not acquired through the senses. Can you think of counterexamples to this? Can you think of how to alter the definition to deal with the counterexamples?
4. Knowledge acquired through the senses seems to require a sensory input. There does not seem to be any analogous input in the case of 'a priori' knowledge. Can you think of any input that might do the job – even if it's not sensory?
5. Following on from (4): if there is no input, can you think of another way of explaining how we seem to be able to use 'a priori' knowledge to tell us about the world around us?
6. If 'a priori' belief is knowledge that is independent of sensory input (in some sense) do you think that we should deny that it is knowledge?
7. If we could not define what 'a priori' knowledge is, would that pose a problem for the notion?
8. Would it matter if there was no such thing as 'a priori' knowledge?
9. Some 'a priori' knowledge might seem to be a product of the meaning of our terms: for instance, we can know 'a priori' that if all women are mortal and that Iselde is a woman, then Iselde is mortal. Does that case seem to you to be importantly different from the mathematical case? Why/why not?
10. Knowledge that is not 'a priori' is normally described as 'a posteriori': it is knowledge that we gain through the senses. Do you think that we might ever need to invoke a third category of knowledge?

Reading

Devitt, M. 2005. 'There Is No A Priori', in M. Steup and E. Sosa (eds), *Contemporary Debates in Epistemology*, Oxford: Blackwell, 105–15.

Field, H. 2006. 'Recent Debates About the A Priori', in T. Gendler and J. Hawthorne (eds), *Oxford Studies in Epistemology*, Oxford: Oxford University Press, 69–88.

Jenkins, C. 2008. 'A Priori Knowledge: Debates and Developments', *Philosophy Compass*, 3(3), 436–50.

Jenkins, C. 2012. 'A Priori Knowledge: The Conceptual Approach', in A. Cullison (ed.), *The Continuum Companion to Epistemology*, London: Continuum Press, 180–98.

---

2 Contextualism

Knowledge is a funny thing. We're mostly happy to attribute it to people. If asked, many of us will claim to know *what the time is*, or *how to get home*. These are things that we know. But, as we shall see in Chapter 9 (Scepticism), it seems that there are ways to get people to reverse their judgements. Bring up the possibility of an all-powerful and all-deceiving demon, and all bets are off. In those kinds of cases, it can seem as if we don't know anything at all.

Perhaps, then, whether or not it's true to say that someone knows something varies according to the context that we're in. To give a comparison, suppose that I tell you that Bob is tall (for the record, Bob is 6ft 2in. or 187.96cm). My claim is true. But now suppose that Bob's playing basketball with some friends of his. If I said 'Bob is tall' in that context, I'd likely be corrected. In *that* context, it's just not true that he's tall. So, what else can we say about knowledge that might make it seem as if knowledge really does depend upon context? Here's a thought experiment:

Thought experiment

Part I

> Arnold is going to the bank. It's a Saturday morning.
>
> Arnold is walking along with Beth. Beth asks, 'do you *know* whether or not the bank is open?' Arnold pauses, 'yes.'

Part II

> Arnold is going to the bank. It's a Saturday morning. It is absolutely imperative that Arnold get to the bank before it closes. If he fails to do so he and Beth will both default on a payment that is already late and they might well end up in prison.
>
> Arnold is walking along with Beth. Beth asks, 'do you *know* whether or not the bank is open?' Arnold pauses, 'no, I'm not sure.'

This two-part thought experiment is supposed to bring out the fact that although someone – in this case, Arnold – may have the same justification for believing a proposition in two different circumstances, whether or not their claim to knowledge is correct might vary. It seems that, in cases like Part II, the *stakes are higher*.

It matters much more, now, whether or not Arnold knows that the bank is open. And because of that, it's tempting to see his response as correct. Because it matters so much more that he knows in Part II than it did in Part I, we're more reluctant to say that Arnold knows in Part II than we were in Part I.

In general, the contextualist will appeal to cases like this to motivate the view that in some (low) stakes cases we can know quite a bit. But in some high stakes cases – perhaps cases like that of the evil demon – we know very little indeed.

Questions

1. Do you think that Arnold's responses are reasonable? Explain your answer.
2. Vary the case: in Part II, imagine that Arnold doesn't know that he and Beth are close to defaulting on their payment. Would *we* say that he knows?
3. In the case just described, would *Arnold* say that he knows?
4. Given your answers to (2) and (3), whose context do you think matters the most? The person who is said to know, or the person who is making the claims about knowledge?
5. How, if at all, could this kind of contextualist position be used to explain how, in sceptical scenarios, we know nothing at all?
6. How plausible do you find contextualism, as outlined above?
7. Do you think that the parallel with 'tall' is a good one in the case of knowledge?
8. Do you think that the contextualist is trying to track the way that people normally use the term 'knows', or something else? If 'something else', then what?
9. Would it be useful to track the way people normally use the term 'knows'? How might we do that most effectively, given the technology available to us as philosophers?
10. Why might the contextualist position be important or useful?

Reading

Brown, J. 2005. 'Comparing Contextualism and Invariantism on the Correctness of Contextualist Intuitions', *Grazer Philosophische Studien*, 69(1), 71–100.

DeRose, K. 1999. 'Contextualism: An Explanation and Defense' in J. Greco and E. Sosa (eds), *The Blackwell Guide to Epistemology*, Malden, MA: Blackwell, 187–205.

DeRose, K. 2005. 'The Ordinary Language Basis for Contextualism, and the New Invariantism', *Philosophical Quarterly*, 55(219), 172–98.

Schaffer, J. and Knobe, J. 2012. 'Contrastive Knowledge Surveyed', *Noûs*, 46(4), 675–708.

3 Epistemic injustice

We can agree that prejudice and bias against particular groups of people are bad. It would be wrong, for instance, for us to pay Sally less than Mark for doing the same work in the same way, just because Sally is female and Mark is male. It would be wrong for us to pay Mark less than Sally for doing the same work in the same way, just because of the colour of Mark's skin. In purely moral terms, so much seems entirely right.

But are there epistemic considerations (considerations to do with knowledge) here, too? We might well think so. Following Miranda Fricker, we borrow the following case from Harper Lee's *To Kill a Mockingbird*:

Thought experiment

> It is 1935 and a young black man, Tom Robinson, is charged with the rape of Mayella Ewell – a young white girl. Tom Robinson's left arm is disabled following an accident he suffered when he was a boy. It is clear to the reader and any unprejudiced witness that Tom is innocent and has been proven such by Atticus Finch, the counsel for the defence.
>
> Nonetheless, with this being a case of her word against his, Tom is (wrongly) found guilty. The jury do so in spite of the evidence and in no small part because they go along with the automatic and unreflective distrust generated by the prejudices that structure their perception of the speaker. The jury *see* Tom as someone who is not to be trusted. In that sense, just because of the colour of his skin, Tom is not viewed as trustworthy by his peers.

This case, and others like it, are what Fricker describes as epistemically unjust. It is what she describes as *identity-prejudicial credibility deficit*. Very roughly this is the idea that because of his social identity (being a black man) and because groups are prejudiced against people with that identity, his credibility is judged to be low (or lower than that of someone else), merely because of some feature of his identity.

For obvious reasons, this is morally unjust: neither the colour of someone's skin nor their gender makes the slightest bit of difference to their trustworthiness, and it is morally wrong to think otherwise.

But notice that Tom is also done an *epistemic* injustice. That is, he is judged to be less credible, a less good source for justification, testimony and evidence, because of the colour of his skin. Not only, then, is this to do Tom a moral disservice, it is also to do him an epistemic disservice. Within the confines of the story, Tom's testimony is simply held to be of less value than it really is. Epistemically, we should respect how reliable someone really is, rather than allow prejudice to deflate their worth.

Questions

1. In the case described, we find a case where a fictional black man is not trusted by his peers. Find a real case from the news where, plausibly, exactly this has happened.
2. Are there other, similar cases that you can find that involve other people who may be negatively stereotyped, leading to identity-prejudicial credibility deficit?
3. Without copying what is written above, write down what you understand to be involved in identity-prejudicial credibility deficit.
4. Write down the difference between epistemic injustice and legal or moral injustice.
5. Why is epistemic injustice a bad thing?
6. In what way(s) is this topic relevant to the ways in which we live our lives?
7. How might we try to prevent identity-prejudicial credibility deficit from occurring?
8. Do you think that there could be identity-prejudicial credibility excess; cases where people's testimonies are given undue weight because of prejudices we have about social identity? If so, write down some examples.
9. Is identity-prejudicial credibility excess problematic? Why?
10. What would you say to someone who denied the existence (or problematic effects of) identity-prejudicial credibility deficit?

Reading

Anderson. E. 2012. 'Epistemic Justice as a Virtue of Social Institutions', *Social Epistemology*, 26(2), 163–73.

Coady, D. 2010. 'Two Concepts of Epistemic Injustice', *Episteme*, 7(2), 101–13.

Dotson, K. 2011. 'Tracking Epistemic Violence', *Hypatia*, 26(2), 236–57.

Fricker, M. 2007. *Epistemic Injustice*, Oxford University Press.

---

4 Intuitions

Throughout various different parts of philosophy, it's likely that you'll see various positions described as 'intuitive', or 'seeming right'. This is often used as an indicator that the theory or position in question has something going for it. After all, we want our best theories to capture our intuitions, don't we?

Well, according to some philosophers, it's not quite so clear. There's quite a bit of uncertainty as to exactly what an intuition *is*, whether such a thing might be useful to us, and the extent to which we might trust our intuitions. To bring out at least a part of the concern, here's a thought experiment:

Thought experiment

> Consider the question: 'is red better as a colour than green?' The topic is absurd, but put that to one side. Suppose that we get folks to consult their intuitions and it turns out that we have a rough consensus: most people engaged in academic philosophy report the intuition that red is a superior colour to green. Perhaps we could then suppose that this gives us some minimal, defeasible reason to think red better than green.
>
> But now suppose that we engage in some fieldwork, some experiments, and we find that these are only the intuitions reported by white, male, philosophers. When we ask other people (women, people who are non-white, people who are not philosophers), we find that their intuitions vary hugely. It turns out that there is no agreement *across* these groups as to which colour is the best.

The example is clearly silly. But the point is not. In the case described, with such varied intuitions being reported, it might seem ridiculous to give any weight at all to the intuitions of one small group.

Yet, in a range of studies, similar kinds of result seem to have been uncovered, on more serious subject matters. Thus, when investigating the intuitions of diverse groups on Gettier cases, experimenters found that these alleged counterexamples were not held to be counterexamples by some groups, to the same extent.

Indeed, right the way across philosophy, there are experiments into the kinds of intuitions that people have about particular topics, and these experiments seem (in a good number of cases) to be in tension with the kinds of claim that philosophers often report as being intuitive, and in cases where those philosophers also report intuitions as being important motivational factors supporting their theories.

This has led to some folks suggesting that vast swathes of academic philosophy need to be re-thought. After all, if intuitions are being used so extensively but are fundamentally unreliable guides to our best theories, then it would hardly seem reasonable to persist with the same methodology.

In contrast, the data has also led some to try to offer a defence of the philosophical methodology against this charge. In so doing, some philosophers have claimed that intuitions are playing only a very small (and non-essential) role in philosophy. Others have tried to defend the line that the only intuitions that matter (or, at least, those that matter most) are those of philosophers.

Questions

1. How valuable do you think it might be taken to be if a particular theory is intuitive?
2. Can you think of a specific case where it's claimed as a virtue of a theory that it's intuitive? How much difference did that seem to make to the theory? Would it make any difference if we jettisoned the claim?

3 What reasons can you give for thinking that a lack of agreement across groups might be troubling?
4 A philosopher might respond: all that matters is the intuitions of the experts (in this case, the philosophers), the intuitions of other groups doesn't matter. What do you make of such a claim?
5 Suppose that intuitions are like quick, compressed thought processes, giving us quick and easy ways to navigate the world around us. Might that support the idea that the intuitions of the experts matter more? After all, they've acquired lots of expertise in the area so their intuitions are more likely to be refined.
6 What do you think that an intuition *is*?
7 Suppose that everyone in philosophy had the same intuitions about all philosophical cases. Would that show that they were right, or that people who are good at philosophy are able to reproduce the intuitions of people already in the discipline?
8 Why might the answer to (7) matter?
9 Do scientists use their intuition? Does it matter?
10 How worrying do you think philosophers should find this varied empirical data?

Reading

Brown, J. 2013. 'Intuitions, Evidence and Hopefulness', *Synthese*, 190(12), 2021–46.
Nichols, S., Stich, S. and Weinberg, J. M. 2003. 'Meta-Skepticism: Meditations in Ethno-Epistemology', in S. Luper (ed.), *The Skeptics*, Aldershot: Ashgate, 227–47.
Weinberg, J. 2007. 'On How to Challenge Intuitions Empirically Without Risking Skepticism', *Midwest Studies in Philosophy*, 31(1), 318–43.

---

5 Foundationalism

Likely, as you read this, you think that the United States has a president. You believe this *for a reason*. If pressed, you could justify your belief to someone else. Most of our beliefs seem like that. I think that I'm writing this right now. I think this *because* I can see my hands moving at the keyboard, feel the keys under my fingers, and see the words forming on the screen in front of me.

It can also seem as if justification is essential to knowledge. Suppose that (by sheer fluke) my 4-year-old daughter says 'the Earth goes around the Sun'. She knows nothing about the solar system. She knows nothing about orbits. This really is a belief that she has formed for no good reason at all. To fill out the details, perhaps she's just stringing together words and, on this occasion, got lucky. Many folks would be uncomfortable in attributing knowledge to her, in the absence of her having a justification for her belief. It would certainly seem strange to say, in this case, that she *knows*.

But, as we'll see in the following thought experiment, there seem to be reasons to think that *not all* of our beliefs can be justified:

Thought experiment

> Mark is talking with his daughter, Esme. 'Time to go, Esme', says Mark. 'Why?', asks Esme. 'Because I can see the hands of the clock and I know how long it will take us to get there.' 'How?' 'Look, Esme, I can see the clock!' 'But dad, why is seeing the clock good evidence about the time?' 'Because, Esme, it's a clock! It keeps time.' 'Sure, dad, but why think it is now?'
>
> The conversation goes on for some time. Every time Mark provides some justification for his beliefs, Esme presses with another *why* question. In the end, Mark opts for a response familiar to many parents: '*Because*, Esme! Now get your coat!' In effect, Mark reaches a belief for which he can provide no further justification.

Though a little everyday, the case is somewhat illuminating. Begin with a belief, b. That belief needs justification or (presumably) we shouldn't hold it. So we offer that justification, j. But why hold that j is right? Again, if we don't think that we have good reasons to think j right, presumably we shouldn't hold it. So we offer another justification, j*. But why hold that j* is right? And so off we go on an infinite regress.

One very natural response to this is to respond, as Mark does, by saying that there is bedrock. There are some beliefs that we have (or can come to have) that are epistemically basic. These are beliefs for which we need no further justification – often called *properly basic beliefs*. A properly basic belief will be something that we can believe without further justification, though it is itself still justified, and so can be the base upon which all of our other justification snakes upwards.

The other response that we might offer – similarly obvious – is to simply *deny* that justification is required in some cases. But how comfortable are we with that? After all, we saw, above, that in the case involving my daughter I'd be really reluctant to say that she knows something in a case where she has no justification for her belief.

Questions

1. Assuming that there are properly basic beliefs, what kinds of belief do you think might be good candidates for being properly basic?
2. Do you think that a basic belief could be false?
3. How plausible do you find the idea that some of our beliefs might be properly basic?
4. Do you think that it's true that, if a belief isn't justified then (normally) we shouldn't hold it? Try to give examples in your answer.
5. What do you think is the difference between an infinite regress that is benign and an infinite regress that is vicious?
6. Do you think that the justification regress described above is vicious? Explain your answer.
7. If no beliefs are properly basic, do you think that we should say that any of our beliefs are justified?

8 Above, we talked about individual beliefs justifying others. Might there be another way to go? For instance, is there a way of using particular rules to justify some of our beliefs?
9 Would rules help deal with the justification regress?
10 Does it matter whether we can solve the problem of justification regress?

Reading

Howard-Snyder, D. and Coffman, E. J. 2006. 'Three Arguments Against Foundationalism: Arbitrariness, Epistemic Regress, and Existential Support', *Canadian Journal of Philosophy*, 36(4), 535–64.

Klein, Peter. 2004. 'What Is Wrong with Foundationalism Is That It Cannot Solve the Epistemic Regress Problem,' *Philosophy and Phenomenological Research*, 68(1), 166–71.

McGrew, T. 2003. 'A Defense of Classical Foundationalism', in Louis Pojman (ed.), *The Theory of Knowledge*, Belmont, CA: Wadsworth, 194–206.

Pryor, J. 2001. 'Highlights of Recent Epistemology', *The British Journal for the Philosophy of Science*, 52(1), 95–124.

---

6 Coherentism

In Chapter 5, we looked at foundationalist theories of justification. But, although we didn't explicitly focus on this issue in the last section, there might be a very general kind of worry with a foundationalist approach. In particular, we might worry that the foundationalist position doesn't reflect our epistemic practices to us.

Recall that, for the foundationalist, justification is a matter of drilling down towards a base – a base constituted by basic beliefs that do not, themselves, require further justification. But, says our objector, we *never* engage in the process of justifying our beliefs in this way (except, perhaps, when dealing with small children or philosophers, when 'why?' gets repeated over and over). Rather, when wanting to justify our beliefs we look for another feature, as brought out by the following thought experiment:

Thought experiment

> Ruth believes a number of propositions: a, b, c, d, e and f. For our purposes, it doesn't matter what these are. She also believes that today is Tuesday. Her friend, Helen, then tells her that today is, in fact, Wednesday. Helen reflects momentarily. She then calls to mind various memories that she has of the week so far. She remembers starting off on Monday with a day off work. Then, aha, she remembers that she went into work on Tuesday. Today must be Wednesday. She'd forgotten that she'd had a day off at the start of the week and had confused Monday and Sunday! Consequently, Ruth revises her belief about what day it is today. Having forgotten that she'd had a day off on Monday she updates her beliefs to now include the belief that today is Wednesday.

What is striking about this case, we might think, is that Ruth opens with a set of beliefs that are coherent. She seems perfectly justified in holding the beliefs that she does. But then Helen tells her something new and crucially it's something that does not cohere with her belief set. Ruth then inspects her beliefs and memories and recalls something that she's forgotten. She now has an inconsistent belief set.

It is precisely because this set of beliefs is inconsistent, we might say, that Ruth is then not justified in holding on to it. Something must go. And so Ruth does what any reasonable person would. She updates her belief as to what the day is today in order that her set of beliefs forms a coherent whole. This, roughly, is the coherentist position. Beliefs are justified just to the extent that they cohere with the other beliefs that we hold, and not justified to the extent that they do not cohere with our other beliefs.

One reason to find coherentism attractive is that it does seem to mirror the kind of epistemic practice that I've just described, where we aim to have a belief set that coheres and find beliefs that do not so cohere epistemically problematic.

Questions

1 Write down a proposition that you believe to be true. Explain how we would justify this belief, if we are coherentists, using specific examples.
2 Suppose that I hold two beliefs, 'a' and 'b', alongside a number of others. Given coherentism, do you think it's right to say that 'a' plays a part in justifying my belief in 'b' (and vice versa)?
3 Given your answer to (2), does this pose a problem for coherentism?
4 Would it be consistent with coherentism for Ruth to simply ignore what Helen says and still be justified in her belief? Is this a problem for coherentism?
5 Would you make any modifications to coherentism? Explain your answer.
6 Do you think that other disciplines (perhaps the sciences) will ever deliver an account of what it is to be justified in believing a proposition?
7 Why does justification matter to us?
8 It seems as if I am justified in believing what I'm seeing. Does this have anything to do with coherence or is there another form of justification in play?
9 It seems consistent with most people's beliefs both that the Universe is infinite in extent and also that it is finite. Thus, given coherence, we might think that two different and contradictory beliefs are justifiable. Does that seem right?
10 'Coherence' has been presented as requiring merely that a given belief not contradict another within the believer's belief set. Do you think that's the right kind of condition? Do you think that we might want something stronger, for justification, than merely that a belief not contradict any of our others?

Reading

BonJour, L. 1999. 'The Dialectic of Foundationalism and Coherentism', in J. Greco and E. Sosa (eds), *The Blackwell Guide to Epistemology*, Malden, MA: Blackwell, 117–42.

Cross, C. B. 1999. 'Coherence and Truth Conducive Justification', *Analysis*, 59(263), 186–93.

Lehrer, K. 1997. 'Justification, Coherence and Knowledge', *Erkenntnis*, 50(2–3), 243–58.
Olsson, E. J. 2002. 'What is the Problem of Coherence and Truth?', *The Journal of Philosophy*, 99(5), 246–72.

---

7 Analysing knowledge

I know some things – not as many as I would like to. You all know things. Some of you may know more than I. We're very familiar with the concept of knowledge. It's more than mere opinion. I can only know things that are true. Further, if I know something to be the case, then I'm not just guessing that it's true. If I know some proposition to be true, then it seems reasonable to think that I must have some good reason for believing that proposition to be the case.

Putting these ideas together, it seems to make sense to suppose that knowledge involves having true beliefs. It also seems that knowledge involves having justified beliefs. Thus, we might think, we can give an analysis of what knowledge is: knowledge *is* justified true belief. If I know that there is a glass on the table, then I have a true belief that the glass is on the table, and that belief is justified by (say) my perceptual experience of there being a glass on the table.

But there are some problems with this analysis, as shown by the following case (adapted from a problem due to Gettier):

Thought experiment

> Samantha drives a blue Ford. Bert knows this. As he's approaching Samantha's house, Bert sees a blue Ford being towed away. It seems to Bert as if the car came from Samantha's house. Bert tells Samantha that her car is being towed. She forms the belief that it is. She has justification for doing so: Bert is a reliable witness who knows what her car looks like. As it happens, the car that Bert saw wasn't Samantha's. However, her car was in fact being towed away, though it was towed in the opposite direction from the one Bert approached from, so he couldn't see it. Samantha now has a justified true belief that her car has been towed.

But this does not seem to be a case of knowledge. The problem seems to be that the belief that has been formed is, in some way, 'lucky'. It seems that Samantha has formed her true belief without consulting the kinds of source that, in this particular instance, would have appropriately connected her with the object of her belief.

To illustrate this idea, it's worth reflecting upon what would have happened if Samantha's car hadn't been towed. Presumably, Bert would still have seen the other car being towed, and would still have told Samantha. That being the case, Samantha would still have formed the belief that her car was being towed. It seems somewhat accidental that her belief turned out to be true in this case. Thus, it seems that there's an important sense in which Samantha's belief was formed by the wrong kind of process.

The difficulty, of course, comes in specifying quite what that process would be. For instance, it's tempting to say something like the following. X knows that p if and only if X non-accidentally believes that p, and p is true. This rules out the Samantha and Bert case as an instance of knowledge because it seems like Samantha's belief is only true by some kind of accident. But we must then say what 'non-accidentally' means here. The most obvious thing to say is that 'non-accidentally' means 'believes in the right sort of way to yield knowledge'.

So, to put all of that together: X knows that p if and only if X believes in the right sort of way to yield knowledge of p, and p is true.

But that can't be a good definition, because the definition makes use of the term 'knowledge' in trying to describe what knowledge is. This is, obviously, circular. If the analysis of knowledge is circular, then it seems that we do not really have a proper analysis at all.

Questions

1 Explain, in your own words, why the thought experiment case involves Samantha's having a justified, true belief. (The experiment is an example of a 'Gettier case'.)
2 Explain, again in your own words, why this isn't an instance of knowledge.
3 How would you describe the notion of 'justification'?
4 Historically, a number of people would have said that they knew that the Earth was flat. Using the analysis where knowledge is treated as justified true belief, explain why they were wrong.
5 We said, above, that this is a case where Samantha has a justified true belief, but where that belief seems to be true only by accident. Can you think up your own case with this structure?
6 It's typical to try to respond to this kind of case by offering a third condition that, in addition to a belief's being true and justified, allows us to complete the analysis. Can you think of a condition that we might add and that would help?
7 Suppose that the case had been different, and that if Samantha's car had not been towed, then Bert would have realized his error and then would have reported to Samantha that her car had not been towed: would Samantha then know that her car had not been towed?
8 Suppose that Samantha knew Bert to be drunk. Would you say, then, that she would have had a justified, true belief?
9 Does it matter if we can't provide an analysis of what knowledge is? Why?
10 Do you think that everyone would agree that the case involving Samantha is one where she doesn't know that her car is being towed? What are the consequences of your answer for our thoughts about knowledge?

Reading

Gettier, E. 1963. 'Is Justified True Belief Knowledge?', *Analysis*, 23(6), 121–3.
Morton, A. 2002. *A Guide through the Theory of Knowledge*, Hoboken, NJ: Wiley, chapter 1.

Sosa, E. and Kim, J. (eds) 2008. *Epistemology: An Anthology*, Oxford: Blackwell, section II.
Zagzebski, L. 1999. 'What Is Knowledge?', in J. Grecco and E. Sosa (eds), *The Blackwell Guide to Epistemology*, Malden, MA: Blackwell, 92–116.

---

8 Perception

Perceiving objects is very natural. Simply look around you and you will see a great many things. But it is also quite complex. Neurologically, we know that the brain processes required are tremendously involved. Philosophically, things are quite complicated as well. It would seem to be a natural starting point to suppose that if I see an object (an orange) in front of me, then I *see that thing*. It is the *orange* that is the entity with which I am immediately acquainted through my perceptual experience. There is nothing intermediary. However, the following thought experiment looks like it might make problems for this view:

Thought experiment

> Archie is trekking across the desert. After many days, he has run out of water. He is tired, thirsty and hungry. As more days pass, still without water, Archie begins to hallucinate. It seems to him as if an oasis has appeared on the horizon. Joyously, he rushes towards it. Upon reaching the point where it seemed to him that the oasis was located, he finds nothing but sand. But notice: the perceptual experience that Archie had was precisely the same (from Archie's point of view) as a *perception* of an oasis. The two are not distinguishable. (Were they distinguishable, then presumably Archie would not have rushed towards this particular spot in the desert.)

Some people moved by this kind of concern will then argue that what this case shows is that perceptual experience (veridical or otherwise) does not involve a direct and immediate experience of the object *out there in the world*. Rather, it shows us that the objects we are directly and immediately acquainted with via experience are what they call *sense data*. These are something like mental objects, or properties, that we are aware of, and that are located in our mind/brain. Thus, we do not see objects in the world *directly*, but only as they are mediated to us through the sense data.

There are a variety of responses. In particular, there are two theories that oppose sense data theories and that will concern us here. Both theories are (at least in part) motivated by concerns with sense data theories. What *are* sense data? Where do they reside? Are they physical or mental?

- The first alternative to sense-data theory that we will look at is adverbialism. According to this view, to say that I perceive something to be orange is to say that I experience *orangely*; that is, *orangely* is a way that my sensory experiences can be.

- The second alternative to sense-data theory that we will consider is disjunctivism. According to this view, *either* we are perceiving directly, or else we are undergoing a distinct kind of mental process. Crucially, the disjunctivist will say that Archie's experience of the hallucination *will not* be the same as the experience had by someone who is *perceiving* an oasis.

Questions

1. Suppose that Clare perceives an orange. Write down how each of the four theories described will describe the act of perception.
2. In addition to the case involving hallucination, sense data theorists sometimes appeal to cases involving time lag. To illustrate: look up at the night sky. You see stars. Some of those stars do not exist. So, how can perception be *directly of* the object being perceived? Such objects don't exist.
3. Another sense data argument involves difference of perspective. Suppose we perceive a car, close up. It looks one way. Now suppose that we move away from the car. It no longer looks the same. But the car has stayed the same. So what I am acquainted with in my experience cannot be the *car*, for the experience of it is no longer the same.
4. What are sense data? How plausible do you find the idea that these things exist?
5. Do you think that sense data would commit us to dualism about the mind?
6. Consider adverbialism: how would you describe the experience of seeing an orange? Now, how would you describe the experience of seeing a green square and a brown triangle? Can you see any problems with your description?
7. The disjunctivist says that a perceptual experience is not like the experience of an illusion or hallucination. Does that strike you as plausible?
8. If forced to pick one, which theory would you adopt? Why?
9. Do you think that we should wait for the science of perception to improve before trying to answer this question? Why? In detail: what kind of information do you think might be forthcoming?
10. Do you think that it matters whether we can give an account of perception?

Reading

Lowe, E. J. 1981. 'Indirect Perception and Sense Data', *Philosophical Quarterly*, 31(125), 33–42.

Macpherson, F. 2013. 'The Philosophy and Psychology of Hallucination: An Introduction', in F. Macpherson and D. Platchais (eds), *Hallucination*, Cambridge, MA: MIT Press, 1–38.

Shoemaker, S. 1990. 'Qualities and Qualia: What's in the Mind?', *Philosophy and Phenomenological Research*, 50 (supplement), 109–31.

Tye, M. 1984. 'The Adverbial Approach to Visual Experience', *The Philosophical Review*, 93(2), 195–225.

9 Scepticism

I know that there's a cup on the desk in front of me. I can see it. If you were to ask someone whether or not I knew that there was a cup in front of me, she'd say that I did. But the following idea seems pretty plausible. If I know that there's a cup on the desk in front of me, then I know that there's not just a picture of a cup on the desk, about which I'm mistaken. Indeed, if I *know* that there's a cup on the desk in front of me, then it seems that I know that a lot of other things don't obtain: it's not that there's just a picture in front of me; it's not that there are some cleverly arranged mirrors on the desk; it's not that there are some fancy holograms generating the mere appearance of a cup. And, indeed, I can check that this is so. Simply by touching the cup I rule out all of these possibilities.

But notice: as in the case just described, it seems intuitive to say that if I know some proposition (for instance, about the cup), then I know that none of the illusory cases described are present. I can be confident that they're not the case. I can rule them out. If I didn't know that all of those illusions were ruled out, then I wouldn't know that the cup was in front of me. Roughly, the idea seems to be this: if I know p, and q entails not-p, then I know that not-q is the case. (And so if I don't know not-q, then I don't know p.) This is something that the epistemic sceptic, who will claim that we can never know anything, can try to exploit.

Thought experiment

> In the film *The Matrix*, the main character, Neo, discovers that he's lived his whole life plugged into a computer simulation. If you'd asked him, within the simulation, and at the point the computer simulated the experience in him of walking down the road, whether or not he knew that he was walking down the road, Neo would say that he did. Similarly, at the point at which the computer simulated the experience in him of eating a nice steak, Neo would have said that he knew that he knew that he was eating the steak. He was wrong, of course.
>
> But now think about yourself. Can you rule out the possibility of your being a brain in a vat that's being programmed to have particular experiences? If you can't rule out the possibility, then you don't know that you're *not* a brain in a vat. And if you don't know that you're not a brain in a vat, then it seems, just as in the case of my cup that we started with, that you don't know that you're really seeing what you're seeing, or hearing what you're hearing.

What matters to this setup is whether certain states are possible. If I *know* that I'm not a brain in a vat, then it's *not possible* that I'm a brain in a vat (I can only know truths, after all). So if it's *possible* that I am a brain in a vat, then I don't know that I'm not. And since it is *possible*, I don't know that I'm not a brain in a vat.

Of course, if I know that I'm sitting here writing this, then I know that I'm not a brain in a vat. But since we've just said that I don't know that I'm not a brain in a

vat, I don't know that I'm writing this. Indeed, it seems that I will know very little, for most of my putative knowledge is incompatible with my being a brain in a vat. Accordingly, if the ideas in the thought experiment are right, I don't know that I'm writing this and you don't know that you're reading it!

Questions

1 The sceptic looks to deny that we have much knowledge. What is so important about knowledge? How might it be importantly different from mere belief?
2 What do you think about the case involving Neo? Is such technological sophistication possible for us? Does it matter if it's not?
3 Recall that: if I know p, and q entails not-p, then I know that not-q is the case. Explain why it follows that if I don't know not-q, that I don't know p.
4 An early forebear of the 'brain in a vat' case is Descartes' evil demon. Descartes thinks it's *possible* that he might be deceived by an all-powerful evil demon, capable of giving him completely non-veridical experiences. Are there any differences between this case and the 'brain in a vat' case?
5 Suppose that you might be a brain in a vat: is there *any* knowledge that you would retain?
6 Quite a bit of the sceptics' argument follows from the principle: if I know p, and q entails not-p, then I know that not-q is the case. How plausible would it be to reject the principle?
7 Consider the following line of argument, very similar to one offered by G. E. Moore. If I know that I have hands, then I know that I'm not a brain in a vat. I do, in fact, know that I have hands. Thus, I know that I'm not a brain in a vat. What do you think of this argument?
8 Sceptics argue that we know nothing (or very little). Do you think that it makes a practical difference whether they're right? If not, is that a problem?
9 Suppose that you go along with the sceptics' argument: not because you think it persuasive, but because you can't find a way to block it. You concede that you know nothing (or very little). You then go out into the world and claim to know many things. Have you contradicted yourself? If so, isn't that a problem? After all, we normally try to avoid falling into contradiction.
10 How do you think that we should respond to the sceptic?

Reading

Frances, B. 2005. 'When a Skeptical Hypothesis Is Live', *Noûs*, 39(4), 559–95.
Stroud, B. 1984. *The Significance of Philosophical Scepticism*, Oxford: Clarendon Press.
Unger, P. 1975. *Ignorance: A Case for Scepticism*, Oxford: Oxford University Press.
Williams, M. 1999. 'Scepticism', in J. Greco and E. Sosa (eds), *The Blackwell Guide To Epistemology*, Malden, MA: Blackwell, 35–69.

10 Trust

Trust features centrally in all of our lives. You trust your friends and family to support you and to look after you. You can trust your friends to keep a secret. *You* will be trusted *by* people to not let them down; to keep their secrets. Nonetheless, the notion of trust seems quite hard to pin down precisely. What *is* trust? And how does it mesh with our other epistemic concepts?

Perhaps a little surprisingly, the notion of trust has only come to the fore quite recently in epistemology. A significant part of that work has focused upon two related tasks. First, giving an adequate account of what trust is. Second, differentiating trust from other, closely related, concepts. As the thought experiment demonstrates, that can be a challenging task:

Thought experiment

> Sasha and Mark, a couple, are discussing Mark's future. Mark needs Sasha's help. He's planning on taking some time away from work. He will need financial support from Sasha in order to keep paying some of his debts. They have, coincidentally, just put up some new shelves on the wall of their flat and put some ornaments onto the shelves. The conversation has been going for some time. As a couple they haven't always dealt with money very well.
>
> At the end of the conversation, they agree that Mark will take the time away from work and that Sasha will pay more of the household bills and will also ensure that Mark has some money for himself. Mark is very grateful, but a bit concerned about the arrangement. Sasha says: 'it's OK, Mark, you can trust me – just like you can trust that shelf to hold up our ornaments.'

Clearly, there is a sense in which what Sasha says is very reasonable. We can trust people (and she is pretty trustworthy). And *in a sense* we can trust shelves to do their jobs. But in another sense, the sentence doesn't sound quite right. It seems altogether more natural to say that shelves (and other inanimate objects) can be *relied upon* to do their jobs, rather than trusted. At least, there is a sense of the word 'trust' that seems to work in that way; to apply to people and our relationship with them. It is that sense of trust that we're after.

Of course, trust also seems to involve reliance. It would be odd to say that Mark trusts Sasha but does not think that he can rely upon her. This leads to the idea that we might think of trust as being reliance *plus* something else, s, such that when we have reliance plus *s* we have trust. For some, s is appropriately understood as something like a positive view of the motivations of the trusted person.

Other approaches to the concept of trust take a slightly different route. For instance, we might think of trust as involving believing that someone has a commitment to perform some task and also relying upon them to perform it (see Hawley, 2014). Since both of these accounts just sketched make reference to

people, it turns out that they deliver the result that we cannot trust inanimate objects (like shelves), but that we can trust people. This seems right. Though of course, whether either of the analyses suggested will succeed is quite an involved and difficult question.

Questions

1 Do you think that it's right to draw a distinction between the way in which we trust a person and the way in which we trust (for instance) a shelf? Why?
2 We normally think of trust as a positive thing. Why is that? Do either of the accounts just sketched help us to understand why?
3 Using the two accounts of trust sketched, explain why it's true that Mark trusts Sasha.
4 Using the two accounts of trust sketched, make up your own case involving trust and explain how the analysis delivers the correct result.
5 Can you think of a counterexample to either account? That is, a case where we would trust but in which we would not satisfy these conditions, or else a case where we would satisfy these conditions, but not trust?
6 Which of the proposed accounts do you prefer, and why?
7 Is distrust the mere absence of trust? If not, what is distrust?
8 Do you think that we could trust a very advanced computer or robot? What implications (if any) does your answer have for the accounts of trust sketched?
9 It's not uncommon to hear discussion in industry and commerce of customers 'trusting a brand' or a particular company. Do you think that the accounts of trust sketched are getting at the right sort of trust? Why?
10 Do you think that a philosophical analysis of trust might be useful to people working in commerce and industry? Why?

Reading

Hawley, K. (2014). 'Trust, Distrust and Commitment', *Noûs*, 48(1), 1–20.

Holton, R. 1994. 'Deciding to Trust, Coming to Believe', *Australasian Journal of Philosophy*, 72(1), 63–76.

Jones, K. 1996. 'Trust as an Affective Attitude', *Ethics*, 107(1), 4–25.

Marsh, G. 2011. 'Trust, Testimony and Prejudice in the Credibility Economy', *Hypatia*, 26(2), 280–93.

Part II

PHILOSOPHY OF LANGUAGE

11 Frege and names

We give names to a lot of objects and a lot of types of object. Both of the authors of this book work at a university with the name 'the University of Nottingham'. The present queen of England has the name 'Elizabeth'. The current president of the USA is called 'Obama'. These facts are very obvious and straightforward. What is less clear is quite how names work. We know that names, in a sense, *stand for* particular objects: they 'point' to objects. But how does that 'standing for' or 'pointing' work? The name 'Elizabeth' successfully picks out the queen of England: but how exactly does it do that?

One very natural starting point is that names directly refer – just literally *stand for* – objects. In slightly more formal language, we might say that names *denote* individuals. Thus, 'Elizabeth' *denotes* (stands for) the queen of England. However, as Gottlob Frege pointed out, this doesn't seem to be all that there is to naming. Consider the following:

Thought experiment

> Suppose that names do just *denote* or *stand for* particular objects. Consider, then, identity statements. The claim that 'Jonathan is identical to Jonathan' is trivial and uninformative. Nothing is learned. And if all that names do is *denote* then this should be the case in all instances of naming. After all, *all* that a name does is stand for an object.
>
> However, we learned a little while ago that, for some time, the famous author J. K. Rowling had been writing under a pseudonym: Robert Galbraith. In other words, we learned that 'J. K. Rowling is identical to Robert Galbraith'. If *all* that names did was to stand for objects, then we should not have learned anything by being told that J. K. Rowling is identical to Robert Galbraith. But we did. And so it must be the case that names do more than *just* stand for objects.

The case bears a little unpacking. On the view that names *just* denote, names are being taken to be a little like metaphorical pointers: they point to objects in the world. If 'J. K. Rowling' is a pointer, and 'Robert Galbraith' is a pointer, then to use those names properly we must understand what they point at – what the names denote. But since both names denote the same object, once we know how to use both names, we should just grasp that both names pick out the same object. And, of course, we didn't. It came as something of a shock to the literary world when Galbraith's identity was revealed!

Since Frege was writing in the nineteenth century, he didn't have this case in mind. But other cases like it motivated him to suggest that names must do more than merely denote. Names must also have some other feature. The feature that Frege thought important was that of 'sense' or 'mode of presentation'. The basic idea is that, in the above case, the name 'J. K. Rowling' does more than just denote: it also

connotes. In this case, the *sense* of the name includes the following features: author of the Harry Potter series; wealthy; gives significant funds to charity, etc. In contrast, the name 'Robert Galbraith' doesn't have that sense. It would seem that the sense of the name 'Robert Galbraith' would include: author of *The Cuckoo's Calling*; relatively obscure author; etc. Thus, in this particular case, we can learn something from the identity statement 'J. K. Rowling is identical to Robert Galbraith.' What we learn is that: the famous, wealthy author of the Harry Potter series, who gives lots of money to charity, is identical to the relatively obscure author of *The Cuckoo's Calling*.

Thus, Frege thought, we have a motivation to think that names have senses as well as playing the role of denoting particular objects.

Questions

1 Imagine that Superman and Clark Kent are real. What would be the sense of the names 'Superman' and 'Clark Kent'?
2 Lois Lane does not believe that Clark Kent and Superman are the same person (at some points). Then she learns that they are. Using the senses that you've described in answer to (1), write down what it is that Lois learns, if Frege was right.
3 Using 'sense', explain why the identity statement 'a is identical to a' is uninformative.
4 Call to mind the theory that names *just* denote (that sense is not involved). Consider the following two beliefs: 'John believes that J. K.Rowling is rich'; 'John believes that Robert Galbraith is poor'. Can John consistently believe both, if the denotation-only theory is true? Explain your answer.
5 Can John consistently believe both, if names have senses as well as denoting? Explain your answer.
6 In the discussion, 'senses' were described as 'modes of presentation'; based on what we've said so far, how would you describe what sense is?
7 What connection do you think there is, if any, between the sense of a name and what the name denotes?
8 Do you think that there is a difference between how we think of an object and the sense of the name that refers to it? For instance, I have particular thoughts about J. K. Rowling. Are my thoughts about her identical to the sense of the name 'J. K. Rowling'?
9 To properly understand how to use a name, do you think that we have to have perceived the object that the name denotes, fully grasped the sense, both, or neither?
10 Can you think of any problems for Frege's view?

Reading

Dummett, M. 1973. *Frege: Philosophy of Language*, London: Duckworth, chapter 5.
Frege, G. 1892. 'On Sense and Reference', reprinted in A. W. Moore (ed.), *Meaning and Reference*, Oxford: Oxford University Press, 23–42.

Luntley, M. 1984. 'The Sense of a Name', *Philosophical Quarterly*, 34(136), 265–82.
McDowell, J. 1977. 'On the Sense and Reference of a Proper Name', *Mind*, 86(342), 159–85.

---

12 Russell and definite descriptions

As noted elsewhere in this book, it seems reasonable to think that the primary function of names is to refer to, or stand for, particular objects (see, for example, Chapter 11). One (loose) way of thinking about this is to say that what names contribute to a sentence is the object that they refer to. Thus, when I use the name 'David Cameron' the name is making a specific contribution to the sentence: the name is standing for David Cameron himself.

Now, for reasons touched on in other chapters (such as Chapter 11), we may want to augment this approach by saying that names *also* make other contributions. But there are a range of cases that seem to suggest that names *can't even* play that role. That names *can't* stand for particular objects. These are cases of 'bearerless names'. Consider the following:

Thought experiment

> To grasp the meaning of a name, it seems that we have to grasp what it stands for. Consider the name, 'The present queen of France'. Whilst there is no object, it nonetheless seems that we can grasp the meaning of the expression, for we can understand the sentence, 'The present queen of France is bald.' If we can understand that sentence, then it cannot be the case that the name 'The present queen of France' is a name that we cannot grasp. Thus, 'The present queen of France' must *stand* for something. But it does not. It is a name without a bearer.

Now, at this point, we might be tempted to say that the contribution made by the name here is *just* that it expresses the sense of the name 'The present queen of France', and that we need not grasp what the name picks out, or stands for. But there is a problem. Suppose we have two sentences:

1 The present queen of England is bald.
2 The present queen of France is bald.

England, but not France, has a monarchy. So, 'the present queen of England' denotes; 'the present queen of France' does not.

If two expressions have the same kind of grammatical function (in this case, if two terms are both names), we should give them the same semantic treatment. It seems unfeasibly *ad hoc* to think that *sometimes* names function to pick out a denoted object; other times they pick out a sense. For one thing, it would make

learning how to use names extraordinarily complex. Thus, what the problem of bearerless names may be taken to show is that names can't stand for objects.

In place of treating names as standing for objects, Bertrand Russell proposed a universal treatment of names. Names are taken to be *definite descriptions*. In order to appreciate Russell's own account of naming, we need a little bit of terminology; we need the distinction between definite and indefinite descriptions. To see what the distinction consists in, consider the following two sentences:

1 The lecturer is bald.

Notice that, in (1), the referring term serves to pick out a *specific* lecturer.

2 A lecturer is bald.

By way of contrast, (2) does not serve to pick out a specific lecturer.

Now that's interesting; names seem to function just like the referring term 'the lecturer'. That is, names seem to function in such a way to pick out an individual.

On the face of it, that seems OK but a little tangential. Names aren't descriptions like (1). 'Jonathan Tallant' doesn't describe anything. It's a name. Names aren't descriptions. Russell obviously had to concede this point. However, he responded with the claim that names are *disguised* definite descriptions. Thus, the name 'Jonathan Tallant' in fact is a *disguised* way of using a definite description to pick out an individual (me). The definite description involved will probably take the form: 'philosophy lecturer; has brown hair; has a cat; about 6ft tall; Sagittarius . . . etc.'

Questions

1 If names are disguised definite descriptions, then work out what disguised definite description your name is.
2 How would you describe the difference between a definite and an indefinite description?
3 Building on the chapter on Frege, explain, using Russell's account, why true identity statements (e.g. 'a = a') are, or can be, informative.
4 What definite description does 'the present queen of France' stand for? Is the resulting description true or false?
5 Building on your answer to (4): is 'the present queen of France isn't bald' true or false? Why does (or might) this matter?
6 Call to mind the theory that names *just* denote. Consider the following two beliefs: 'John believes that J. K. Rowling is rich'; 'John believes that Robert Galbraith is poor'. Can John consistently believe both if names are definite descriptions? Explain your answer.
7 Suppose I think of J. K. Rowling as 'rich author who wrote the Harry Potter series', and you think of her as 'person born on 31 July 1965 who was named

in 2007 as a runner up in the *Time* magazine person of the year'. Both of these are definite descriptions. Does the fact we think about her differently pose any problems for this theory of naming?

8 If this theory is right, then at the point of introducing a new name, must it be the case that a new description is brought along with it? Is your answer plausible given the way that names are introduced in real life?

9 Most people change an awful lot over the length of their lifetime. Does this change pose any problem for a theory that treats names as disguised definite descriptions?

10 What, if anything, is the difference between Russell's view and Frege's view?

Reading

deRossett, L. 2011. 'Reference and Response', *Australasian Journal of Philosophy*, 89(1), 19–36.

Donnellan, K. 1966. 'Reference and Definite Descriptions', *The Philosophical Review*, 75(3), 281–304.

Russell, B. 1905. 'On Denoting', *Mind*, 14(56), 479–93.

Strawson, P. F. 1950. 'On Referring', *Mind*, 59(235), 320–44.

13 Kripke on names, rigid designators and a causal theory

The thesis that names have some descriptive component to them is quite powerful. As is noted in other chapters (e.g. Chapter 12), Russell even thought that we could have names to *be* disguised definite descriptions. But even moving away from that, the connection between names and descriptions is vivid. Think, for instance, about the name of the town 'Exmouth'. The town is so-named because it sits at the mouth of the river Exe. Indeed, the name 'Exmouth' might seem to *be* a description as well as a name.

But in spite of the fact that there might seem to be a close connection between names and descriptions, there are some reasons to think that the connection is not in fact all that close.

Let's begin with a piece of terminology: 'modal profile'. When I talk about the 'modal profile of names and terms', I'm talking about how they work when we're thinking about what's possible and what's necessary. Talk about what's possible and what's necessary is just another part of the way in which we talk about objects – so our theory of naming ought to work perfectly well in such contexts.

So, let's borrow an example from Saul Kripke. Let's talk about and think about the name 'Hitler'. What description seems to be involved with that name? Let's start easy: Führer from 1934 to 1945; lived 20 April 1889–30 April 1945; author

of *Mein Kampf*; 'man who succeeded in having more Jews killed than anyone else managed to do in history' (Kripke 1980: 75).

Let's now focus on the last clause. In particular, I want us to focus upon the modal profiles of 'the man who succeeded in having more Jews killed than anyone else managed to do in history' and 'Hitler'. Do the two terms have the same modal profile?

Thought experiment

> Hitler, the man referred to by the name 'Hitler', might have led a very different life. Had his early forays into art been successful (he was twice rejected by the Vienna School of Fine Arts), then it seems entirely plausible that Hitler would not have been correctly described by the description 'the man who succeeded in having more Jews killed than anyone else managed to do in history'. Thus, in particular modal contexts (where we describe particular possible situations) the terms 'Hitler' and 'man who succeeded in having more Jews killed than anyone else managed to do in history' clearly pick out different things.

For this kind of reason, Kripke thought that names are what he called 'rigid designators': they pick out the same object in any possible world (see Chapter 77), where they refer to any existent at all. In contrast, descriptions only contingently pick out the objects that they describe. In that case, of course, names can't *be* descriptions, because the two things work in different ways in different contexts.

According to Kripke, the way in which the reference of a naming term is developed is via an 'initial baptism' (1980: 96). Thus, when a name is first introduced it is accompanied either by an ostensive act (something like pointing) or a description. Let's give examples. Suppose I introduce the term 'blug' and, as I speak the word, I point to the ceiling. You might reasonably infer that I mean the term 'blug' to refer to the ceiling. This would be an example of reference being fixed by ostension. It seems pretty plausible that very early naming occurred in this way.

Suppose, however, that I introduce a name: 'Verity'. I might say that 'Verity' is the name of my daughter. This is a case where I'm fixing the reference of the name with a *description*. Thus, Kripke's thought is that although names aren't identical to descriptions, descriptions can be used to fix reference at the point of baptism.

Once the term has been introduced and had its reference fixed, then it is passed from person to person via a causal chain. So, for instance, once the name 'Moses' was introduced, the name itself was then told by one person to another, and then another, and so on through a long causal process – a causal relation, if you will. There is, therefore, a causal relation between the initial baptism and subsequent uses of a term to refer to a particular individual originally baptized, because of the place they occupy in that causal chain. For this reason, we may say that in order to refer to a particular individual successfully it must be the case that the user of the

name must stand in a causal relation to the act of baptism. And this explains why we call this the causal theory of reference.

Questions

1 Write down a description that you satisfy. Explain why that description does not 'rigidly designate' you in all possible scenarios.
2 Explain why your name *is* a rigid designator.
3 Explain how Kripke might account for the name 'Exmouth' having come into being, if the name is not, in fact, a description.
4 Do you think that it's right to say that the only involvement that descriptions have with the naming process is that involved in baptism?
5 Suppose that aliens came down to Earth and referred to you by your name, despite never having been told of your name (or being made aware of it via other means). According to the causal theory, they have not referred to you (they don't stand in the right kind of causal relation to the baptism). Do you think that's right?
6 It seems possible that I could have been given a different name, 'Bob' let's say. Does this threaten Kripke's view that names rigidly designate? Explain your answer.
7 Can a name change, over time, with respect to what it refers to (such that it starts out referring to one object, then changes to refer to another)? If so, does this pose any problems, either to the claim that names rigidly designate, or to the causal theory of reference?
8 Suppose that an infant is named 'Sally', but that, due to an error in the hospital, the infant is swapped for another, and some other infant gets known by the name 'Sally', despite never having been baptized. How can the causal theory of reference deal with this?
9 Many species are given names: *Apis mellifera* is the name given to Honey Bees. If it's a name, it should rigidly designate. Does it? To work this out, begin by thinking about what the name picks out, and then try to decide whether or not it picks out the very same thing at all worlds.
10 The question of whether or not names and descriptions are the same seems, in Kripke's hands, to turn into a debate about how 'we intuitively' think about names and descriptions working. At no point has data on people's intuitions and whether they agree with Kripke been presented. Do you think this lack of data is a problem?

Reading

Evans, G. 1973. 'The Causal Theory of Names', *Proceedings of the Aristotelian Society*, supplementary volumes, 47, 187–208.
Kripke, S. 1980. *Naming and Necessity*, Oxford: Blackwell.

Sosa, D. 2001. 'Rigidity in the Scope of Russell's Theory', *Noûs*, 35(1), 1–38.
Stanley, J. 1997. 'Names and Rigid Designation', in C. Wright and B. Hale (eds), *A Companion to the Philosophy of Language*, Oxford: Blackwell, 555–85.

14 Quine on the analytic/synthetic distinction

There's a reasonably intuitive distinction to be drawn in language between two different sorts of claim that can be made. The first type of claim that will interest us is like this: all bachelors are unmarried men. The second type of claim that will interest us is like this: John is a bachelor. We will, following philosophical convention, call the first sort of claim 'analytic' and the second sort of claim 'synthetic'. The *very* rough idea is that an analytic truth is a claim that is true *in virtue of the meaning of the terms involved*. A true claim that is not analytic is taken to be synthetic. However, there's a problem with the distinction between the analytic and the synthetic, and, as we will see, it's a problem that threatens to do away with our notion of meaning altogether:

Thought experiment

> Sally thinks that the following principle is true: (P) If we cannot define what it is to be x, we should not believe that there are instances of x. But Sally also thinks we can't give a (non-circular) definition of what it is to be analytic. She worries as follows: x and y are synonymous if 'all xs are ys and all ys are xs' is an analytic truth. Now plainly, since we're trying to define what it is for two terms to be analytic, this isn't going to help us define what synonymy is.
>
> You might think this is just foolish; two terms are synonymous if they mean the same thing. But Sally then wants to know what it is for two terms to 'mean the same thing'. Will you say that two terms, x and y, mean the same thing if 'all xs are ys and all ys are xs' is an analytic truth?
>
> The problem here is coming to this: we can define sameness of meaning in terms of synonymy and analyticity; analyticity in terms of synonymy and sameness of meaning, and synonymy in terms of sameness of meaning and analyticity. But these definitions are all circular; we are defining (e.g.) synonymy in terms of analyticity; analyticity in terms of sameness of meaning, and sameness of meaning in terms of synonymy.
>
> In that case, of course, we're really defining synonymy in terms of synonymy and that is not an acceptable definition. We can't credibly maintain that 'all synonymous terms are synonymous with one another' is a non-circular definition of synonymy.

Now note: if we can't say what it is for two terms to 'mean the same' as one another, it is far from clear that we have any genuine concept of meaning at all.

Thus, from a worry about whether there is a distinction between the analytic and the synthetic, we find ourselves now worrying about meaning itself!

One very natural way in which one might try to respond is to talk about being able to swap terms, but doing so without affecting the truth of the sentence. To illustrate: we might think 'bachelor' and 'unmarried man' are synonymous *because* in any sentence in which the word 'bachelor' appears is one where we can swap 'bachelor' for 'unmarried man' and the truth – or falsity – of the sentence will stay the same. But there are a number of problems with this option.

Willard Van Orman Quine – who introduced something like the arguments presented as Sally's – thought that he had a sensible solution. According to Quine, there is no such sharp distinction to be drawn between the analytic and the synthetic. Rather, we should think of our beliefs as forming a web. Some of them – presumably roughly corresponding to the synthetic – will be towards the outer edge of the web. These beliefs are relatively easy to revise; we don't necessarily need much evidence to get us to change our minds about these. In contrast, the beliefs at the centre of the web – again, roughly corresponding to the analytic – will be those that are much harder to find reasons to revise.

Questions

1 Write out a list of some truths that, intuitively, would seem to you to be analytic, and some that seem to you to be synthetic.
2 Can you think of any way of spelling out that difference between the analytic and the synthetic in a non-circular fashion?
3 Consider the sentence: 'The word "bachelor" has eight letters.' Can you swap 'bachelor' for 'unmarried man' in that sentence, without altering the truth-value? What consequences are there – if any – for the analytic/synthetic distinction?
4 There seems to be an assumption in the argument presented that if we can't define something in a non-circular fashion we can't really grasp what we're trying to define. (Thus, because we can't describe the analytic/synthetic distinction, we should deny that we have any grip on that distinction.) How persuasive do you find that idea?
5 Were there any cases in your answer to (1) that you weren't sure about? Of what consequence (if any) is your answer to that, to your answer to (4)?
6 It's a consequence of Quine's 'web of beliefs' that *any* of our beliefs could – at least in principle – be revised, given appropriate evidence. What kind of evidence would we need in order for 'all unmarried men are bachelors' to be revised?
7 What do you think of the move from denying the distinction between the analytic and the synthetic, to declaring that there is no genuine concept of meaning?
8 You (hopefully!) understood Quine's argument. Quine's argument was expressed using language. If that language wasn't meaningful, then presumably you wouldn't have understood it. Does this threaten Quine's position?

9 Would mathematical truths be analytic or synthetic?
10 Do you think that there is a distinction between the analytic and the synthetic?

Reading

Boghossian, P. 1997. 'Analyticity', in C. Wright and B. Hale (eds), *A Companion to the Philosophy of Language*, Oxford: Blackwell, 334–68.
Grice, P. and Strawson, P. 1956. 'In Defense of a Dogma', *The Philosophical Review*, 65(2), 141–58.
Quine, W. V. 1951. 'Main Trends in Recent Philosophy: Two Dogmas of Empiricism', *The Philosophical Review*, 60(1), 20–43.
Wright, C. 1997. 'The Indeterminacy of Translation', in C. Wright and B. Hale (eds), *A Companion to the Philosophy of Language*, Oxford: Blackwell, 397–426.

15 Indeterminacy of translation

It seems quite apparent that I can teach my daughter how to use a word and explain to her what its meaning is. Of course, some cases are easier than others. But there are paradigmatic easy cases. For instance, if I want to teach her how to use words like 'table', 'chair' and so on, all that I need to do is to point to the object in question and then repeat the word in question. Eventually, she'll get there and grasp the meaning.

But there are some reasons to think that, even in these seemingly very easy cases, we should pause. For even in these easy cases there might be no straightforward way in which we can *be sure* that we're fixing on quite the right meaning and grasping the right way to use the term. To help illustrate this point, here's a thought experiment that we'll borrow (at least in outline) from Quine:

Thought experiment

> One day, an explorer meets a tribe of people. They speak a language unique to their tribe. As an explorer, you wish to befriend these people, to help understand their local environment and better appreciate their place in the local ecosystem. Suddenly, away to your right, you spot a rabbit dash from a bush. At the same moment, the tribespeople yell out, in unison, 'Gavagai!' You assume that 'gavagai' means 'rabbit'. To check, you point to the rabbit and say 'gavagai?' The tribespeople respond, seemingly positively, and say 'gavagai'.

Given the thought experiment, it would be very natural to suppose that the suggested translation is the right one; that 'gavagai' means 'rabbit'. Indeed, at the outset it would seem very odd to say that 'gavagai' means anything else.

But there's a worry. Importing just a few background assumptions, it's easy to see that the term *could* mean 'creature whose foot is lucky', or 'long-eared mammal'.

There is no *reason* to prefer one translation to the other. Each of these proposed options is consistent with the evidence that we have. Indeed, if we were really determined to force the issue, we might suppose that 'gavagai' means 'undetached rabbit part'.

Quine's point, made using the thought experiment, is that while it's true to say that each of us could point to a rabbit and say 'rabbit', all that establishes is that we have a convention to do so. It does not establish that there are any *facts* as to what terms mean, over and above a convention to behave in a particular way. Thus, when Verity is learning the term 'table' all that she sees is me and other family members pointing to things and saying 'chair', 'table' and so on. She learns to mimic and replicate this behaviour.

When we learn a language (and when Verity learned a language, of course) we are like the explorer being confronted by a tribe that we do not understand. We must learn their language. In so doing, just like the explorer, all that we can hope to pick up on is the evidence, and we can mimic and replicate what we see. But that is the limit of what we can do, and it is a limit of our language that we are forced to learn it in this way. There is no reason to think that there is anything else – anything we can call 'meaning' – involved.

Questions

1. Do you think that there is any reason to prefer one translation of 'gavagai' to another?
2. Would the length of time involved in talking with the tribe make any difference?
3. The case involves only the term 'gavagai'. When engaging with a new language, we are typically presented with more than one term in that language. Would this make a difference to how the case goes? Explain your answer.
4. Would the addition of a third language to the thought experiment make any difference?
5. The case ignores the fact that languages have specific grammatical structures. Would those structures – which may be detectable – make any difference to the case?
6. The case described is often described as demonstrating the 'indeterminacy of translation'. It's often alleged, in science, that theories are underdetermined by the evidence (that more than one theory is compatible with any set of evidence). Is this anything more than a particular instance of this phenomenon?
7. Does it feel to you as if you are just using terms according to a convention, or does it seem like you're doing something else? Does this 'what it feels like' kind of evidence have any weight or bearing here?
8. Does the argument have any practical implications?
9. What *is* a convention?
10. Will a convention do the work that Quine sets out for it? That is, is it enough to simply say that we use language in accordance with a convention that we have learned? Explain your answer.

Reading

Evans, G. 1975. 'Identity and Predication', *The Journal of Philosophy*, 72(13), 343–63.
Quine, W.V. 1951. 'Main Trends in Recent Philosophy: Two Dogmas of Empiricism', *The Philosophical Review*, 60(1), 20–43.
Quine, W.V. 1970. 'On the Reasons for the Indeterminacy of Translation', *The Journal of Philosophy*, 67(6), 178–83.
Wright, C. 1997. 'The Indeterminacy of Translation', in C. Wright and B. Hale (eds), *A Companion to the Philosophy of Language*, Oxford: Blackwell, 397–426.

---

16 Kripkenstein

It is very natural to think that words and sentences have meanings. We do not think that in using a language, all that we are doing is making meaningless sounds or inscribing meaningless marks onto pieces of paper. After all, it hopefully seems to you as if you've just read and understood this! But there are reasons to worry about that claim.

We should start by noting that there seems to be a 'fact of the matter as to what a term means'; there is mind-independent *fact* as to what a given term means. Pick any term you like – 'happy' for instance – and it seems to have some meaning. There is just a *fact* as to what 'happy' means.

The charge that we're going to explore, however, is that there are no facts of the matter as to what a term might mean. To begin, we need some clarifications. Suppose that you had *idealized* epistemic access. That is, you're not deceived and that you have perfect access to all of the facts that are in your experience. The charge is that, even if you had such idealized access to the facts, you still wouldn't grasp any 'fact of the matter' as to what a term means. If, under *those* conditions, you can't be said to grasp what a term means, then there's absolutely no way that we should think that there *are* such facts of meaning. After all, if you can't grasp a fact in idealized epistemic circumstances then you shouldn't think that there is any such fact.

Thought experiment

> Let's assume that you've never before deployed the addition function involving terms greater than 57. OK, so how do we know that '58 + 67 = 125' is true?
>
> There's a very obvious answer: because the symbol '+' means *addition* and when we add these two numbers (67 and 58) the result is 125. But how do we know that '+' means addition? For instance, how do you know that '+' here doesn't mean 'quaddition' which is a very different function? Indeed, the function 'quad' behaves just like addition does, except when we have involved values greater than 57; in that case, the function returns the value of 5. Thus, in virtue of *what* is the right answer to (57 + 68) 125, rather than 5? What fact is it that you grasp that makes 125 correct, rather than 5? It's hard to see that there is one.

There are three reasonably natural responses:

- First: the symbol '+' means addition and I know that 58 + 67 = 125 via extrapolation from previous experiences; because '+' has meant addition when I've used it (correctly) in sums like 1 + 1 = 2; 2 + 2 = 4 and the like, so knowledge of *these* uses is fully sufficient for me to know, in this situation, that I'm using '+' to mean addition. This fails. What my previous experiences show me is that, when dealing with values below 57, the function '+' is addition. But I'm not performing such a sum here. I'm using values of 58 and 67. Now, what in your previous experiences makes it that case that, when you're using '+' in connection with values greater than 57 it means addition? Well, since you've never used '+' in this connection before, nothing.
- Second: what we do is learn *and internalize* rules. In this case, you might think, my internalizing of the rule to add, rather than quad. This fails: we can't be clear which rule we've internalized: adding or quadding. Don't forget, the position we're in is one of not yet having performed operations involving '+' for values greater than 57. So the rule that we've internalized could be add; it could be quad. Until we perform a sum with values greater than 57 we simply can't tell which rule we've internalized. And in order to say that we've got the *right* answer, we must be in a position to say that the rule we've internalized *is* addition rather than quaddition. Since we can't tell between the two rules, the fact that we think we're grasping can't be a rule that we've internalized.
- Third: to understand a term is to 'visualize' what it means. In the case we've been considering that might then lead us to the following strategy: we visualize (or imagine) 57 marbles and 68 marbles and then 'see' that this gives us 125 marbles. This fails: visualization is neither necessary nor sufficient for understanding. For instance, I can understand the term 'quark' without being able to visualize a quark. I can also (seemingly) understand the term 1,000,000,000 without being able to visualize 1,000,000,000 of anything – which is just as well. Notice, also, that I can visualize 7 things without knowing anything about number-talk. So although it might be true that meaning and our capacity to visualize are closely associated with one another, it's clearly not *that* tight a connection.

Thus, we can't discern a *fact of the matter* as to what the term we're considering means. Since we're granting idealized epistemic access, so there simply cannot be a fact here as to what terms mean. Therefore, terms have no meaning.

Questions

1 What does 'idealized epistemic access' mean? Do you think that the notion is coherent?
2 What is meant, do you think, by 'mind-independent fact of the matter about meaning'?
3 What is the result of quadding 57 to 68?

4 The thought experiment requires us to imagine a scenario in which we've never performed a sum involving terms greater than 57. But most of us *have*. How, if at all, does this bear on the outcome?
5 The thought experiment focuses on cases involving mathematics. Can you think of a similar case that involves logical concepts, like conjunction and disjunction?
6 What do you think of the first response to the thought experiment?
7 What do you think of the second response to the thought experiment?
8 What do you think of the third response to the thought experiment?
9 If this argument succeeds, how might we explain the seeming success of communication? Why have you tried to answer these questions if the terms used all have no meaning?
10 Do you think that it is a fact that terms have mind-independent meaning?

Reading

Boghossian, P. 1989. 'The Rule-Following Considerations', *Mind*, 98(392), 507–49.
Fodor, J. 1990. *A Theory of Content and Other Essays*, Cambridge, MA: MIT Press.
Goldfarb, W. 1985. 'Kripke on Wittgenstein on Rules', *The Journal of Philosophy*, 82(9), 471–88.
Kripke, S. 1982. *Wittgenstein on Rules and Private Language*, Oxford: Oxford University Press, esp. chapters 1 and 2.

---

17 Verification in language

Think about two different sorts of attempt to make claims about what kinds of thing exist. Here's the first: there is a cat on the mat. Here's the second: the nothing noths. It can seem as if the first of these is meaningful and the second is not. One way in which to try to explain that fact would be that although we know how to verify whether or not the first one is true, we do not know how to verify whether the second one is true. To verify that 'there is a cat on the mat' is true, I need simply to inspect the mat. But I wouldn't know how to begin to verify that 'the nothing noths' is true.

Some – notably the logical positivists – thought that something like this distinction could be used to capture the difference between meaningful and meaningless claims and to give an analysis of meaning. Roughly, the idea is that a claim is meaningful if and only if we can observe what it says to be either true or false. But this rough idea will require some modification. For instance, consider the sentence: 'there is a far side to the moon'. I cannot verify this claim to be true. I lack the rocket-ship required. Nonetheless, it's meaningful.

A more interesting version of a verification principle of meaning might say something like: a claim is meaningful if and only if we know, *in principle*, how to

verify whether it's true or false. Since I know how *in principle* to verify that it's true that there's a far side of the moon (get hold of the required rocket-ship, etc.), this deals with the initial worry. But there may be a problem:

Thought experiment

> Newton's law of universal gravitation states that any two bodies in the Universe attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them. Newton is often credited with discovering this law. But Newton did not know, even in principle, how to test it. Neither does Sally. Sally has two worries about it: first, she doesn't know – even in principle – how *she* could check every single body in the Universe. Second, Sally thinks that, as a scientific law, the law should apply to massive bodies that *could* exist, but that don't. Sally doesn't think that she knows, even in principle, how to check whether merely possible massive bodies would attract one another with the described force. If Sally is right, and the principle of verification is also right, then Newton's law of universal gravitation isn't meaningful.

One response would be to accept the result of the thought experiment. But that seems extreme. It would seem better to try to modify the verification principle of meaning that we gave. One suggestion, due to Ayer, is as follows, and will treat any factual or experiential proposition as meaningful:

> Let us call a proposition which records an actual or possible observation an experiential proposition. Then we may say that it is the mark of a genuine factual proposition, not that it should be equivalent to an experiential proposition . . . but simply that some experiential propositions can be deduced from it in conjunction with other premises without being deducible from those other premises alone.
>
> (Ayer 1946: 38–9)

Let's follow this through. 'There is a cat on the mat' is an experiential proposition because it is an observation we could make – I could observe a cat being on a mat. The case involving Newton's law is also dealt with. That any two bodies in the Universe attract each other with a *force* that is *directly proportional* to the product of their masses and inversely proportional to the square of the distance between them can be used to deduce, in conjunction with premises that state which massive bodies exist, that these bodies will move towards one another. That these bodies will move towards one another is not deducible merely from their existence.

One notable consequence of this kind of verification principle is that many claims in metaphysics (and perhaps also other areas of philosophy) will turn out to be meaningless, for many claims in metaphysics (and other areas of philosophy) are not such that they can be used to generate observable or testable claims. Indeed, this is the purpose to which at least some verificationists wished to put the principle of verification.

Questions

1 Do you think that something like the verification principle is intuitive? Do you think that gives us any reason to accept or reject it?
2 Ayer himself had some doubts about the principle, as stated. In particular: if O is an observation statement, and N is the putative nonsense claim that 'the nothing noths', we can render N meaningful by asserting: 'if N, then O', in conjunction with 'N'. Prove that this leaves N meaningful.
3 In view of question (2), can we successfully modify the principle of verification to avoid the counterexample?
4 In light of the reading, how do you think we should best state the verification principle?
5 Suppose that it's true that metaphysical claims are not meaningful, given the verification principle: do you think that metaphysical claims constitute a counterexample to the verification principle?
6 The general insight that the verification principle is trying to capture is that a claim is meaningful if and only if its contents can be verified. Do you think that claim, itself, can be verified? If it cannot be, is that a problem?
7 If you were going to reject a principle of verification, what kind of explanation would you give as to why 'the nothing noths' is not meaningful?
8 How bad would it be if scientific claims were meaningless? Would the view be defensible?
9 Is there a reason to think that it would be worse if scientific claims were meaningless than if metaphysical claims were meaningless?
10 Do you think that we should adopt a verificationist principle of meaning?

Reading

Ayer, A. 1946. *Language, Truth and Logic*, London: Gollancz, esp. the introduction, which gives an overview of Ayer's project.
Foster, J. 1985. *Ayer*, London: Routledge.
Gower, B. (ed.) 1987. *Logical Positivism in Perspective: Essays on Language, Truth and Logic*, London: Croom Helm.
Wright, C. 1989. 'The Verification Principle: Another Puncture – Another Patch', *Mind*, 98(392), 611–22.

---

18 Grice and an intentional theory of meaning

Suppose that I tell you that there's a pen on the desk. Why would I do that? Presumably, I'd tell you that there's a pen on the desk because I intended you to form the belief that there's a pen on the desk! You recognize my intention to get

you to form the belief that there's a pen on the desk, and, let's suppose, you *do* form the belief that there's a pen on the desk.

Reflection on this case – and others like it – make it tempting to some philosophers to think communicating intention is what meaning consists in. Slightly more formally, that when you say something:

1 You intend your interlocutor to form a belief that B.
2 Your interlocutor recognizes your intention to get them to form a belief that B.
3 The reason that your interlocutor forms the belief that she does is explained by her recognizing your intent that she form the belief that B.

These are our necessary and sufficient conditions for meaning. But it looks like this might require modification:

Thought experiment

> Suppose that you turn up in my office hours to ask me a question. You ask your question, which is (sadly), 'what course is it we're taking with you?' To vent my frustration, I say 'fffaffaa ffaffaa jaffaa'. I do this with the intention that you form the belief that 'Jonathan thinks you're a fool with jam for brains'; you recognize my intention that you form this belief, and, of course, it is the case that you form the belief, in part because of recognizing my intent that you form the belief. Given our analysis, that then means that 'fffaffaa ffaffaa jaffaa' means 'Jonathan thinks I am a fool with jam for brains'. But it doesn't mean anything; it's a random string of noises. As a consequence, this theory fails.

One way in which we might try to rescue the analysis is to avail ourselves of the distinction between illocutionary and perlocutionary intentions. In the previous case, you recognized my intention to insult you. By virtue of this, you were insulted. Other cases of illocutionary intentions include warning (if I shout 'look out', you are thereby warned) and promising (if I tell you, 'I promise to give you a First', then you are thereby promised to). But other intentions (perlocutionary intentions) don't seem to be that way. If you recognize my intention to embarrass you, you are not thereby embarrassed; if I have the intention that you form the belief, p, you do not thereby form the belief that p. What we must do, then, to deal with this first counterexample is add an additional clause:

4 The intentions of the speaker are illocutionary.

In the case, above, where I said 'fffaffaa ffaffaa jaffaa" my intention was to insult you. You can know that without thereby being insulted. Thus, the intention involved in the putative counterexample is *perlocutionary* and so is not a counter example.

But there might be worries about even this. Suppose that I am a soldier, dropped behind enemy lines during the Second World War. I meet up with the local French resistance. I do not speak French (poor choice to send, really), and they do not speak

English. None of us speaks German. As we make our way towards the centre of the local village, I spot, up ahead, some German soldiers. In order to *warn* my resistance colleagues, I shout the only word of German that I know; Kartoffelkopf. This translates as 'potato head'. I shout this with the intention of warning my resistance colleagues as to the presence of German soldiers and intend them to form the belief 'there are German soldiers'. They recognize my intention, and form the belief that 'there are German soldiers' and, clearly, my intention plays a part in their forming the belief that 'there are Germans soldiers'. The intention, to warn, is illocutionary; if you recognized the intention to be warned, you are thereby warned. Thus, this is a counterexample to even this revised version of Grice's thesis.

Questions

1 Using the first Gricean analysis, explain what the meaning of the expression, 'Jonathan likes eating cheese' consists in.
2 How was the intention recognized in your answer to (1)?
3 The first analysis given focuses on the meaning of what speakers *say*; can you think of how we might apply this more generally, to cases involving what is written?
4 In the thought experiment, do you think that it's right that 'fffaffaa ffaffaa jaffaa' is treated as a counterexample? For instance, do you agree that 'fffaffaa ffaffaa jaffaa' really is meaningless?
5 Would cases involving sarcasm generate problems for (1)–(3) as an analysis? Explain your answer with reference to a case.
6 Describe, in your own words, the difference between illocutionary and perlocutionary intentions: give some examples of your own.
7 In the case involving the Second World War soldiers, why *might* one think that 'Kartoffelkopf' does not mean 'Warning: German soldier'?
8 Do you think that there's any scope for saying that, in the case described, Kartoffelkopf' *does* mean 'Warning: German soldier'?
9 Can you think of any other conditions that we might add to (1)–(4) that would help us solve the problems?
10 In general, do you think that trying to analyse meaning in terms of intention is a project worth pursuing (even if the analysis here turns out to fail)?

Reading

Blackburn, S. 1984. *Spreading the Word*, Oxford: Oxford University Press, esp. chapter 4.
Davis, W. 1989. *Implicature: Intention, Convention, and Principle in the Failure of Gricean Theory*, Cambridge: Cambridge University Press.
Grice, H. P. 1989. *Studies in the Way of Words*, Cambridge, MA: Harvard University Press.
Miller, A. 1997. 'Tacit Knowledge', in C. Wright and B. Hale (eds), *A Companion to the Philosophy of Language*, Oxford: Blackwell, 146–74.

19 Truth-conditional semantics

There is a very strong connection between truth and meaning. It would seem very natural to say that to know what an expression means is to know when it could be applied truthfully. For instance, if I know what it means to say 'the box is red', then I will know when that sentence can be used truthfully. So, if we are after an analysis of meaning, it would make sense to use truth to help us deliver that.

One way in which we might try to do that is to offer the following analysis of the meaning of an expression: a sentence, s, in a language, L, is true iff p. Here, p is intended to be a statement in a language that is, formally at least, distinct from the language in which s is stated. This language will, for our purposes, look exactly like English, but we need to keep in mind that it is formally distinct. We will call it the 'metalanguage'.

With this in hand, it's easy to give the meaning of a sentence: 'snow is white' is true iff snow is white. However, there seems to be a problem for views that place truth at their core:

Thought experiment

> Brian, a small child, is learning English. His parents are Briony and Paul. Brian does not already grasp the meaning of many expressions in English. Yet he is able to learn new concepts. Briony and Paul think it's hard to see how he can do that. For instance, Brian does not know what the words 'snow is white' *mean* – in either English or in the metalanguage. Although Briony and Paul are good believers in a truth-conditional semantics, and would like to endorse a truth-conditional semantics, their experiences with Brian make it impossible for them to see how they can. The concern is one of learning: how can someone who does not already understand the meaning of a particular term come to recognize whether it's being truthfully applied?

Happily, within this model, there are the resources to help them out. According to Donald Davidson, we are to employ something known as the 'principle of charity', such that we should ascribe to a speaker the thoughts and beliefs that we would have in their shoes.

Thus, the reason that Brian is able to pick up the meaning of 'snow is white' is that when his parents are teaching him how to understand the meaning of the expression, there are a variety of contextual factors: his parents' behaviour, their pointing to the snow, their contrasting it with other white objects in his environment. Because Brian is able to recognize something of his parents' intentions or beliefs in this circumstance, he is able to learn.

Questions

1 Write out a straightforward sentence of English. Use the truth-conditional model, above, to give the meaning of the expression.

2 Using the truth-conditional model, give the meaning of 'Schnee ist weiss'. Are there any problems in giving the truth-conditions of a sentence in another language?
3 Do you think that there is the close connection between truth and meaning described at the outset? Explain your answer.
4 Why might it be important to give a reductive analysis of the meaning of an expression?
5 There are potentially infinitely many sentences of English. Is there a reason to worry about the truth-conditional model because of that? Explain your answer.
6 Do you think that we normally employ the principle of charity when engaging with fellow-speakers? Give examples that support your conclusion.
7 There are, plausibly, some occasions when we fail to use the principle of charity. See if you can write one down.
8 Does this failure of the principle of charity in some cases mean that we should reject it?
9 How should the Davidsonian, who endorses a truth-conditional semantics and the principle of charity, make sense of the meaning of sentences that are questions or commands that, at least intuitively, do not seem to have truth-conditions?
10 Can you think of a context in which the Davidsonian theory has practical consequences?

Reading

Davidson, D. 1980. *Essays on Actions and Events*, Oxford: Oxford University Press.
Davidson, D. 1984. *Inquiries into Truth and Interpretation*, Oxford: Oxford University Press.
Glock, H. 2003. *Quine and Davidson on Language, Thought and Reality*, Cambridge: Cambridge University Press, esp. chapter 5.
LePore, E. 2005. *Donald Davidson's Truth-Theoretic Semantics*, Oxford: Oxford University Press.

---

20 Indexicals

There is a particular class of expression, the truth of which depends upon the location (broadly understood) at which it is uttered or written down (to use a technical term, we may say that the truth of the expression depends upon the location of its tokening). Examples will be familiar to everyone. Sentences like 'it's raining *now*' depend for their truth upon how things are *at that location in time*. Sentences like 'it's warm in *here*', depend for their truth upon how things are *at that location in space*. Sentences like 'I am hungry' depend for their truth on how things are with the person who utters them. The following thought experiment is borrowed from John Perry:

Thought experiment

> Arthur is doing the shopping at his local supermarket. Whilst pushing his trolley around the supermarket, he notices a trail of sugar. Whilst wandering around the store, looking for all of the different items that he requires, Arthur comes to think to himself, 'well, someone is making a right mess here'. More time passes and Arthur continues to see trails of sugar all the way around the store. Eventually, Arthur looks down into his trolley and notices that his bag of sugar has been leaking. Arthur thinks, 'I am the messy shopper.'

There are a number of connected problems that this case raises. We will focus here on just one. Let's think about how we might offer a truth-conditional analysis of the sentence. Recall from the last chapter that a sentence s is true iff p. But, in that case, think about what analysis we might look to provide: 'I am the messy shopper' is true iff I am the messy shopper. But note: the right-hand side of the bi-conditional is not something that is true, except if uttered by Arthur (or by some other messy shopper).

Thus, were Arthur to say it to his friend Beryl, Beryl could not treat 'I am the messy shopper', said by Arthur, as having the truth-conditions just described. If she did, then she would be committed to the claim that: 'I am the messy shopper' (said by Arthur) is true if and only if *I* (Beryl) am the messy shopper. Clearly, that's not right.

In order to make sense of the truth-conditions of indexical expressions, we need some technical apparatus. In particular we will need the notions of a context and a character.

A context is something like an agent, a time, a possible world, and a spatial location. So, for instance, my utterance (right now) of 'I'm hungry' has the context <Jonathan Tallant, 16:20, w, C45>. (My office is C45.) The context of assertion will vary depending upon the agent, time, world and location of the utterance.

Next, character. The character of an expression is the function that determines, within a particular context, what the content is. Now, in the context just specified, the sentence 'I am hungry' serves to pick out the agent, *me*, and the time *16:20* because of the character; that is, it is a part of the function of the expression that it picks out that individual and time, given that context. To give you a slogan, a character is a function from context to content.

This enables us to say the following: the content of 'I' with respect to a context, c, is the agent of c; the content of 'here' with respect to a context, c, is the spatial location of c; the content of 'now' with respect to a context, c, is the time of c; the content of 'actually' with respect to a context, c, is the world of c.

To return to the case of Perry's messy shopper, the truth-conditions of 'I am the messy shopper' will be <Arthur, time, w, aisle 14>.

Questions

1 Write out three indexical expressions and, using the analysis specified, state their truth-conditions.

2 Explain the problem, in your words, giving a straightforward truth-conditional analysis of the form '"I am the messy shopper" is true iff I am the messy shopper'.
3 The example given involved the indexical 'I'. Explain how a truth-conditional analysis of the form 's' is true iff p won't deliver the right result for a case involving 'here'.
4 The example given involved the indexical 'I'. Explain how a truth-conditional analysis of the form 's' is true iff p won't deliver the right result for a case involving 'now'.
5 Consider the sentence, as we might speak into an answering machine, 'I am not here right now'. Does this sentence pose problems for the analysis suggested? If so, how might we try to patch the analysis?
6 It seems right to say that Arthur learns something when he learns that he is the messy shopper. What do you think it is that he learns?
7 Do you think that it's plausible to suppose that the proposition <Arthur, time, w, aisle 14> is something that Arthur believes? If not, does this pose a problem?
8 Does this make any practical difference to anyone? For instance, do you think that understanding how indexical expressions work might help us code artificial intelligence? Do you, for instance, think that anything said here would help to code a machine to understand human communication?
9 Demonstrative expressions (this, that, etc.) seem to be very similar to indexicals. Can you work out how we might try to provide truth-conditional analyses for these expressions?
10 Do you think that we should be trying to provide a truth-conditional analysis of indexicals, rather than some other form of analysis (perhaps one that focuses on the intentions or desires of the speaker)? Explain your answer.

Reading

Kaplan, David. 1989. 'Demonstratives', in J. Almog, J. Perry and H. Wettstein (eds), *Themes from Kaplan*, Oxford: Oxford University Press, 481–563.

Perry, J. 1979. 'The Problem of the Essential Indexical', *Noûs*, 13(1), 3–21.

Predelli, S. 1998. 'I Am Not Here Now', *Analysis*, 58(2), 107–12.

Wettstein, H. 1986. 'Has Semantics Rested on a Mistake?', *The Journal of Philosophy*, 83(4), 185–20.

Part III

METAPHYSICS

21 Abstract objects

If they exist, abstract objects are very unlike physical objects. They will exist outside space and time. They will not *cause* events to occur. Given these peculiar features, we would need to have a good reason to think that they do exist. One such reason might come in the form of (what's known as the enhanced version of) the indispensability argument. The idea, roughly, is that some abstract objects will be mathematical objects – numbers being a good example. For numbers, if they exist, will be unchanging and will not obviously exist in space and time (I've never seen them, though of course I've seen the figures that stand for them). Thus,

- We ought rationally to believe in the existence of any entity that plays an indispensable explanatory role in our best scientific theories.
- Mathematical objects play an indispensable explanatory role in science.
- Hence, we ought rationally to believe in the existence of mathematical objects.

Proponents of this argument-form then typically try to give some reason to think that mathematical objects do indeed play an indispensable role in our best scientific theories. One such attempt has roughly the following structure: the life-cycle of the North American periodical cicada is 17 years in the north of the territory it inhabits, and 13 years in the south. This life-cycle has evolved, the evolutionary biologists tell us, because it is advantageous. It is advantageous because, 17 and 13 both being prime, having a 17-year or 13-year life-cycle minimizes intersection with other lifecycles – including those of predators. It minimizes this intersection because it is prime. Primeness is a feature of numbers, and nothing else. Thus, numbers feature as an indispensable part of our best scientific theories.

The position that there really are such things as numbers, motivated by the above argument, can be challenged. The following thought experiment outlines one such way in which to develop that challenge:

Thought experiment

> Let's allow that there are numbers and that numbers are abstract objects. Imagine a way that the world *could be*. Include in such a possible scenario all of the physical objects that actually exist. Arrange those physical objects exactly as they are now. However, from this scenario, excise any numbers and any abstract objects. Such a scenario seems conceivable. We normally take the conceivability of some scenario to indicate its possibility. Thus, it seems possible that the world might be exactly as it is, even if there were no numbers. And, in that case, it seems as if numbers aren't really explaining how the physical world is the way it is. It therefore appears that we should not believe that they exist.

If successful, this argument undermines the indispensability argument, given above. And if the argument is successful, then presumably we ought to give up on the existence of abstract objects. It would then behove us to say something about the truths of mathematics. If there are *no such things as numbers*, then why is it true that 2 + 7 = 9? Why is it true that there is a number between 3 and 5 – if there are no numbers?

Two obvious strategies come to mind. The first strategy – a view known as mathematical nominalism – is to try to understand mathematical claims as about something other than mathematical objects. To illustrate, to say that 'the number of electrons is 2' is better interpreted as saying: 'there is an electron; there is another electron'; to say that 'the number of electrons is 3' is better interpreted as saying: 'there is an electron; there is another electron; there is a further electron'. This process can simply be iterated.

The other obvious strategy is to deny that mathematical talk *is* true, though to agree that it's very useful. Call this view 'mathematical fictionalism'. The thought might go like this. Mathematical discourse seems to be about numbers. But there are no numbers. (Abstract objects are just too strange for us to countenance.) In that case, it's straightforward to see that all of our mathematical discourse is false. However, mathematical discourse is still *extremely useful*. But we shouldn't confuse its usefulness for its truth.

Questions

1 What are the strengths and weaknesses of the enhanced indispensability argument?
2 Do you think that the thought experiment shows that numbers cannot be playing a genuine explanatory role? Why?
3 How do you think that the proponent of the existence of numbers should respond to the thought experiment? How plausible do you find that response?
4 How plausible do you find the idea that abstract objects exist?
5 As we saw in the introduction of nominalism, the view requires quite extensive paraphrasing of mathematical discourse. How problematic do you think it will be to provide that paraphrase? Can you think of cases that will be impossible to paraphrase in this way?
6 How would you try to deal with the cicada case if you were a nominalist?
7 How plausible did you find the view that mathematical claims could all be false but useful?
8 How would you try to deal with the cicada case if you were a fictionalist?
9 The thought experiment might seem to share some superficial features with the subtraction argument (see Chapter 28). Do you think that the similarity is close enough to allow strategies from other debate to influence the other? If so, how would you try to make sense of this?
10 Of those outlined, which view do you find most plausible?

Reading

Baker, A. 2003. 'Does the Existence of Mathematical Objects Make a Difference?', *Australasian Journal of Philosophy*, 81(2), 246–64.

Baker, A. 2005. 'Are There Genuine Mathematical Explanations of Physical Phenomena?', *Mind*, 114(454), 223–38.

Daly, C. 2006. 'Mathematical Fictionalism: No Comedy of Errors', *Analysis*, 66(291), 208–16.

Melia, J. 2000. 'Weaseling away the Indispensability Argument', *Mind*, 109(435), 455–79.

---

22 Causation

Causation is very familiar. Suppose that I drop my glass and it smashes. It seems wholly correct to say that my dropping the glass caused it to smash. Likewise, suppose that my daughter throws a ball at a window and the window breaks. It seems wholly correct to say that the throwing of the ball causes the smashing of the window. But how should we analyse the notion of a cause as it's being used here?

One very simple and intuitive thought is to say that an event, c, is the cause of another event, e, if and only if both c and e occur, and had c *not* occurred, then e would not have, either. I say that the idea is simple and intuitive precisely because it seems to capture our first-up intuitions about the cases described above. For instance, take the case involving my daughter: she threw the ball; the window broke. It seems to be true that, had she not thrown the ball, then the window would not have broken. Thus, it seems, we have a good starting point. But, as we'll now see, there are some problems with taking this line:

Thought experiment

> My daughter throws the ball at the window. It strikes the window. Her throwing the ball is the cause of the window breaking. Nonetheless, shortly after she threw the ball, I also threw one at the window. If my daughter hadn't thrown the ball, then my ball would have struck the window and the window would have broken. This means that *had my daughter not thrown the ball, the window would still have broken*. Our tentative analysis, suggested above, would then rule my daughter's actions not to be the cause of the broken window. But this is wrong. She did cause the breaking of the window. Something has gone wrong with our analysis.

The analysis in question is a very simple version of the counterfactual theory of causation and the problem expressed in the example is a basic form of the problem of late pre-emption. There are, as you might imagine, attempts to save the analysis as well as further iterations of the problems.

But aside from these technical problems, there are some that strike a more simple note. For instance, one point that is missed by the counterfactual analysis is the seemingly close tie between the transfer of energy and causation itself. There is an intuitive sense in which once energy is transferred from one process to another, the second process is acted upon, and caused to behave in a particular way. An example involving billiard balls makes the case nicely: suppose that one ball rolls along the surface of a billiard table and strikes another ball. The second ball moves off. It moves off because of the transfer of energy from one ball to another. Consideration of this case, and others like it, have tempted some people to speak of causation as requiring some transfer of energy from one state to another.

All the same, these energy transfer stories might seem not to give us a particularly elegant solution to some other problems we encounter in considering causation: particularly, causation by absences. If my daughter neglects to feed our cat for long enough, the cat will die. In such cases it seems as if her *not* performing some particular action – feeding the cat – is the cause of the death of the cat. But this is odd, particularly if we consider an energy transfer theory: there is no action that she *did* perform, such that energy can flow from *it* to the death of the cat. In general: it seems wrong to say that there is any particular action that she did perform that caused the death of the cat. It can be difficult to see how to account for such cases.

Questions

1 What, if anything, do you think is good about the counterfactual analysis?
2 How do you think that the proponent of the counterfactual analysis of causation might respond to the problem of pre-emption? Can you think of any difficulties with that response?
3 How would you feel about a case in which my daughter throws, but I do not. But the only reason I do not throw is that my daughter does. Do you think her action still causes the breaking? Is my intention relevant to the success (or otherwise) of the counterfactual theory?
4 It seems to be true of the case involving my daughter, that, had the Big Bang not started the Universe, then the window would not have broken. Does this make the Big Bang the cause of the breaking of the window? Does this absolve my daughter of responsibility?
5 What do you think about the energy transfer theory?
6 Do you think that my daughter's not feeding our cat would be a cause of its death? If not, would it be wrong to punish her for it?
7 Can you think of other cases in which the lack of some action is a cause? How would you propose to analyse the causation in this case?

8 Given the difficulties of analysing causation, do you think that we should deny that there *is* causation? What would be a problem with such a causal eliminativism?
9 There seems to be a connection both between causation and moral responsibility and between causation and explanation. Do think it's right to say that if an event, c, causes another, e, then c gives us an explanation of e and someone performing c is thereby responsible for e? If not, how would you connect causation to responsibility and explanation?
10 Do you think that there could be different conceptions of causation: for instance, a scientific conception, an 'everyday' conception or a legal conception?

Reading

Dowe, P. 2001. *Physical Causation*, Cambridge: Cambridge University Press.
Lewis, D. (1986). 'Causation', in *Philosophical Papers 2*, Oxford: Oxford University Press, 159–213.
Paul, L. A. and Hall, N. 2013. *Causation: A User's Guide*, Oxford: Oxford University Press.
Schaffer, J. 2008. 'Causation and Laws of Nature: Reductionism', in T. Sider, J. Hawthorne and D. W. Zimmerman (eds), *Contemporary Debates in Metaphysics*, Oxford: Blackwell, 82–107.

---

23 Composition

In his excellent *Material Beings*, Peter Van Inwagen asks us to consider *when* composition occurs. That is, what we have to do to two or more objects to get them to compose another. The question is easily asked, but is surprisingly hard to answer. What we're after is a universally applicable answer to the question. We want our answer to apply just as well to atoms (that are composed of protons, neutrons and electrons) and chairs (that are composed of legs, seat and back). The question is known more formally as the 'Special Composition Question':

The way in which we are encouraged to proceed is to try to find an answer and then look for counterexamples to the answer. If we can find a counterexample to the answer, then the answer is to be rejected.

Thought experiment

> Think about some lego bricks, separated out in a clear plastic bag. It seems very natural to think that what we have to do to get these to compose is to fasten them all to one another. So, it makes sense to think that a good candidate answer to the Special Composition Question is that the bricks compose if, and only if, we fasten them to

> one another. Once, and only once, we've fastened the bricks to one another do they compose a further object. Call this answer 'Fastening'.

The problem with Fastening is that there seem to be cases where objects compose but aren't fastened to one another, and there seem to be cases where objects don't compose but *are* fastened to one another. For an example of the first, think about a deck of playing cards. There is an object – a deck of playing cards – but it doesn't seem to be essential to the existence of the deck that the cards are fastened to one another. Even if we deal out the cards, we still have a deck.

For an example of the second, think about two people shaking hands with firm grips. The two people are fastened to one another, via their grips, but do not then compose a new object. As noted, the Special Composition Question is very hard to answer. In part because of this, some people think that we might reasonably conclude that there are *no* composite objects. That is, although there are (presumably very small) objects that lack parts, they never come together to *compose* a further object. All that exist are loose arrangements of very small things without parts – collections of objects without parts arranged *as if* they are composite objects. To borrow some language from the literature, rather than saying that (for instance) 'there are chairs', we should say that there are collections of mereological simples (objects without parts) arranged 'chair-wise'. The view that there are no composite objects is called 'mereological nihilism'.

Others think this too extreme, but agree that finding a particularly intuitive response to the SCQ is problematic. Instead, though, these mereological universalists claim that in order to compose a further object, all that objects x and y have to do is to exist! Thus, whenever there exist two objects, there exists another object made out of them. The view has certain counterintuitive consequences, since it implies that there is a composite object (for instance) composed out of me and Saturn! But although it does commit us to the existence of these strange objects, it also commits us to the existence of all the ordinary objects that we naturally suppose to exist.

And there's a question as to how much we should worry about these strange objects. Universalists will often say that the only reason we find these objects strange is that we don't often think about them. We don't engage with them. And because of that, we don't talk about them much. If we don't talk about something, then, of course it's going to be the case that when we introduce them they strike us as a bit strange – on first inspection, at any rate.

Whatever the merits of universalism and nihilism, it seems clear that they're pretty extreme views. It seems most natural to think that there will be some view – in the middle – that gives us a good answer to the Special Composition Question. But what that more moderate answer is, is a tough question to answer.

Questions

1 What do you think of mereological universalism?
2 How plausible did you find the universalists' quick defence of strange objects?

3 What do you think of mereological nihilism?
4 How plausible did you find the suggestion that we replace talk of composite objects with talk of 'simples arranged object-wise'?
5 Ignoring nihilism and universalism, how do *you* think we should answer the SCQ?
6 One more moderate view, defended by Van Inwagen, is that the only composite objects are living things; we should be nihilists about the non-living. Does such a view strike you as plausible?
7 Given that the SCQ asks us to write down necessary and sufficient conditions that describe when we think composition occurs, do you think we really need to answer the SCQ? After all, doesn't asking this question imply that we can already tell when composition occurs? What value do we then add by explicitly answering the SCQ?
8 Can you think of any reason why someone might be sceptical that there's a genuine difference between universalism and nihilism?
9 Why should anyone care about the answer to the SCQ? Could an answer ever make a difference to the way that someone lives their life?
10 If there are no objects with parts, how does this bear on the questions of persistence and time travel (discussed in Chapters 25 and 29)?

Reading

Markosian, N. 2008. 'Restricted Composition', in T. Sider, J. Hawthorne and D. W. Zimmerman (eds), *Contemporary Debates in Metaphysics*, Oxford: Blackwell, 341–64.
Rosen, G. and Dorr, C. 2002. 'Composition as a Fiction', in R. Gale (ed.), *The Blackwell Guide to Metaphysics*, Oxford: Blackwell, 151–74.
Sider, T. 1993. 'Van Inwagen and the Possibility of Gunk', *Analysis*, 53(4), 285–9.
Van Inwagen, P. 1990. *Material Beings*, Ithaca, NY: Cornell University Press.

---

24 Constitution

Let's start off with a relatively simple bit of theory. Suppose that we have two terms, x and y. These terms might refer to the same object; they might not. If x is identical to y, though, then what do we know? Well, one very natural thought is that if x and y are identical, then they share all of the same properties. After all, it would be pretty strange if Jonathan was identical to Dr Tallant, and yet Jonathan was 6ft tall and Dr Tallant was 5f 8in!

It would also seem very sensible to think that if we worked out that x and y both had all of the same properties, then they are in fact the same thing. Thus, if Jonathan has the property of being 6ft, and being a philosophy lecturer, and living in Nottingham and . . . , and it also turned out that Dr Tallant has the property of

being 6ft, and being a philosophy lecturer, and living in Nottingham . . . then they are in fact one and the same thing. Jonathan *is* Dr Tallant.

Let's hold on to a principle, then, which encapsulates both of these principles and is called Leibniz's law. If x and y have exactly the same properties, then x is identical to y; and if x and y are identical, then x and y have exactly the same properties.

What of it? Well, now consider this thought experiment:

Thought experiment

> Suppose that a statue maker forms a statue from a lump of clay. It seems pretty clear that we have a single entity: a statue, albeit one that is constituted by clay. But now, note the following: the statue cannot survive a squashing. Were a particularly belligerent patron of the statue maker to come along and squash the statue flat, the statue would cease to exist. We may say that the statue has particular properties, including: not being able to survive a squashing. Note, though, that the clay would survive the squashing. Simply squashing a clay statue does nothing to make the clay go out of existence. We may say that the clay has particular properties, including: being able to survive a squashing.

If that is all correct, then note the following: the clay has a property that the statue does not (and vice versa). We said, above, that Leibniz's law entails that objects with different properties are not identical. And that means that there *cannot* be just a single entity – a statue. Rather, it seems that we must have both a statue *and* a lump of clay, present throughout the entire lifetime of the statue.

The natural conclusion of the thought experiment is that there are two objects, the statue and the lump. But this strikes many philosophers as very strange. After all, we don't normally think that two objects can coexist at exactly the same space. My chair and my mobile phone *certainly* cannot coexist at exactly the same points in space.

A number of points can be made in response. If one wished to defend the view that there really are two objects here and that the clay constitutes the statue – for we'll call this the 'constitution view' – one might note that there is an important difference between the statue and lump on the one hand, and my chair and mobile phone on the other. The statue and the lump *share all of their parts*, in a way that the telephone and chair do not. Since they share the same parts, the statue and lump share all of the same matter, too.

But there are other solutions that we could look to. For instance, we could deny Leibniz's law. Here is one thought. Let us call the result of the sculptor's actions, 'David' and the clay 'Lump': then, although it makes sense to say that David is the same statue as Lump and that David is the same lump of clay as Lump, that is all that we can say. Specifically, we cannot make judgements about identity considered in the absence of a *kind*. Sameness doesn't work that way. Sameness judgements are always to be considered relative to a kind. This sort of view – one that commits us to the relativity of identity – denies Leibniz's law, for the law is

supposed to be unrestricted to a kind. There is, then, no problem, for the law – which led to the problem – is to be denied.

Questions

1 What's your first impression of how we should respond to the thought experiment?
2 Suppose that we tweak the case: the statue and lump are now taken to coexist for the whole of their lives. Does this change how you think that we should respond?
3 What do you think of the concern that no two objects can occupy the same location? Do you think it compelling?
4 Do you think that we should endorse Leibniz's law?
5 Do you think that we should adopt the relativity of identity?
6 In Chapter 23 we surveyed mereological nihilism, according to which there are no composite objects. What would the nihilist say about this case? Does this case give us any reason to endorse nihilism?
7 Is there a temptation to treat this dispute as merely a verbal one (it's just about how we use words), without any serious metaphysical import? Explain your answer.
8 What difference, if any, could a resolution to this question make to the way we live our lives? Does this matter?
9 How do considerations of parsimony bear on this matter? How much difference to they make, if any?
10 Is there any scope for using counterpart theory (see Chapter 26) to help us solve this problem? (Hint: which modal properties an object might be thought to have may depend upon context, given counterpart theory.)

Reading

Fine, K. 2003. 'The Non-Identity of a Thing and Its Matter', *Mind*, 112(446), 195–234.
Geach, P. 1967. 'Identity', *Review of Metaphysics*, 21(1), 3–12.
Lewis, D. 1971. 'Counterparts of Persons and their Bodies', *The Journal of Philosophy*, 68(7), 203–21.
Thompson, J. J. 1998. 'The Statue and the Clay', *Noûs*, 32(2), 148–73.

---

25 Persistence

The basic challenge faced by any theory of persistence is to give an account of how one thing can stay itself – that is, stay the very same thing over a period of time – whilst also changing the properties that it has.

Thought experiment

> Pick an object – just about any object. Whatever object it is that you picked, it's highly likely that the object in question isn't quite the same now as it was. For instance, suppose that we focus our attention upon a particular blade of grass in my garden. That blade of grass *was* standing straight and tall. However, my cat recently came bounding across the garden and flattened it. Where once it was straight, it is now bent in half. But, although this has certainly damaged the blade of grass, it seems that it has survived. Although it has a slightly different nature to the one it had previously, it nonetheless seems right to say that the blade of grass is the same blade of grass that it was before. In language more appropriate to the topic, the grass has persisted through a change in its properties.

But how can that be? The grass is bent. The grass is straight. Both straightness and bentness are properties of the grass. It cannot have them both, for nothing can be both straight and bent. But if the grass *is literally the same object* at both times, then *it* exists at both times. And since *it* has both properties, we have a contradiction.

The natural response to this thought experiment is to say that the grass has different properties *at different times*. Thus, although, yes, the grass really is *the same blade of grass* at each time, the properties it has aren't genuinely contradictory. Rather, the blade of grass instantiates the properties being-straight-at-t and being-bent-at-t* (where t* is later than t).

Such a position about persistence is known as 'endurance', whereby objects persist through time by being 'wholly present' at each moment through which they persist and instantiate 'time-relativized' properties. Having the property of straight-at-t and having the property of bent-at-t* is no contradiction. Thus we have solved the puzzle. Call this view an 'endurance' account of persistence through time.

Another solution is to treat persistence through time much as we treat being extended in space. Just as we might say that a table is both scratched and smooth, by saying that a *part* of the table is smooth and that a *part* of the table is scratched, so in the temporal case we might say that there is a single object that persists through time – in our case, the blade of grass – and that each of its parts has distinct properties.

Thus, the temporal part of the blade of grass that exists at t is straight. The temporal part of the blade of grass that exists at t* is bent. Of course, having different parts with different properties is no contradiction at all, and so we have solved the puzzle. Call this view a 'perdurance' view of persistence.

There are a range of different considerations adduced by theorists from each camp who are looking to argue in favour of their view. There are also further views available! One point to note here, though, concerns best-fit with our natural way of thinking about the world. It seems very natural to think that objects are *wholly present* at each moment of time through which they persist.

Thus, this seems to favour the endurance model. In contrast, the idea of properties being relativized to time seems a bit of a departure from our natural way

of thinking. The idea of objects actually being large, temporally extended things, not wholly present at each moment through which they persist, seems a little odd (at least on first inspection). However, since this model allows us to preserve the idea that properties are not relativized to times, the perdurance view seems to gain a little motivation from matching up well with our natural conception of time.

Questions

1. Which view of persistence to do you prefer? Why?
2. How much do you think it matters whether or not an account of persistence fits well with our natural way of thinking?
3. Do you think that the supposed natural ways of thinking match up with your way of thinking about persistence? Do you think this matters?
4. One natural conclusion to take from the thought experiment would be that objects *cannot* persist through time: why do you think that philosophers are traditionally reluctant to take this option?
5. Are there any other ways of resolving the puzzle that you can think of? If so, what are the strengths and weaknesses of such an approach?
6. Given a perdurance model, *I* do not exist *now* (properly speaking): rather, I am a temporally extended object, with only a temporal part that exists now. Do you think that's a problem? Explain your answer.
7. Suppose time travel is possible and that bent blade of grass travelled back to when the blade of grass was straight. How would the two solutions proposed deal with that case? Can you see any potential problems with any of the solutions?
8. If properties are all relativized to times, it would seem to follow that objects cannot exist in the absence of time whilst also having properties. What do you think that we should say in response?
9. If you focus on people as persisting objects, does your view of which theory is best change?
10. How would you sum up the difference between the two views presented?

Reading

Kurtz, R. 2006. 'Introduction: What's the Problem?' in S. Haslanger and R. Kurtz (eds), *Persistence*, Cambridge, MA: MIT Press, 1–26 (see also many other excellent papers in this wonderful volume).

Lewis, D. 1986. *On the Plurality of Worlds*, Oxford: Blackwell, esp. 202–5.

Lowe, E. J. 1988. 'The Problems of Intrinsic Change: Rejoinder to Lewis', *Analysis*, 48(2), 72–7.

26 Possible worlds

If you read anything much at all in contemporary philosophy, you'll likely see mention of possible worlds. Though by no means a modern invention, possible worlds have taken centre stage in a number of metaphysical debates in the last few years. In outline, the idea is that every possibility is to be *literally* identified with some world; we are offering an analysis of possibilities in terms of worlds.

One very radical and well-known view is that of David Lewis. According to the Lewisian view, the term 'world' picks out objects that are spatio-temporally connected to one another. Thus, I'm a part of the same world as you; we're all a part of the same world as both Caesar and alpha Centauri. We're all spatio-temporally connected to one another. This world is the actual world. There are other worlds. That is, there are other collections of spatio-temporally connected objects, and since these worlds are not part of the actual world, they are not spatio-temporally connected to our own. These worlds are the merely possible worlds. Each world is a way that things could be; to every possibility there is a world. And since many things are possible, so there are many worlds. Indeed, there are probably infinitely many of them. This is his *modal realism*.

One way to get a feel for just how many possibilities Lewis thinks there are is to consider his principle of recombination: any two things that exist in this, the actual world, could be put together to create a new possibility – a new world. Thus, there is a possible object somewhere composed from my left leg being stuck to your nose. A terrifying thought!

Some claims about what's possible involve actual individual objects – tables, chairs, people – and what could have happened to them. For instance, even though I have two legs, it's possible for me to have only one leg. Now, recall, for Lewis possibilities are worlds and worlds are collections of objects that are spatio-temporally connected to one another. It would seem to follow that I exist at this world with two legs and that I exist at another world with one leg. How can this be?! I can't have both two legs and one leg! Nothing can manage that. This is sometimes known as the problem of 'transworld identity'.

Lewis's answer is to say that individuals exist only at one world. But true claims about them are grounded by counterparts of theirs that exist at other possible words. Thus, what makes it true that 'I could have had only one leg' is that there exists at some merely possible world, a counterpart of mine – where a counterpart is understood to be an entity very similar to me, though not identical to me – that has only one leg. This is Lewis' counterpart theory; though it's distinct from his view of what worlds are, it's an important part of his overall theory.

Thought experiment

> Suppose that Elizabeth wants to know whether it's *possible* for her to have a career as a musician. According to the analysis suggested, what matters to whether or not

> this is true is whether or not there exists a world *other than this one*, where someone *other than Elizabeth* has a career as a musician. Given that what matters to Elizabeth is whether or not *Elizabeth*, who exists at this world, can be a musician, how can that be right?

The worry described in the thought experiment is sometimes known as the Humphrey objection, so called because when first formulated it focused on a character called 'Humphrey'. The obvious way to respond to the objection is to remind ourselves of what the proponents of possible worlds are doing. They claim to be analysing possibility. Thus, what it is for it to be possible for Elizabeth to carry out some particular action is for her to have a counterpart to carry out that action, at some world other than this one.

It's also worth flagging here that not all theories of possible worlds are quite so extreme as Lewis's modal realism. Indeed, most philosophers who take seriously this talk of possible worlds opt for something else. Typically this will take the form of a belief in worlds as representations; most usually the representations are taken to be maximal, consistent sets of propositions. One of these sets represents how the actual world is; all of the others represent how the world could have been.

Questions

1. How plausible do you find the idea that there exist possible worlds, that are collections of physical objects, that are not spatio-temporally connected to our own? Explain your answer.
2. Sometimes people respond to modal realism with what Lewis called 'the incredulous stare'; the view is just too strange to accept. What is the best response to this 'incredulous stare' that you can think of? How would you critically assess your own response?
3. The thought experiment describes a worry with counterpart theory. Do you think that the worry is pressing? Explain your answer.
4. If we are forced to give up on counterpart theory, how do you think that we might avoid the worry about transworld identity that motivated it in the first place?
5. Do you think that the principle of recombination is correct? Why/Why not?
6. The view that possible worlds are sets of maximal, consistent propositions (aka: propositional ersatzism) was sketched: how plausible did you find that view? Can you think of any good objections to it?
7. According to some other views, possible worlds are maximal consistent sets of sentences. Can you think of any advantages this view (sometimes called 'linguistic ersatzism') has over propositional ersatzism?
8. If each possibility is a world, and a world is nothing more than a set of sentences, then do you think that there are enough sentences to describe all of the possibilities? If not, why might this be a problem?

9 According to some views, modal claims (that is, claims about what's possible and what's necessary) are false, but useful. How plausible does this view seem to you? Does the fact that this view denies the existence of possible worlds strike you as a theoretical virtue of the view?
10 What, if anything, would be wrong with saying that there are no possible worlds, but that modal claims were true regardless?

Reading

Bricker, P. 2008. 'Modality and Possible Worlds', in T. Sider, J. Hawthorne and D. W. Zimmerman (eds), *Contemporary Debates in Metaphysics*, Oxford: Blackwell, 111–34.
Daly, C. 2006. 'Mathematical Fictionalism: No Comedy of Errors', *Analysis*, 66(291), 208–16.
Melia, J. 2008. 'Ersatz Possible Worlds', in T. Sider, J. Hawthorne and D. W. Zimmerman (eds), *Contemporary Debates in Metaphysics*, Oxford: Blackwell, 135–52.
Sider, T. 2003 'Reductive Theories of Modality', in M. J. Loux and D. W. Zimmerman (eds), *The Oxford Handbook of Metaphysics*, Oxford: Oxford University Press, 180–208.

---

27 Properties and their bearers

Look at this book. It is a particular colour. But notice: that sentence seems to say two things.

- First, the sentence seems to suggest that there is a thing – a book.
- Second, the sentence seems to say that the book *is a particular way*, or that it has a particular colour property.

One very natural way in which to understand this sentence is as committing us to both the existence of books and colours. Thus, there are particulars and there are properties. The particulars can be said to be 'property bearers'.

Let's start with properties. It would seem obvious (on first inspection) that properties can be repeated. There are many objects in the world that are *exactly the same colour*. There are certainly many objects with exactly the same charge (electrons, for instance). We might say, then, that there are property bearers (each particular electron) and properties (the property of charge), and that the bearers bear (or 'instantiate') the properties.

But in that case, why think that there are objects *and* properties at all? Why not just take the view that objects are *bundles* of properties?

Thought experiment

> Suppose that properties are repeatable, such that two objects being red is a matter of their instantiating the very same properties. Consider, then, two large spheres: we will call them Castor and Pollux. Both spheres are the same shade of red. They are also the same diameter and mass. Indeed, both spheres have exactly the same intrinsic properties. There is no property that will allow us to differentiate them from one another. But now: suppose that the two balls are nothing more than bundles of properties. In that case, there is in fact no difference between them. And in that case, we should not say that we have two spheres, at all, we should say that there is only one! After all, there are principles of identity that tell us that if there is no difference *at all* between x and y, then x and y are in fact the same thing. But this seems wrong. After all, we began by conceiving of two distinct spheres, with the same properties. Something here has gone very strangely indeed.

There are a number of ways in which we might look to respond to this particular puzzle. The first thing that we might do is to posit property *bearers* in addition to properties. If we do that, then we have a difference between Castor and Pollux. Castor is a different property bearer than Pollux.

This does raise an awkward question, however, about the nature of property bearers. If they are to bear all of the properties, then they do not themselves *have* properties. They are featureless. Can we really make sense of featureless entities?

Instead, then, some have tried to respond by denying that properties are literally repeatable. Rather, what we have here are *particularized* properties. That is, Pollux will consist of the properties: Pollux's mass of 30kg; Pollux's diameter of 2m, etc. (The typical example to illustrate here is to contrast wisdom (the repeatable property) with Socratic wisdom, which is a property particular to only Socrates.)

Of course, this is still to commit to the reality of properties. Some views are more radical and try to argue that there are no such things, and that any appearance of similarity that objects have is just that: an appearance. For such views, often called 'nominalist' views, there are only particulars and no properties.

Questions

1. Write down a list of properties that your favourite book has.
2. Now try writing these as *particularized* properties.
3. In the thought experiment, do you think that we really can imagine two spheres? If so, what does this show?
4. Does the idea of a propertyless property bearer strike you as plausible? What might be the problems that such a view would face?

5 What problems might we run into if we deny that there are properties? For instance, would we have to deny that electrons have charge? Would that be a problem?
6 Suppose that there are properties and property bearers: what connects them?
7 Suppose that objects are just bundles of properties. What connects the properties to one another? Why don't the properties simply float apart from one another?
8 If there are repeatable properties (often called universals) it would be possible for one instance (the negative charge on one particle) to move towards another (the negative charge on another particle). Thus, we'd have an entity moving towards itself. Would that be a problem?
9 In the case of Castor and Pollux, what helped generate the problem was a principle of identity. Do you think that the principle is correct? Why?
10 Could we ever have empirical evidence that would help us settle the questions posed here? Does this matter?

Reading

Armstrong, D. 1989. *Universals: An Opinionated Introduction*, Boulder, CO: Westview Press.
Hoffman, J. and Rosenkranz, G. S. 1994. *Substance Among Other Categories*, Cambridge: Cambridge University Press.
Lowe, E. J. 2006. *Four Category Ontology*, Oxford: Oxford University Press.
Oliver, A. 1996. 'The Metaphysics of Properties', *Mind*, 105(417), 1–85.

---

28 Something rather than nothing

Consider this book. Why does it exist? Presumably, it exists because the authors got together and decided to make it so. With a lot of help from the publishers and referees, that led to the thing you hold in your hands. Well and good. But, just as we can ask this question of individual things in the world, so we can ask it of the whole cosmos itself. Thus, why is there anything at all? Why is there something rather than nothing? One answer is quite bold: the reason that there is something rather than nothing is that it's *impossible* for there to be a situation where nothing at all exists. Thus, there has to be at least *something*, because matters couldn't have been otherwise. This at least gives us *some* explanation of why there's something rather than nothing. But there are some problems:

Thought experiment

Suppose that objects o_1, $o_2 \ldots o_n$ exist. Each of these objects is contingent: it might not have existed. It follows, then, that it's possible for each of $o_1 \ldots o_{n-1}$ to exist. *That*

> world is a possible world. In crude terms: we can subtract an object from the world: the result is a possibility. Now of course we can iterate that process. We can continue subtracting. Thus, there is a world where only each of $o_1 \ldots o_{n-2}$ exists; there is a world at where only each of $o_1 \ldots o_{n-3}$ exists, and so on. It seems that we can carry on subtracting, right up until the result is that just o_1 exists. If we can subtract each of the other objects, and the result is a possibility, then it would *seem* to be the case that we could subtract o_1, too, and the result would be a possibility.

Thus, contra the claim we made above, it would seem that it is possible that no concrete objects (a philosophical term that means roughly the same as 'physical object') exist. Call such a possibility 'metaphysical nihilism'.

The argument loosely outlined in the thought experiment is known as the subtraction argument. And on first inspection it looks to be compelling. But there are some interesting responses available. For one thing, notice that there is a large difference between each of the steps that leads up to the subtraction of o_1 and the subtraction of o_1 itself. Each of the steps leading up to o_1 being the only object at a world trades upon the idea that each object is contingent. That thought is supposed to license the thought that we could then remove o_1, too. But merely that each object is itself contingent doesn't actually do anything to show that the last step in the argument is good. Each of those subtractions is carried out in the context of there being *some* existing objects. The induction – that we've been able to remove each of $o_2 \ldots o_n$ and have the remainder be a possibility, so that we can also remove o_1 and have the remainder be a possibility – doesn't tell us anything *at all* about cases where what we're talking about subtracting *is the last object*. It seems, then, that more might need to be said.

There's also a further metaphysical question that requires us to think quite hard about the nature of possibility itself. As we saw in the discussion of possible worlds (Chapter 26), Lewis thinks that possible worlds are collections of spatiotemporally connected objects. If there *are* no objects, then there is no world. If there is no empty world, then there is no such possibility. This, the view that there is an empty world simply makes no sense. (Although, equally, one might worry about Lewis's theory of possible worlds if it can't allow us this possibility.) This is another place in metaphysics where different areas intersect with one another – where a theory about the nature of one thing seems to have implications for another debate.

Questions

1. The subtraction argument is suggested by the thought experiment: do you find it convincing?
2. If each of o_1 to o_n is itself such that it might have failed to exist, do you think that this *suggests* that metaphysical nihilism is possible, even if there's no direct argument to the conclusion?
3. Can you think of any other argument to show that there might have been nothing?

4 What do you think of the reply to the subtraction argument, outlined above?
5 If it's impossible for there to exist a metaphysically nihilistic world, do you think that *really explains* why there is something rather than nothing?
6 How convincing do you find the worry that an empty world would not, by Lewis's lights, be a world at all? Do you think that, instead, this is a problem for Lewis's view of modality? Explain your reasoning.
7 What theory of modality would you suggest in place of Lewis's? How would that help with the subtraction argument?
8 What answer would you give to the question 'why is there something rather than nothing'?
9 Existence monism is the view that there exists only one (very large!) physical object, 'the world' (there are no people, no plants, no planets, etc.). Suppose that this view is true and explain how that bears on the question of metaphysical nihilism.
10 Does the question of why there is something rather than nothing matter? Explain your answer.

Reading

Baldwin, T. 1996. 'There Might Be Nothing', *Analysis*, 56(4), 231–8.
Cameron, R. 2006. 'Much Ado About Nothing: A Study of Metaphysical Nihilism', *Erkenntnis*, 64(2), 193–222.
Efird, D. and Stoneham, T. 2005. 'The Subtraction Argument for Metaphysical Nihilism', *The Journal of Philosophy*, 102(6), 303–25.
Rodriguez-Pereyra, G. 2004. 'Modal Realism and Metaphysical Nihilism', *Mind*, 113(452), 683–704.

---

29 Time travel and the grandfather paradox

Representations of time travel in science fiction are commonplace. But there are philosophical worries about time travel. In particular there are some philosophical concerns that time travel may be shown to be impossible. This would be a little surprising, given that some plausible scientific models of reality seem to imply that time travel is possible – even if we currently lack the means to take advantage of the fact. This lack of means isn't something that we worry about, either. What we are interested in is not whether it's within our current means to travel in time, but whether or not it is possible *in principle*.

When we talk about 'time travel' it seems that we have a very specific phenomenon in mind, though it's surprisingly hard to pin down the notion precisely. As in films like *Back to the Future* and programmes like *Dr Who*, what we seem to

require in order for there to be time travel are cases where a person or other object starts off at one time and then 'jumps' either forward or backward in time, arriving at a point in the (relatively) distant past or future, without having experienced the intervening period.

One of the threats to the possibility of travelling through time was suggested by David Lewis. Consider the following:

Thought experiment

> Sara spends most of her formative years fascinated by guns. Over many years she trains, methodically, until she's a crack shot. One day she meets Ellie. Ellie has a time machine. Ellie sends Sara back to the time at which her grandparents live. Sara – with a love of time travel paradoxes – takes careful aim at her grandfather. The weapon has been checked and double-checked, as has the ammunition. It's all perfect. The weather conditions are perfect, her grandfather is in a wide open space, and Sara is quite close to him – though he cannot see her. Can Sara kill her grandfather? It seems not! After all, if she did, then she would not have been born and so could not have travelled back in time to shoot him. But, all the same, it seems that she can. After all, there's nothing to stop her.

There are a couple of interesting features of this case that need bringing out:

- The first is that it really does seem to be a paradox: Sara seemingly both can *and* cannot kill her grandfather.
- The second is the way in which Lewis thinks that we should respond. It seems clear that there are two senses of the word 'can'. For instance, I *can* speak English: indeed, were you to hear me speak, it's likely that you'd hear me speak English. And although I have no knowledge of the language, I *can* speak Arabic. That is, I have all of the physical and biological tools that I would require to do so. But I don't speak Arabic. I have never learned.

That might be useful in the grandfather paradox case. Sara *can* kill her grandfather. That is, she has all of the physical and biological tools that she would require to do so. But she doesn't shoot her grandfather. After all, she is there to take aim. And had she been successful, that simply wouldn't have happened.

Questions

1. What do you think of the suggested response to the paradox?
2. Can you think of any differences between the sense of 'can' evoked by the thought experiment and the sense evoked by the discussion of learning a language?
3. Suppose that the meaning of 'can' was held fixed: why do you think it would be a problem if Sara both could and could not shoot her grandfather? What's wrong with saying that she could and could not do so?
4. How do you think we should respond to the thought experiment?

5 Does the lack of *actual* time travel make the debate pointless?
6 Can you give an analysis of what is meant by the term 'time travel'?
7 Consider the theories of time discussed on pp. 98–9 do you think that time travel is compatible with all of them? If not, does that raise any problems?
8 Think back to the theories of persistence discussed on pp. 57–9: do you think that time travel is compatible with all of them? If not, does this raise any problems?
9 Do you think that time travel *is* possible?
10 Would it be reasonable to respond to the thought experiment that Sara can travel in time, just not back within her own past?

Reading

Effingham, Nikk and Robson, Jon. 2007, 'A Mereological Challenge to Endurantism', *Australasian Journal of Philosophy*, 85(4), 633–40.

Lewis, David. 1976. 'The Paradoxes of Time Travel', *American Philosophical Quarterly*, 13(2), 145–52.

Sider, T. 2005. 'Traveling in A- and B-Time', *Monist* (special issue on time travel), 88, 329–35.

Smith, Nicholas J. J. 1997. 'Bananas Enough for Time Travel?', *British Journal for the Philosophy of Science*, 48(3), 363–89.

---

30 Time

Time is hard to think about. As Augustine of Hippo had it, 'What then is time? If no one asks me, I know what it is. If I wish to explain it to him who asks, I do not know.' Nonetheless, philosophers have tried to put forward theories about the nature of time. The issue that can sometimes seem to drive the debate is that of temporal passage and change. It seems to us as if objects are in a constant state of flux; that time is passing; that there is a constant *change in what there is*. It seems, then, that our theory of time ought to try to reflect and capture that.

But there are various other pressures on our best theory of time. Our best theory of time should be consistent with how we perceive time to pass. Our best theory of time should be consistent with what our best physics tells us about time. Our best theory of time should be an accurate reflection of how we speak about events, past and future. In particular, though, we often seem to return to the notion of passage. To get to the heart of the issue, consider the following:

Thought experiment

> Suppose that we could pause reality, much as we might pause a movie. This would give us a collection of physical objects, all frozen in space. Suppose that we call that a slice of reality. Then, suppose that we add to that first slice another, except this time

> with all (or most) of the objects moved slightly from their positions in the first slice. Then add another slice, again with all (or most) of the objects moved on a little further. Carry on adding slices until we have a collection of slices long enough to create a history of the Universe. Is there anything more that we need to add to this collection of slices, in order to capture the reality of time? It *seems* so. There is nothing in the block – the collection of slices – that seems to correspond to the passage of time.

There are a number of different responses to this kind of case. The first response is to note that there really *isn't* much that needs to be added to the case. All that we need to do is to add some temporal relations into the mix. Thus, what we require is that each slice is arranged such that it is 'earlier than' or 'later than' some other slice. These temporal relations, that relate the various slices, give rise to a 'block Universe' picture, where time is real.

But not everyone agrees that the addition of the earlier than/later than relation (sometimes called a B-relation, due to language given to us by McTaggart (1908)) is enough for the reality of time. For in the picture so described, there doesn't seem to be any passage. One way in which to add what's missing, or so they say, is to add a property of 'presence' to the picture. This property moves along the block, first illuminating one slice, and then the next. This moving-spotlight view of time treats the property of 'presence' (sometimes called an A-property) as giving rise to the reality of time.

Indeed, not everyone thinks that we need to talk of a block at all. Some – presentists – typically agree that there is a property of presence, but deny that all of the different slices exist. Presentists think that only present objects exist, and that those objects have the property of presence, and that there is a continual change in which objects exist and have that property. Others think that the past and present objects exist, but that the future objects do not. These 'growing block' theorists think that the passage of time is constituted by the addition of present facts to the block of the past.

Questions

1. Have a go at describing what you think of as being the passing of time. (The notion is very familiar, but very hard to describe.)
2. Do you think that the proponent of the B-relation does enough to capture that sense of time's passing?
3. The moving-spotlight view of time has a property of presence move along the block. Most (all?) movements take time. Are there any difficulties in describing this 'movement' of the present as a movement?
4. Presentists think that only present objects exist. How do you think that they should try to capture the notion of time's passing? Do you think that they do enough to capture it?
5. Do you think that too much emphasis is being put on temporal passage? Explain your answer.

6 It's a common view in philosophy, that truths require grounding: that a truth about p is true if and only if p exists. But both the presentists and growing block theorists deny that there are future objects. Does that mean that they should deny that there are truths about the future? If they do, is that a cost or a benefit of their view?
7 Some people think that there's an asymmetry between the past and the future: where the past is fixed, the future is open. Which view (if any) do you think best captures the idea and how much does it matter?
8 Presentism denies that past objects exist. Is presentism therefore incompatible with time travel? Does it matter?
9 Which theories of time are compatible with which theories of persistence (see Chapter 25)? If there are any incompatibilities, then does this raise any problems or advantages for particular views?
10 All of this discussion has presupposed that time is real: do you think that's a correct assumption?

Reading

Markosian, N. 2004. 'A Defense of Presentism', in D. W. Zimmerman (ed.), *Oxford Studies in Metaphysics*, Oxford: Oxford University Press, vol. I, 47–82.

McTaggart, J. E. 1908. 'The Unreality of Time', *Mind*, 17(68), 457–74.

Price, H. 2011. 'The Flow of Time', in Craig Callender (ed.), *The Oxford Handbook of Time*, Oxford: Oxford University Press, 276–311.

Savitt, S. 2002. 'On Absolute Becoming and the Myth of Passage', in Craig Callender (ed.), *Time, Reality and Experience*, Cambridge: Cambridge University Press, 153–68.

Part IV

PHILOSOPHY OF MIND

31 Chinese room

The idea that computers (or computer programs) can think is intuitively seductive. It becomes ever more so with the increasing complexity and sophistication of the tasks that such systems can accomplish. And with that thought comes an equally natural thought about the human mind and its consciousness: doesn't it make sense to treat the mind and its consciousness as a computer? A very complex and sophisticated computer, of course, and not one that runs on familiar silicon hardware, but a computer nonetheless.

The following thought experiment, due to John Searle, is intended to put pressure on this picture:

Thought experiment

> Suppose that we place an English speaker, who knows no Chinese, into a locked and sealed room. He is kept fed and watered. In the room with him is a manual: it tells the man that when a particular Chinese symbol is passed into the room via a small postbox, he is to pass out a given symbol in reply. The manual specifies a whole range of outputs in response to the symbols that come in. People outside the room end up believing that they are conversing with a Chinese speaker. But, to repeat, our English speaker knows no Chinese and there is no translation going on. The man is merely following the instructions in the manual and matching appropriate ciphers.

It seems wrong to say that the man understands Chinese. If we asked him a question in Chinese, he wouldn't have the first idea how to respond. Indeed, take away the manual and he wouldn't know how to respond to the symbols being passed into the room. But, in that case, it seems that there's no understanding in the case described. And since people do understand things, this kind of system isn't one that can be anything like a person.

But, of course, a computer (or computer program) is very similar to this: signals are input and, in accordance with a pre-programmed set of rules, signals are output. Here's a nice quote from Searle:

> Computational models of consciousness are not sufficient by themselves for consciousness. The computational model for consciousness stands to consciousness in the same way the computational model of anything stands to the domain being modelled. Nobody supposes that the computational model of rainstorms in London will leave us all wet. But they make the mistake of supposing that the computational model of consciousness is somehow conscious. It is the same mistake in both cases.
>
> (Searle 2002: 16)

There are, of course, various responses to this and it's fair to say that the thought experiment is extremely contentious. A couple of the replies centre on the following kinds of issue. In what follows, though, we'd do well to keep in mind that this

discussion is not concerned with the tough question of whether or not machines like computers can think. Rather, the focus here is supposed to be on whether or not running a program analogous to that described in the thought experiment is sufficient for thought and understanding.

- Complexity: the case given by Searle is very unlike the brain because it is so simple. The brain involves billions of neurons. This case involves just one person. It's not possible to derive a conclusion like this given the differences in complexity.
- Location: Searle's right that the person in the room doesn't understand Chinese. But that's the wrong question. The right question to ask is whether or not the system that includes the manual, room and person understands Chinese. That seems a more plausible option.
- No consciousness: according to one extreme response, there's no such thing as consciousness anyway. In that case, there's nothing missing from the picture. According to these eliminative materialists, Searle is looking for something that simply isn't there.

Questions

1 People tend to have quite strong intuitive responses to the Chinese room case. What were yours?
2 What did you think of the complexity response?
3 What did you think of the location response?
4 What did you think of the 'no consciousness' response?
5 Would it make any difference to your view if the 'system' were able to move around and investigate its environment (like a robot might)?
6 Can you think of another response? If so, how does yours compare to the ones just mentioned?
7 Suppose that we stripped out parts of the brain, over time, and replaced them with nano-processors: could a person survive one such replacement and still be conscious? Two? All of them?
8 Do you think that future advances in computing could have any bearing?
9 Do you think that computer programmers should care about this result? Why?
10 The case involves our intuitions about what can result from complex computation in the brain. Given the degrees of complexity involved, do you think that those intuitions are reliable? (Compare: do you have intuitions about what could arise from 15 billion atoms being thrown together in particular ways?)

Reading

Churchland, P. M. and Churchland, P. S. 1990. 'Could a Machine Think?', *Scientific American*, 262(1), 32–7.

Dennet, D. 2013. *Intuition Pumps and Other Tools for Thought*, New York: W. W. Norton.

Searle, J. R. 1983. 'Can Computers Think?', in D. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*, Oxford: Oxford University Press, 669–75.
Searle, J. R. 2002. *Consciousness and Language*, Cambridge: Cambridge University Press.
Sharvy, R., 1985, 'It Ain't the Meat It's the Motion', *Inquiry*, 26(1), 125–34.

32 Artificial intelligence

As we saw in the last section, some are opposed to the idea that computers can think. Indeed, the very idea of an artificial intelligence seems (to them) to be a little misguided – at least in its current form. But not everyone is so disposed. Some are of the view that were computers able to manifest particular sorts of behaviour, then that would give us good reason to think that there was genuine intelligence present. The following thought experiment goes some way to illustrating those conditions:

Thought experiment

> Sally goes into a room. In the room she finds a computer with a chat-screen open. She notices that there are two named users available to chat with. She spends a while swapping messages with one of them – 'Bill'. She learns a bit about Bill and they spend some time swapping stories. Bill then signs out. Sally then engages in a brief conversation with the other user – 'Maggie'. Sally chats for a similar length of time with Maggie. Again, she learns a bit about Maggie, and they swap some stories. There is nothing unusual about either of the conversations. As far as Sally is concerned, both of her conversational partners are normal human beings. However, it turns out that, unbeknown to her, there is no Bill. In fact, all of Bill's responses have been generated by a computer. But despite having 15 minutes of discussion with 'Bill', Sally didn't notice this.

Some think that, were a computer to behave in this way, it should be regarded as being intelligent. After all, the computer has effectively behaved *as if* it were an intelligent human agent for 15 minutes. Behind the scenes, here, is the famous 'Turing test' – the idea being that any computer capable of fooling someone into believing that she is conversing with a human being is a computer that we might describe as being genuinely intelligent.

One natural response to this might be that what matters isn't the external behaviour of the computer (including its ability to maintain a conversation) but its *internal* workings. Is the computer *really* thinking? That's what we should want evidence for. But what kind of evidence would we want? In the case of human beings, the only evidence that we have that *they* are intelligent is their external behaviour. It's not clear that we should put the bar any higher in the case of artificial intelligence.

A slightly different response might then be one that suggests that what matters is the total replication of human neurological behaviour. In the Turing-test type of

case described, we're seeing such a limited subset of intelligent behaviours that we really don't have enough to go on. Once again, though, we would presumably be happy enough with Sally's judgement that Bill *is* an intelligent person, just based upon the interactions she had with him in the chat-room. As such, this response seems at least a little lacking.

Questions

1 Do you think that a computer passing the Turing test would be intelligent? Why/why not?
2 What else might we ask a computer to do to demonstrate that it is intelligent?
3 Do you think that it would make any difference to the test if the participant (Sally) was aware that one of her interlocutors was a computer? Why?
4 How much might it matter how clever the people are in Sally's role in the test? Are there traits other than intelligence that might be salient?
5 Should we care about the Turing test? When making planes we don't look to mimic birds. So why, when looking to make intelligence, should we look to mimic humans?
6 Historically, people have identified many different objects as having intelligence (sometimes for religious reasons; see Chapter 38 on panpsychism). Rocks, trees, clouds, and so on, have all been thought to have intelligence. Against that backdrop, of what value is the Turing test?
7 What about a change to the test: suppose that we ask a computer to differentiate, after conversation with another computer and a human being, as to which was which. Would a computer passing that test be intelligent?
8 How much difference would it make if the interaction was with a lifelike robot, that moved and spoke in such a way as to not be distinguishable from a human?
9 Would a genuinely intelligent computer have to be capable of engaging in emotional behaviour?
10 Is human intelligence the only kind of intelligence?

Reading

Boden, M. (ed.) 1990. *The Philosophy of Artificial Intelligence*, Oxford: Oxford University Press.

Copeland, B. J. 1998.'Turing's O-Machines, Searle, Penrose and the Brain', *Analysis*, 58(2), 128–38.

Copeland, B. J. and Proudfoot, D. 1996. 'On Alan Turing's Anticipation of Connectionism', *Synthese*, 108(3), 361–77.

Penrose, R. 1990. Précis of 'The Emperor's New Mind', *Behavioral and Brain Sciences*, 13(4), 643–55

33 The mind and the brain

Dr Tallant and Jonathan are, in fact, the same person. They have all of the same properties. These two facts about them seem to be intimately linked. Things that are identical have all of the same properties. This principle can then be put to work in a very famous case involving the mind and the brain.

For several centuries, the prevailing view was that the mind is distinct from the brain. The brain is physical, the mind is mental. But as our physics improved, it became clear that interaction between the physical and the mental would be problematic. For one thing, where would it occur? What could we identify as the locus of contact between the immaterial mind and the material body? For another thing, physical systems seem to conserve energy: none is lost. Interactions require energy transfer; so why don't we see a loss of energy in the interactions between the physical and the mental? This has led to a steady move away from treating the mind and body as distinct and towards a form of physicalism where we try to identify the brain with the mind or else reduce the mind to the brain – in some way or other. Nonetheless, there are still some arguments that can give some pressure to go the other way, and to treat the mind and brain as distinct. Here, we'll run together two of these in a case:

Thought experiment

> Sam is a keen experimentalist who wishes to know everything that there is about minds, brains and thoughts. In order to do so, Sam sets about carving up a region of the brain. As Sam is doing so, Mark walks into the lab and asks: 'all very well cutting up the brain, but if you want to know about these thoughts, why don't you try cutting up one of those?' Sam laughs, but then pauses: how could Sam cut up a thought? The very idea doesn't seem to make any sense. Thoughts *can't* be cut up. But brains can. Sam begins to think: what else could minds or thoughts be capable or incapable of that brains can do? One thing strikes her immediately: it seems to her that her mind might be able to survive the death of her body. After all, Sam thinks, although maybe it doesn't happen, plenty of people claim to have imagined the idea of an afterlife. They could only do that if they could imagine minds surviving bodies and if we can imagine something like that happening, that gives us some reason to think that it's possible. Disheartened, Sam trudges from her lab, across campus, and into the Philosophy class.

There are a lot of different ideas in here:

- First, the fact that brains but not thoughts are divisible seems to suggest that they are different things: after all, if we could cut up Dr Tallant, but not cut up Jonathan, then that would seem to suggest that Dr Tallant and Jonathan are not the same sort of thing.
- Second, there was the contentious claim that Sam can imagine surviving her bodily death and that this gives her some reason to think this could happen.

> From there, of course, it was just a short step to saying that Sam's mind and brain must be distinct: after all, Sam's body cannot survive her bodily death. If her mind can, then her mind is not identical to her brain.

These two arguments are, roughly, lifted from Descartes, and are sometimes known as the divisibility argument and the conceivability argument (respectively).

Questions

1. Do you think that Sam is right that brains can be divided, where thoughts cannot?
2. What, if anything, does your answer to (1) tell us about whether the mind is the brain?
3. The divisibility argument relies upon the premise that identical things have all of the same properties: do you think that premise is true (compare with the case involving Dr Tallant and Jonathan)?
4. Do you think that you can imagine surviving your bodily death? How does that bear on the conceivability argument?
5. *If* someone could conceive of surviving their bodily death, do you think that this tells us anything about whether or not it's possible for this to occur? How does this bear on the conceivability argument?
6. We can often work out what it's possible for us to do: I know that it's possible for me to finish writing this sentence. How are we working that out? How does this bear on the conceivability argument?
7. If minds aren't brains, what are they?
8. Write down a list of properties that you think (intuitively) that thoughts have. Write down a list of properties that you think (intuitively) that brains have. If they're not the same, explain how one thing (a brain) can have such different properties.
9. Do you think that considerations of 'the soul' are relevant here? Why?
10. Are minds identical to brains?

Reading

Chalmers, D. 2002. 'Does Conceivability Entail Possibility?', in J. Hawthorne and T. Gendler (eds), *Conceivability and Possibility*, Oxford: Clarendon Press, 55–100.

Churchland, P. M. 1984. *Matter and Consciousness*, Cambridge, MA: Bradford Books, esp. chapter 2.

Lowe, E. J. 2000. *An Introduction to the Philosophy of Mind*, Cambridge: Cambridge University Press, esp. chapters 1 and 2.

Lowe, E. J. 2003. 'Physical Causal Closure and the Invisibility of Mental Causation', in S. Walter and H. Heckman (eds), *Physicalism and Mental Causation*, Exeter: Imprint Academic, 137–54.

34 Eliminative materialism

As you read this, it will likely seem to you as if you are having distinctive mental experiences – a feeling *of what it's like* to do the reading. It will also likely seem to you that you are forming beliefs about what you're reading. It seems to be an essential part of what it's like to be a person that you experience the world in this way.

But some philosophers of mind – the eliminative materialists, most notably Patricia and Paul Churchland – have challenged this idea. To be an eliminativist about some particular subject matter is to *eliminate* that subject matter; to say that the putative subject matter is not real. And, say the eliminativist materialists, we should eliminate the category of the mental: there are no thoughts, there are no feelings. In short, there is nothing other than the material.

Thought experiment

> Joey is scientifically and historically minded. Joey notices that the history of science is populated with a number of cases with the following structure: people began with a 'folk' understanding of something – physics, meteorology, mental illness, etc. This folk understanding turned out to be pretty useless. In the meteorological case, sacrificing animals to Zeus didn't really make it rain. Ultimately, it turned out that there was simply no place in our meteorological theories for gods. This being so, the folk theory had to be rejected. Likewise in cases involving mental illness. Folk theory would have us believe that such things are caused by demons. But there is no place in our best theories for demons. Such things should be eliminated. Joey then focuses upon the mental. We have a folk theory that is entirely unregimented by science. We may *talk* of beliefs, feelings, and the like, but these are hardly scientific categories of being in the way that 'electron' or 'proton' are categories of being. On the basis of an inductive argument, Joey then decides that it's highly likely that our 'folk psychology' is false, too, and that we should eliminate the category of the mental from our best theory.

In the case described, Joey engages in an inductive argument. Inductive arguments are, though certainly fallible, the kinds of argument that we often find credible. (If I told you that the Sun had risen every day for the last several hundreds of thousands of years, you'd probably agree that we should therefore think that it will rise tomorrow.) This being so, though this argument might be able to be defeated, it's certainly one that we should take very seriously.

One might want to be an eliminativist about the mental for other reasons, too. For instance, we may think that we have beliefs. But beliefs are propositional in structure. They are entities with *semantic* and *syntactic* structure and content. But where in the head will we find *semantic* or *syntactic* structure? Neurological processes just aren't like that. We don't literally *have* sentences written into our neurons.

In reply to the eliminativist, we might note that this kind of concern need not rule out all mental kinds. We might eliminate thoughts, for instance, but not *feelings*. That I am in pain, for instance, seems to not be something that we can doubt. But it should be noted that things here are complex. What we think of as pain seems to most of us to be such that it is essential to the experience being painful that it be unpleasant. But there are those (patients with certain kinds of psychological conditions or who are on morphine) who will report the experience as not being unpleasant. What the materialist will likely say here is that this shows us that our very firm beliefs about mental kinds are based upon mere illusory impressions and that they cannot be trusted.

Questions

1. Write out all of the different kinds of mental activity that you can (thinking, believing, etc.).
2. Do you think that some of these kinds are easier to eliminate than others? Why?
3. What did you think of the inductive argument presented in the thought experiment? Is there a potential way to improve it?
4. Do you find eliminative materialism plausible?
5. Did you find the argument about beliefs (that we do not find the right kinds of structures in the brain) to be convincing? Why?
6. What would a good argument against materialism need to show in order to persuade you? Explain your answer.
7. What would a good argument *for* materialism need to show in order to persuade you? Explain your answer.
8. Materialists claim that there are no thoughts. How does this differ from the reductive claim that thoughts can be identified with portions of the brain? Does this elimination strike you as in any way contradictory?
9. What do you think is the biggest challenge facing the materialist?
10. How does materialism connect up to other theses? For instance, are there implications for views about free will, moral responsibility, etc.?

Reading

Churchland, P. M. 1981. 'Eliminative Materialism and the Propositional Attitudes', *The Journal of Philosophy*, 78(2), 67–90.

Churchland, P.S. 1994. 'Can Neurobiology Teach Us Anything about Consciousness?', *Proceedings and Addresses of the American Philosophical Association*, 67(4), 23–40.

Dennett, D. 1988. 'Quining Qualia', in A. Marcel and E. Bisiach (eds), *Consciousness in Contemporary Science*, New York: Oxford University Press, 42–77.

Jackson, F. and Pettit, P. 1990. 'In Defense of Folk Psychology', *Philosophical Studies*, 59(1), 31–54.

35 Emotions

Lots, if not all, of the different experiences that we have involve emotions. I can have a fear of failure, the love of my family and the sorrow of departing. But the philosophical analysis of these emotions is less clear. There are a number of different questions that can be considered. By way of a non-exhaustive example: we can ask how emotions came to be, whether they should be subordinate to reason, and even whether or not they're real. But our interest, here, is in what they are.

On one (common sense) view, emotions are awareness of particular physiological responses:

Thought experiment

> Whilst out trekking, Annie sees a bear. The bear spots her and pauses. Annie freezes. She can see the bear staring at her. Annie can feel her heart beating more and more quickly and more and more strongly. She becomes hyper-aware of her limbs. She can sense the muscles tightening in her stomach and the adrenaline flowing into her muscles, preparing her for flight. She is terrified. Annie's recognizing these physiological changes *is* her terror. The bear turns its head to one side and then walks away. Annie sighs in relief and makes her way away from the area.

It seems right to ask, of the case described, what, over and above the feelings described, an emotion *could be*. There might not seem to be a lot missing. But there are at least two concerns with this as an analysis of emotions, so presented:

- First, it might seem that we have a problem with many different emotions. It might seem that fear, terror, nervousness, and many other emotions would all present in the same kind of way; that each of them would share common physiological traits. If the view just described were correct, and each emotion were simply to be identified with the recognition of the physiological change, then each of these apparently distinct emotions would end up being the same. That seems like the wrong result.
- Second, there is an intuitive difference between a *mood* and an emotion. Feeling depressed, or low, seems to be very different from experiencing a particular emotion – such as being afraid of something. One thing that we might want to say, then, is that emotions are *directed* towards particular thoughts and, as such, have a cognitive component. Thus, the thought is, Annie fears *that the bear will eat her*. Annie is then later relieved *that the bear has gone*. And so on. If that's right, then we're going to have to say more about emotions than that they're just to be associated with some physiological state.

Questions

1 Do you think that we could identify different physiological states for all different emotional states?
2 Do you think that there is a difference between emotion and mood?
3 Do you think that emotions have the cognitive element described? Do you think that we can be afraid, without being afraid *of* some particular thing?
4 Would it follow from emotions having a cognitive element that young infants couldn't have emotions? Would the theory be plausible if it said they couldn't?
5 Think about an emotion (e.g. love). Do you think that we should look to identify that emotion with a *single* state, or with a range of tendencies or dispositions to behave in a given way? For instance, is being in love *just* a matter of looking doe-eyed at someone, or are there different expressions of love in different circumstances?
6 Given your answer to (5): are there any implications as to whether we should opt for one of the particular theories of the nature of emotions, already discussed? How does this all connect up to emotions and film, as discussed in Chapter 86?
7 Consider the fear of flying. Most naturally, the propositional or cognitive content involved would be the thought that flying is likely to lead to death. But I can be afraid of flying even though I am aware that flying is *not likely* to lead to death. How might we deal with this concern?
8 So far, nothing has been said about how emotions help us see the world. But we speak of things as *looking terrifying*, and similar. Might there be a parallel between emotions and perception? If so, how might we try to draw it out?
9 Can you think of a potential practical consequence of our being able to give an analysis of emotions?
10 What connection is there – if any – between rationality and the emotions?

Reading

D'Arms, J. and Jacobson, D. 2003. 'Emotion and Action', in Robert C. Solomon (ed.), *Thinking About Feeling: Contemporary Philosophers on Emotions*, Oxford, New York: Oxford University Press.

Debes, R. 2009. 'Neither Here Nor There: The Cognitive Nature of Emotion', *Philosophical Studies*, 146(1), 1–27.

Scherer, R. K. 2005. 'What Are Emotions? And How Can They Be Measured?', *Social Science Information*, 44(4), 695–729.

Solomon, R. 1980. 'Emotions and Choice', in A. Rorty (ed.), *Explaining Emotions*, Los Angeles, CA: University of California Press, 251–81.

36 Externalism

The contents of your thoughts are in your head. So much seems obvious. When I think about something, I use my mind; the contents of my thoughts are located in my mind. And if we're of the view that minds are (or are intimately related to) brains, then this likely means that the contents of my thoughts are within my brain.

But there are challenges to this natural thought. These challenges often use a form of thought experiment to motivate the idea that some mental content is located *outside* your mind:

Thought experiment

> Prior to 1750, no one knew that water was in fact H_2O. Imagine, then, that in 1749 there exist both Earth, and Twin Earth. Twin Earth is almost exactly the same as Earth. The only difference is that on Twin Earth there is no H_2O. In its place is a clear, colourless liquid that has the same functional properties as water: we can call this XYZ. Suppose that Erika (who lives on Earth) utters the sentence 'water quenches thirst'. What she says is, intuitively, true. Now consider Angela who lives on Twin Earth (but is otherwise a near duplicate of Erika). When Angela says 'water quenches thirst' she is expressing the belief that XYZ quenches thirst. This is a different belief from Erika's, which is a belief that includes H_2O in its content.

The idea, then, is that because H_2O and XYZ are different substances the respective thoughts had by Erika and Angela are different. But this is despite the fact that the thoughts had by both Erika and Angela *seem* to be the same to both of them. It is also the case (we can specify) that Erika and Angela have the same neurological make-up.

Thus, the two thoughts are different (Erika thinks that H_2O quenches thirst; Angela thinks that XYZ quenches thirst), but the only difference between them is to be found in the external world – in the referent of the term 'water'. Thus, it must be the case that some aspect of mental content is located in the external world.

There are, of course, opponents of this kind of view. Some may be inclined to object to the interpretation of the thought experiment. They may object that, in fact, it's *not* true that the term 'water' expresses a different thought for Angela than it does for Erika. In both cases, the content of the thought is exactly the same.

Of course, if one were to accept that mental contents could (in principle) be located in the world around us, then it would seem to follow that we *might* be able to understand a variety of different mental contents in this way: thoughts, memories, perceptions, and so on, could all be held to be (in some sense) located in the world around us.

Questions

1 What is your intuitive response to the thought experiment?
2 Describe the thought experiment to friends who do not study philosophy: do they agree that Angela and Erika have different beliefs?
3 Does the answer to (2) matter?
4 It seems true that Angela and Erika will behave in identical ways in their two worlds. Doesn't it then seem a bit odd to say that they have different beliefs?
5 Do you think that XYZ is water?
6 Do you think that there is a sharp cut-off between what's 'internal' to the mind or brain, and what's 'external'? If not, is this a problem?
7 Does it seem to you that there might be two kinds of content here: an internal psychological content and an external, linguistic content? If so, how might that help with this puzzle?
8 There is no Twin Earth. Does this matter?
9 Why care about this case?
10 Might it be plausible that only some thoughts have external content?

Reading

Brueckner, A. 2002. 'The Consistency of Content-Externalism and Justification-Internalism', *Australasian Journal of Philosophy*, 80(4), 512–15.

Burge, T. 1979. 'Individualism and the Mental', *Midwest Studies in Philosophy*, 4(1), 73–121.

Crane, T. 1991. 'All the Difference in the World', *The Philosophical Quarterly*, 41(162), 1–25.

Wikforss, A. 2001. 'Social Externalism and Conceptual Errors', *The Philosophical Quarterly*, 51(211), 217–31.

---

37 Free will

Arguments about whether or not people have free will are old and familiar. They are also hard to get precise. The central idea, typically, concerns whether an individual has sufficient control over her actions, such that, although she in fact performed a particular action, she did not have to; she could have chosen to do otherwise. To focus our deliberations, consider the following (Locke-inspired) thought experiment:

Thought experiment

> Alex is in a room and the door is shut. Alex is asked to make a decision as to whether or not to leave the room. Alex chooses to stay. However, had Alex chosen

> to leave the room, then (unknown to Alex) via a small chip implanted in Alex's brain, a signal would have been sent in order to reverse the decision. So, had Alex decided to leave, that decision would have been immediately reversed and so Alex would not have been able to leave. Nonetheless, Alex did in fact make the choice to stay in the room.

There are a variety of different theories that will give us different answers to the question of whether or not Alex is free.

- According to some – determinists – all of our actions are predetermined and in fact there is only the illusion of free will.
- At the other extreme are those who claim that free will is very real. Some people, in this camp, are volitionists. They will likely say that Alex is free in the case just described. After all, he was able to make a decision based upon his own volitions – his own beliefs and desires. Of course, ultimately that might not lead to Alex being successful in acting in accordance with his desires. But that's OK. Even if we were free, not all of our actions would be successful.
- In the middle, between these two extremes, are compatibilists. There are a variety of forms of the view. However, the rough idea is that we should look to make two ideas compatible with one another: that our actions are determined by what occurred in the past; that we are free. On the face of it, this seems tricky. But suppose that we think of someone as being free if they can do whatever it is that they want to, based upon their desires and volitions. And suppose that we also think that people are fully determined: that is, which desires and volitions they have can be perfectly predicted, based upon past events and laws of nature. In that case, we would be free (in the sense just outlined), even though we are fully determined.

As this brief survey indicates, there's room for a lot more refinement and subtlety in how we think of free will, and how we might then set about analysing whether or not we really are free.

Questions

1. In the thought experiment, do you think that Alex is free? Explain your answer.
2. Of the views quickly sketched, which account of free will do you think is most plausible?
3. What do you think would be the most pressing challenge to the view of free will that you identified in your last answer? How would you try to deal with that challenge?
4. Given the accounts of free will described above, do you think that we have free will?

5 Suppose that we're not free. Would this make a difference to how we think about moral responsibility? After all, don't we think that you can only be held responsible for an action where you had freedom to do otherwise?
6 What kind of impact might neuroscience have on this debate? What kind of evidence should we be looking for?
7 Do different theories of time make a difference to how we think about free will? According to some views, after all, the future (in a sense) exists. Can we still be free in that situation?
8 Does it matter whether or not we have free will?
9 What kind of impact might science (other than neuroscience) have on this debate? What kind of evidence might we look for from the sciences to help us work out whether or not we're free?
10 Why do you think that the question of free will, despite having been around for thousands of years, is so interesting to people and so hard to answer?

Reading

Dennett, D. 1984. *Elbow Room: The Varieties of Free Will Worth Having*, Cambridge, MA: MIT Press.

Eckstrom, L. 2003. 'Free Will, Chance, and Mystery', *Philosophical Studies*, 113(2), 153–80.

Mele, A. 2006. *Free Will and Luck*, Oxford: Oxford University Press.

Van Inwagen, P. 2001. 'Free Will Remains a Mystery', in R. Kane (ed.), *The Oxford Handbook on Free Will*, New York: Oxford University Press, 158–79.

---

38 Panpsychism

As a starting point, it would be very sensible to suppose that people have minds. We may also want to allow that certain higher mammals have minds. But it would seem ludicrous to extend this too far. We would not, for instance, want to extend mindedness to plants. And we would certainly not naturally imagine that *rocks* are minded.

But there are those who defend a view – called panpsychism – that takes the line that *mind* is a fundamental kind, and that it is to be found throughout the Universe. As with many schools of thought, there are a variety of versions of the view. We will focus on just one version here and treat panpsychism as the view that every object that exists is mental, or intrinsically conscious. Thus, tables and chairs, electrons and flares. All of these things are mental and conscious. Why believe this very strange-sounding view?

Thought experiment

> Marsha believes that she has a mind. She's convinced of it, in fact. But she takes the view that rocks do not. She's uncertain about cats, dogs and other animals. Are they properly described as conscious beings? She's just not sure. However, Rashid challenges her to explain where the physically salient difference appears. After all, people are very similar (cognitively) to certain higher mammals. Higher mammals are very similar to other mammals. We can keep going. There are chains of similarity all the way through to single-cellular organisms and plants. There are also forms of similarity between multi-cell organic life-forms and complexes of inorganic matter – like rocks. So, where *exactly* is the difference that allows for consciousness? Marsha concedes that she cannot find one. But she's conscious. If she is conscious, and she cannot find a legitimate difference between her and all of the other entities that she's considered, then, she feels, she is compelled to allow that all of these other entities are conscious (or, we may say, minded). She grants that this sounds absurd but looks to deflect this concern by noting that the mental lives of such entities will be much *less* complex and involved than the mental lives of human beings.

It is fair to say that, although defended, panpsychsim often strikes people as a very odd view on first inspection, Marsha's reasoning notwithstanding. Something that people have said in its favour is that it allows us to posit minded agents without having to insist on an arbitrary cut-off point at which consciousness arises. Another point in favour of the view is that it does not require us to posit an emergent mental 'stuff' that is somehow dependent upon the physical, whilst being crucially different from it. Nor does it require us to explain how the mental and physical can interact. For, once again, everything is mental and everything is conscious.

Questions

1. Do you think that there's a cut-off point at which we should say that objects are conscious? Where would you draw that line? Why there?
2. Is pansychism just too strange to be believed? Is that a good argument against a philosophical view?
3. Do you think that it makes sense to say that objects like electrons have mental characteristics? Why?
4. Historically, people have believed that all sorts of objects have minds and/or souls. Is that at all relevant here?
5. What do you think is more plausible: the view that everything is mental, or that nothing is?
6. Some panpsychists have argued that paradigmatically physical objects seem to exhibit mental features; that is, they seem to behave as if they're directed towards particular goals or ends. Does that help to motivate panpsychism?

7 If people are minded and animals are minded in a different, less complex way, ought we to say that Siamese twins are also minded in a different way from other people? They are, after all, differently complex.
8 What should we say about artificial intelligence, given panpsychism?
9 Why might we care about this debate?
10 Are there any theological or moral ramifications to panpsychism? Do these matter?

Reading

Chalmers, D. 1996. *The Conscious Mind*, Oxford: Oxford University Press.
Goff, P. 2009. 'Why Panpsychism Doesn't Help Explain Consciousness', *Dialectica*, 63(3), 289–311.
Skrbina, D. 2005. *Panpsychism in the West*, Cambridge, MA: MIT Press.
Strawson, G. 2006. 'Realistic Monism: Why Physicalism Entails Panpsychism', in A. Freeman (ed.), *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism?*, Exeter: Imprint Academic.

---

39 Qualia and the mental

A very famous set of arguments in the philosophy of mind try to motivate a belief in the existence of something other than the physical by consideration of qualia – very roughly, the sensation of *what it's like* to have a particular experience. Again very roughly, the idea is supposed to be that in having the experiences that we do, we are coming to know something that seems to be distinctively non-physical. The thought experiment intended to illustrate this idea here is also very famous, and is due to Frank Jackson:

Thought experiment

> Mary has been brought up in monochrome surroundings: everything in the room where she lives (and works) is black and white. Mary is also an expert neurologist and, over the course of her now extensive education, she has learned everything that there is to know about colour-science. That is, she knows all of the different wavelengths for the colours. She understands everything that there is to be understood about how such light is processed in the brain. Indeed, so far as she is concerned, Mary knows every physical fact about colours that there is to be known. Then, one day, Mary steps outside the room and sees something red. It seems to Mary that she has learned something new: she now knows what it feels like to experience redness.

This thought experiment then makes trouble for those who believe that everything mental is physical. For if we agree with Mary that she knows all of the physical facts that there are to know about colour, and then, through having a new experience she learns something *new*, then by definition that new thing is not physical. Thus, *what it's like* to have the experience is not physical, but something else – most plausibly, something mental.

There are a number of responses to the case. For instance, some are inclined to say that Mary *does not* learn something new. Rather, what she does is acquire new abilities: that is, Mary is now able to imagine, remember (and so on) the experience of what it's like to experience redness. But that is not quite knowledge.

Another response is to argue that although Mary does learn something new, she does not learn any new *propositional* knowledge. Propositional knowledge is normally described as knowledge *that* something is the case: *that* London is in England; *that* electrons are charged. This kind of knowledge is knowledge of facts about the world.

But, says our objector, in the case involving Mary there is no such propositional knowledge and so there are no new facts being learned about the world. Rather, in the same way that I may know *a person* through acquaintance, so Mary is becoming acquainted with redness, and that is a genuinely new experience, even though it is not knowledge of some non-physical fact about the world.

Questions

1 Do you think that the case involving Mary is plausible? (For instance, can we really imagine Mary being brought up in completely monochrome surroundings?) If not, can you revise it to make it plausible?
2 Do you think that Mary has learned something new?
3 Allowing for a moment that Mary is right that she has learned something new: do you think that the knowledge in question is propositional?
4 Do you think that Mary has in fact learned nothing more than a new way to think about or describe facts of which she was already aware?
5 Suppose that Mary has learned something new and that she has learned a new fact: would that persuade you that there are non-physical facts? Why?
6 Do you think that it would be plausible to say that Mary has just acquired a new ability?
7 Do you think that it's right to say that, before her release, Mary knew all of the physical facts about colour?
8 Suppose that Mary learned something 'mental': how should we best understand the nature of something that is distinctively mental?
9 How do you think that we should respond to the thought experiment?
10 Is there anything that would persuade you to change your answer to (9)?

Reading

Conee, E. 1994. 'Phenomenal Knowledge', *Australasian Journal of Philosophy*, 72(2), 136–50.

Jackson, F. 1982. 'Epiphenomenal Qualia', *Philosophical Quarterly*, 32(127), 127–36.

Levine, J. 2007. 'Phenomenal Concepts and the Materialist Constraint', in T. Alter and S. Walter (eds), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, Oxford: Oxford University Press, 145–66.

Lewis, D. 1983. 'Postscript to "Mad Pain and Martian Pain"', in D. Lewis, *Philosophical Papers*, Oxford: Oxford University Press, vol. I.

---

40 Rationality

Very famously, Aristotle described what it is to be human as being a kind of 'rational animal'. The description seems fitting. We seem to be rational creatures. We perform actions for *reasons*. If I am hungry and I want a biscuit, then I will go to the larder to get a biscuit. This is the rational thing to do. It's certainly the logical thing to do. It would certainly not be logical or rational to go to the dishwasher! But there have been studies that have been presented as raising a challenge to the claim that people are rational:

Thought experiment

> Max is asked to take part in a trial. Max is given four cards. On one side of each card is a number. On the other side of each card is a patch of colour. Max (and every other participant in the test) is asked to test, by turning over some cards, the following hypothesis: 'If a card has an even number on one face, then it has a patch of red on the other side.' The cards Max is given currently show: 5, 6, red, green. After thinking a while, Max turns over the cards showing '6' and 'red'. Max is dismayed to learn that her answer is incorrect: turning over the 'red' card is of no use to the test. The hypothesis states that *if* a card has an even number on one face, *then* it has a red patch of colour on the other. As such, she is informed, the correct cards to have turned over are '6' and 'green': only these cards can show the hypothesis to be *false*.

Although we've presented this as a thought experiment, this experiment has been performed and many people chose as Max did. But here is the concern. In this (and some other notable cases) people are asked to test a hypothesis. They fail. If people fail to behave rationally, then they are not being rational (compare with the larder case, above). If people are not always rational (even in contexts where they're being asked to be so), it seems that we should probably not claim that people are rational animals.

There are some notable responses to this line of argument. First, one might simply object that the people engaged in the task were not paying sufficient attention.

In picking the 'red' card, people are just associating the cards with the hypothesis without paying due diligence to the task in hand.

Other responses claim that people don't *engage* with this task in the right kind of way, because it's not central to their interests. In notable further studies, subjects were asked to test the rule 'if you are drinking alcohol then you must be over 25', and given cards with ages on one side and alcoholic or non-alcoholic drinks on the other. In this case, subjects performed much better.

Yet other responses might note that this is just *one* case: a case involving selection. This does nothing to show that we're not *generally* rational. Being rational most of the time might just be enough. But even if we think that there *are* cases where people are rational, it seems quite troubling to the conception of a person as a rational animal that we can fail to make rational decisions in quite this kind of way.

Questions

1 Explain, carefully, why Max's selection isn't correct. Explain what the correct answer is (and why).
2 This task (called the 'wason selection task') shows that even when asked to engage in logical reasoning, people often fail (even at quite simple tasks). Do you think that shows that people aren't rational?
3 What would we have to demonstrate in order to show that people are *not* rational?
4 Explain your answer to (3). Why have you picked *that* as being key to what needs to be demonstrated?
5 Suppose that Max is generally a fan of war films, and also films directed by Stanley Kubrick. What is more likely: that Max likes the war-film *Apocalypse Now*, directed by Kubrick, and works in a library, or that Max works in a library? Explain your answer.
6 The correct answer is *that Max works in a library*: a conjunction is always less probable than either conjunct, provided neither conjunction is certain. And although Max does generally like war films directed by Kubrick, it's not at all certain that she likes *Apocalypse Now*. If someone got that wrong, would you say that she's not rational? Why/why not?
7 If we are able to work out – using logical principles – what the right answer is, and accept that answer if it's pointed out to us, could it ever be right to say that we're not rational? After all, aren't we responding to the correction and able to work out the right answer?
8 What, if anything, do you think is shown by the fact that in the revised test (involving alcohol) people perform much better?
9 Does it matter whether we're rational?
10 Look up the 'base rate fallacy'. How does this impact on your views about human rationality?

Reading

Davies, Paul Sheldon, Fetzer, James H. and Foster, Thomas R. 1995. 'Logical Reasoning and Domain Specificity', *Biology and Philosophy*, 10(1), 1–37.

Sorenson, R. 1991. 'Rationality as an Absolute Concept', *Philosophy*, 66(258), 473–86.

Stein, E. 1997. 'Can We Be Justified in Believing That Humans Are Irrational?', *Philosophy and Phenomenological Research*, 57(3), 545–65.

Stich, S. 1985. 'Could Man Be an Irrational Animal?', *Synthese*, 64(1), 115–35.

Part V

PHILOSOPHY OF SCIENCE

41 Scientific realism

Very roughly, we can think of scientific realism as the view that there exists a mind-independent world; that the language of our best scientific theories literally represents the structure of the world, and that knowledge of our scientific theories *is* knowledge about the mind-independent world. To illustrate: a theory that claims that 'there are electrons and electrons are charged particles', is to be understood as making three claims. That: independently of our existence, there are electrons; talk of there *being particles* and these things *having charge* is literally true; our knowledge of this theory *is* knowledge about the world.

So there is our rough and ready definition of scientific realism. But why believe it? The following thought experiment will go some way to giving us an analogous case that we can use:

Thought experiment

> Suppose that you come across a vending machine. Imagine that you've never seen a vending machine and that you cannot see the inner workings of this one. You notice a piece of paper on the floor nearby. It seems to show something like your vending machine and gives both instructions on use and also a schematic explanation of what's going on inside the machine. So, you decide to test what's shown on the paper. You place a coin in the slot indicated and then press a button. As the writing on the paper suggests, a can of fizzy drink falls out of the bottom. You continue to test the machine, noting that the different buttons reliably produce the different beverages predicted on the piece of paper. Now, either the paper is accurate and has described to you the inner workings of the machine, or it's not. But if it's not, it's a total mystery as to how your button pushings are getting you the right beverage every time (unless you break it open, of course).

The so-called 'argument from miracles' is reminiscent of this. The realist will claim that if our scientific theories are not true (or, at least, not *close* to being true), then many of our technological developments are a complete mystery. Consider, for instance, GPS. It's phenomenally accurate. It's used every day and in a great multitude of ways. Now suppose that our scientific theories weren't true: what *on earth* could be the explanation for this great success?

It seems most straightforward to simply say that the explanation for the success of science is that it's true (or, again, that it's at least close to being true). To say otherwise is to leave unexplained the great success that we have in navigating the world around us.

Questions

1 Write out what you take to be the argument from miracles in premise–conclusion form.
2 Try to write out what you think that scientific realism *is*, using your own words.
3 What is a successful scientific theory?
4 Do you think that the vending machine thought experiment is a good analogue for the case involving realism? What are its limitations?
5 Many false scientific theories have been successful. Find one example.
6 Does the existence of false but useful theories undermine the argument for realism?
7 Would it be coherent to be a realist about some part of science, but not all of it?
8 Suppose that science is not true. How might you set about explaining the success of science?
9 How else might you try to argue for scientific realism?
10 Should scientists be realists?

Reading

Chakravartty, R. 2007. *A Metaphysics for Scientific Realism: Knowing the Unobservable*, Cambridge: Cambridge University Press.

Fine, A. 1986. 'Unnatural Attitudes: Realist and Antirealist Attachments to Science', *Mind*, 95(378), 149–77.

Lipton, P. 1991. *Inference to the Best Explanation*, 2nd edn, London: Routledge.

Wray, K. B. 2007. 'A Selectionist Explanation of the Success and Failures of Science', *Erkenntnis*, 67(1), 81–9.

---

42 Demarcation

On the one hand, we have science: powerful, just and true. On the other hand, we have mere *pseudoscience*: weak, untrustworthy and false. It is clear that there is a difference between the disciplines and enterprises that are genuinely scientific and those that are not. We want to put particle physics on one side of that line; clairvoyance on the other side. Working out where to draw that line is known as the demarcation problem.

So, where does the difference lie? It's natural to view the difference between real science and pseudoscience as something along the following lines: a discipline is pseudo-scientific if and only if it is not in fact scientific; its most serious

proponents try to create the impression that it is scientific. There is a worry, though, that this opening position won't quite do enough to capture the distinction:

Thought experiment

> Bob has his PhD and has a job at a prestigious science lab within a very prominent university. Nonetheless, Bob defends a particular biological theory and performs experiment after experiment, with very poor technique, and does so in the interests of defending a theory that no other scientist believes. He is not behaving in a way that is at all scientific.
>
> Alan has his PhD and has a job at a prestigious science lab within a very prominent university. He is, however, a cheat. These days, Alan never does any work for himself. Instead, he steals all of his results from colleagues at a nearby university. Every night, Alan will hack into their computers, download their files and data, and then work to publish the results first. What Alan is doing is not scientific. Yet folks like Alan will try as hard as they can to create the impression that what they are doing is scientific (though of course also looking to disguise what they are in fact doing).

In Bob's case, we get the right result. Bob is someone that we would naturally describe as a pseudo-scientist. Happily, since what he is doing is not scientific, but it seems as if he is trying to dress it up as such, we correctly classify Bob as a pseudo-scientist. In no part of the investigative process is good science going on.

But our criteria deliver the wrong result when it comes to Alan. Alan is not a pseudo-scientist. There are places in the process of investigation, here, where good science is going on; it's just that Alan steals it! He's simply a liar and a cheat; a scientific fraud.

There might be other problems, though. Although we might quite naturally describe clairvoyance (or astrology) as a pseudoscience, it's not at all clear that its most serious proponents try to create the impression that it is scientific. Indeed, to at least some degree, and with talking of contacting other realms, clairvoyants really could not be charged with trying to create the impression that their project is at all scientific. In that case, it would be quite wrong to insist that a clairvoyant is a pseudo-scientist (though some individual clairvoyants may be).

Another question that lurks here is whether we should be looking to characterize the acts of individual scientists as pseudoscience, or whether we should be looking more widely to particular ways of carrying out research or interacting with the world – often called 'research programmes'. In just the same way that we said that only *some* clairvoyants might count as pseudo-scientists, it might be tempting to treat only some individual scientists as pseudo-scientists. Of course, doing so would mean that we couldn't describe whole tranches of science as being pseudoscience. And *that* might not sit at all easily with some, who might see whole

programmes of science (such as that carried out by creationists) as being pseudo-scientific, quite independently of the practices of individual scientists engaged on the programme.

Questions

1 Write down some instances of pseudoscience.
2 Do all of the instances of pseudoscience that you've picked satisfy the two initial conditions described above?
3 Do you think that it's right to characterize Alan as a fraud and *not* a pseudo-scientist?
4 Is it right, do you think, to describe clairvoyants as pseudo-scientists?
5 Are creationists pseudo-scientists? How can you tell?
6 Do you think that we should use 'pseudoscience' as a description for people, or for larger research programmes? Why?
7 Does pseudoscience have anything to do with the intentions of those involved in its practice? For instance, do they have to be *aspiring* to or *pretending to* do science?
8 It's natural to hear the term 'pseudoscience' as a pejorative. Do you think that we should?
9 Do you think that a scientist should care about the demarcation problem? Why?
10 How would you try to answer the demarcation problem?

Reading

Bunge, M. 2001. 'Diagnosing Pseudoscience', in Mario Bunge, *Philosophy in Crisis: The Need for Reconstruction*, Amherst, NY: Prometheus Books, 161–89.

Hansson, S. 1996. 'Defining Pseudoscience', *Philosophia Naturalis*, 33(1), 169–76.

Lakatos, I. 1981. 'Science and Pseudoscience', in S. Brown *et al*. (eds), *Conceptions of Inquiry: A Reader*, London: Methuen, 114–21.

Pigliucci, Massimo and Boudry, Maarten (eds) 2013. *Philosophy of Pseudoscience: Reconsidering the Demarcation Problem*, Chicago, IL: Chicago University Press.

---

43 Falsificationism

Suppose that you have a theory. You want to know whether or not your theory is any good. You test it. It would seem very natural to focus on looking for *confirmation* of your theory. Positive reinforcement of the view that your theory is the right one is an obvious thing to look for. If I have a theory that every time I press the 'go' button on my coffee machine, I'll get a cup of coffee, then it's natural

to see every successful button-push-coffee-get combination as a vindication of the theory.

But one might worry about this. Consider the difference between a trend and a law. A trend, we can say, is a pattern of events that just so happens to produce particular results. A law, in contrast, is something a bit stronger. Using this contrast, consider the following thought experiment:

Thought experiment

> Carla is sitting in her office. Mark walks into her office and the phone rings. Surprised, they share a laugh about the coincidence. Mark leaves her to the call and then pops back a little later. Lo and behold, Carla's phone rings again. Again, Mark leaves her to the call and returns later, and again the phone rings upon his return. Carla thinks: huh, it seems to me like every time Mark walks into the room, the phone rings. When it happens for the fourth time, she takes this to be confirmation that her theory is correct.

We should be clear: Carla isn't reasoning well. If she works in a busy office, it's not all that surprising that her phone rings quite a bit. Nor is it then all that surprising that her phone should ring when Mark walks in. Worse still, Carla isn't looking to try to test the theory. She's not looking to see if she can find a way to demonstrate that it's false. In this vein, she could try a number of things. She could ask Mark to come through, regularly, at 5-minute intervals. She could ask people not to ring unless it's urgent, change venues, and so on.

Now although the case is somewhat trivial, the important thing to note is the attitudinal difference between Carla looking for easy confirmation, and her looking to try to test the theory and show it to be false.

According to Karl Popper, what we look for in our scientific theories is (or should be) whether we can demonstrate that they are false. This, in fact, then gives us two useful tools with which to approach the scientific enterprise. First, a theory isn't properly scientific unless it is falsifiable – that is, unless we can work out how (in principle) we might demonstrate that it's false. Second, when engaging with scientific theories, we should not hold them to be fixed and true: rather, we should be trying to find ways to falsify them.

Questions

1. Write down a scientific theory.
2. Write out how, in principle, you might try to falsify that theory. What observations would you have to make to show it to be false?
3. Write down a theory that isn't falsifiable.
4. Your unfalsifiable theory will not count as scientific, for Popper, but do you think that theory is good, apart from that? Are you tempted to believe it?
5. Can you think of a scientific theory that isn't (or wasn't) in principle falsifiable? If so, is that a threat to the project of falsificationism?

6 Take falsificationism to be the view that a theory is scientific if and only if it is falsifiable. Falsificationism is a theory. Is it thereby scientific? Does it matter?
7 Do you think that practising scientists use falsificationism as a principle in their work?
8 Do you think that it matters whether practising scientists use a principle of falsification?
9 Could a theory that isn't scientific (in Popper's sense) ever be a good theory, do you think?
10 If you were to replace a falsificationist principle, what would you replace it with?

Reading

Grünbaum, A. 1976. 'Is the Method of Bold Conjectures and Attempted Refutations Justifiably the Method of Science?', *British Journal for the Philosophy of Science*, 27(2), 105–36.

Lakatos, I. 1970. 'Falsification and the Methodology of Scientific Research Programmes', in I. Lakatos and A. Musgrave (eds), *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press,

Mellor, D. H. 1977. 'The Popper Phenomenon', *Philosophy*, 52(200), 195–202.

Popper, K. 1959. *The Logic of Scientific Discovery*, London: Hutchinson.

---

44 Observation in science

It is reasonably clear that science relies upon observation. In order to know that an apple fell from a tree, Newton would have to have seen it on the tree, and then see it fall. These are both observations.

There is an obvious and intuitive difference between observing and experimenting. I might simply observe the apple fall. There, I do not directly interact with the apple. But in a case of a laboratory controlled experiment, I directly influence the set up and make it conform to very specific parameters.

But if observations are central to science, then there are good questions as to what gets to count as an observation. It's clear in the case of the apple, that *I* experience some phenomenon, such that my experience of it is of the apple falling. Yet when we talk of observations in the context of scientific theorizing, we don't mean to be interested in the internal mental life of a lone subject.

Instead, what we seem to be interested in are facts that can be accessed intersubjectively: it isn't merely *my* experience that's crucial, or yours. Rather, what we're interested in are the kinds of facts that we can all agree upon. So, during an experiment, what we're likely to be interested in is whether or not a particular needle is pointing to a given number on a dial, or if a particular light is on, or if a particular digital display is giving us the result that we expect.

But even supposing that's right, it remains a good question as to how much we should trust these observations, as revealed by the following:

Thought experiment

> Cecile is asked to take part in an experiment. She is shown a series of playing cards, one after the other, for a very short period of time. Cecile is able to correctly identify the cards that she is shown from the first pack of cards that she is tested against, so we establish that duration of viewing is sufficient for her to identify the cards. In the second part of the test, Cecile is shown another set of cards. However, this time there is variation. Most of the cards look just as they do in the first pack, but in this second pack, some of the heart-cards are black. When presented with these cases, Cecile reports having seen normal heart-cards – that is, red ones. Upon finding this, what the experimenters decide is that Cecile's observations are *theory laden*; the background expectations that she brings to the task cause her in fact to *see* the cards as red (and she really does believe that she's had the experience of red playing cards).

The experiment described is not, in fact, a mere thought experiment. It's an experiment that's been carried out – as have others like it. The result does indeed seem to be that particular beliefs and theories that observers hold can lead them to 'see' the world in particular ways.

But this is troubling. If we're interested in science, then we're interested in objective and intersubjective observation. Cecile sees a red card. Someone who has never seen playing cards before, and is not familiar with their colour conventions, will see a black card. In that case, we seem obliged to say that reports of colour are not available purely intersubjectively. And, in that case, colour reports should not be treated as playing any role in our science. As you might imagine, though, these experiments threaten to infect more than just our experience of colours; many of our observations seem to be infected by this feature of being 'theory laden' assumptions. Perhaps, then, we cannot trust our observations in science *at all*.

Questions

1. If things really do appear differently to people with different theoretical resources, such that their observations really are 'theory laden', do you think that we can engage properly in scientific practice?
2. How, exactly, would you define 'observation' within the context of science?
3. How widespread do you think that the theory ladenness of observation is? Do we have any reasons to think that it might be restricted to only a few cases?
4. Suppose that observational theory ladenness is very widespread: would it make sense to say that observation is so theory laden that people with different conceptual resources really do systematically *see* different things? For instance, would (e.g.) Aristotle and Stephen Hawking see different things when observing some event (like the falling of a stone)? Would this be problematic for scientific practice?

5 Can you find or think of cases from the history of science where theory ladenness has been easy to nullify?
6 Can you give a clear and precise description of the difference between observation and experimentation?
7 Doing some research first, write down what an electron microscope is.
8 When using electron microscopes, do you think that scientists are making observations? Is your answer of any concern to the idea that scientists do (and should) rely upon observations when forming their theories?
9 Is a scientist using a Geiger counter observing radiation? Is there a sense, do you think, in which she is observing the radiation, but only indirectly? If so, can you write down exactly what it is to observe something indirectly? Does it make a difference that someone has to teach her how to use the Geiger counter?
10 Given that science is very successful, does it matter whether observations are theory laden?

Reading

Kuhn, T. S. 1962. *The Structure of Scientific Revolutions*, Chicago, IL: Chicago University Press.

Lewis, C. I. 1950. *Analysis of Knowledge and Valuation*, La Salle, IL: Open Court.

Pinch, T. 1985. 'Towards an Analysis of Scientific Observation: The Externality and Evidential Significance of Observation Reports in Physics', *Social Studies of Science*, 15, 3–36.

Van Fraassen, B. 1980. *The Scientific Image*, Oxford: Clarendon Press.

---

45 Science

One or many?

The picture that we have been working with throughout this section is that there is a single, unified practice that we might call 'science'. This, we assume, will have a particular methodology: it will begin with hypothesis and prediction formation, involving testing that prediction, and then modifying the hypothesis as necessary. But nothing that has been said here would justify the assumption that there really is a single, choate discipline, set of theories and methodology that we might reasonably call 'scientific'.

We will consider discipline, theory and method, here, but let us begin with a thought experiment designed to probe certain of our thoughts about scientific method:

Thought experiment

> Edwin and Melissa are arguing about religion. They are trying to decide which religion is true. Although they both have some feelings that they would describe as 'spiritual', neither of them was raised within a particular religion, and neither of them has found the religions that they have encountered to reflect their own feelings about spirituality and religion. After several hours of discussion, they agree to adopt a religious pluralism: there are many different religions, each of which has equal claim to be the one true religion, with no obvious reason (that they can see) for preferring one over the other. This is not to say that they think these religions false. Far from it. Both Edwin and Melissa are extremely tolerant of other religions and are prepared to allow that each of these religions has an equal claim to be *true*. Happy that they have agreed on this, Melissa and Edwin then turn to talk about science. Melissa states that she finds it obvious that science is unlike this. Competing scientific theories do not all have equal claim on being *true*. One theory is true; all of the others are false. Edwin, however, finds himself drawn to the same conclusion that they reached about religion: there can be many different scientific theories that do not agree with one another, but have an equal claim to be *true*.

The kind of view sketched in the thought experiment is a form of scientific pluralism. Why might one be attracted to such a view? Well, consider the claim that there is a single, unified scientific method. If that was right, then we might expect the results of scientific inquiry to be genuinely in competition with one another.

However, what if the methodologies involved were very diverse? Would we then be strictly making claims that compete? To give an analogy (which may not be fair – we leave it for you to judge): someone who uses personal historical sources to determine the extent of Heidegger's Nazism (if there is any – it's a controversial subject) may not reach the same conclusion as someone who solely examines his published work. The historian may uncover (from examination of diaries and letters) that Heidegger *the person* was a Nazi. But someone who examines Heidegger's work and finds him not to be a Nazi in his published work does not thereby contradict that. She finds, instead, that his philosophical work is not intrinsically bound up with Nazism.

Now look at the diversity of methods used in different periods in science. Is it really fair to say that we're engaged in the same kind of inquiry as Newton? And, if not, then to what extent is it fair to say that we can now show his theories to be false?

Questions

1 Do you think that pluralism about religion is more or less plausible than pluralism about science? Explain your answer.

2 In the case involving Heidegger, do you think that it's plausible to adopt a form of pluralism whereby there is no contradiction between the two views? Explain your answer.
3 Is there a single, unified methodology employed by all current practising scientists? If so, what is it? If not, provide counterexamples to the claim that there is a single unified approach.
4 Is there a single unified methodology employed by all past, present and future scientists? If so, what is it? If not, provide counterexamples to the claim that there is a single unified approach.
5 Would a difference in the methodology of the groups described favour scientific pluralism? Explain your answer.
6 Do your answers to (2), (3) and (4) suggest that scientific pluralism is true? Explain your answer.
7 Another way to understand the claim that we should be pluralistic about science is to say that we should reject the idea that the only values are 'scientific' values (we should similarly value inputs from other disciplines and modes of thought). Do you think that we should endorse such a pluralism? Why?
8 Suppose we reject pluralism. In that case, when a new theory is being explored, it seems as if the deck may be stacked against it. After all, it must (typically) be shown to be compatible with our current body of theories (otherwise it won't be consistent with our scientific approach). Do you think that sets too high a bar for new theories, illicitly favouring old ones?
9 Do you think that any practising scientists explicitly endorse pluralism? Why? Does it matter?
10 We've mentioned scientific pluralism and religious pluralism: can you think of other areas where a form of pluralism might be attractive?

Reading

Feyerabend, P. 1975. *Against Method*, London: Verso.
Leplin, J. 1969. 'Meaning Variance and the Comparability of Theories', *British Journal for the Philosophy of Science*, 20(1), 69–75.
Moberg, D. W. 1979. 'Are There Rival, Incommensurable Theories?', *Philosophy of Science*, 46(2), 244–62.
Van Bouwel, J. 2009. 'The Problem With(out) Consensus: The Scientific Consensus, Deliberative Democracy and Agonistic Pluralism', in J. van Bouwel (ed.), *The Social Sciences and Democracy*, London: Palgrave Macmillan.

---

46 Science versus philosophy

There are some occasions where philosophical and scientific theories conflict. The nature and severity of that conflict may vary from case to case. But, in these

cases, we must make a decision. Do we persist with our philosophical theory and give up on the scientific theory, or persist with our scientific theory and give up on the philosophical theory?

The following thought experiment will give us some traction on one way in which this kind of issue might play out:

Thought experiment

> Two friends, Simon and Rebecca, are talking. They are discussing structural questions – questions about the fundamental structure of reality. Simon is a student of physics. He believes that the Planck length ($1.61619926 \times 10^{-35}$ metres) is the smallest length that can be measured and he believes this to be a consequence of our best physics. Because there is no length physically measurable that is shorter, Simon doesn't think that we can make sense of talk of distances shorter than the Planck length. Rebecca is a student of philosophy. She agrees with Simon's interpretation of our best physics, and agrees that the physics seems to imply that the smallest measurable length is ($1.61619926 \times 10^{-35}$ metres). However, Rebecca (unlike Simon) thinks that it would be perfectly sensible to talk about distances smaller than a Planck length; she thinks that even the Planck length is divisible into parts.

There is, then, a seeming tension in this case. We must decide who is right. Simon thinks that he's right. He thinks that if we look to the structure of the world as it's described by our best physics, then we'll find the right answers. And because we can't physically measure anything smaller than a Planck length, there is nothing more to be said about the structure of the world.

Rebecca thinks that she's right. She doesn't see why physical measurement needs to be taken as the last word on structure. Rather, although it will never in fact happen that we can measure such a thing, she thinks that we can conceive of smaller distances. And as such, she thinks such distances are possible. She sees no reason why such distances might not also be actual.

What we must now do is try to work out who is right. In cases like this, where we have a tension between physical and philosophical theories, we must decide who is right, and why. But notice: one thing that we must not do is simply follow prejudice. For instance, it would not do for the scientist to simply assert that the more scientific answer is correct, and it would not do for the philosopher to simply assert that their view is correct. Whatever route we pursue, we need to ensure that our position is rational and well grounded.

Questions

1. In the above, who are you inclined to side with, and why?
2. In the thought experiment, do you think the tension between Simon's view and Rebecca's view is a product of the science itself, or of the interpretations of it offered by Simon and Rebecca? Why might this matter?

3 If there is no interpretation involved, are you more or less likely to side with the philosopher?
4 Some people are disposed to find against the philosophers in these kinds of cases: write down as many reasons as you can for why that might be.
5 Now write down as many replies to these as you can.
6 Is there any in-principle reason that a philosophical theory might not overturn a scientific theory?
7 Can you write a description of the difference between science and philosophy?
8 Do you think that our best philosophy should be informed by our best science? Why?
9 Do you think that our best science should be informed by our best philosophy? Why?
10 Does it matter if our science and philosophy conflict? Can't they both be true?

Reading

Hawley, K. 2006. 'Science as a Guide to Metaphysics?', *Synthese*, 149(3), 451–70.
Hofweber, T. 2009. 'Ambitious, Yet Modest, Metaphysics' in D. Chalmers, D. Manley and R. Wasserman (eds), *Metametaphysics*, Oxford: Oxford University Press, 260–89.
Ladyman, J. and Ross, D. 2007. *Every Thing Must Go*, Oxford: Oxford University Press, esp. chapter 1.
Lowe, E.J. 1998. *The Possibility of Metaphysics*, Oxford: Blackwell, esp. chapter 1.

---

47 Anti-realism

Science is immensely powerful. It generates fantastically useful results for us all the time. And, as we have seen already, that inclines some to think that science gives us truths. But not all are so inclined. Anti-realist positions are those that take science to be useful, but not true.

One of the more famous arguments in favour of anti-realism is an argument known as the 'pessimistic meta-induction'. As the name suggests, the argument uses an induction *about* our inductive practices, and leads us to the (pessimistic) conclusion that we should not suppose that our scientific theories are true. To illustrate, here is a thought experiment:

Thought experiment

> Riham is very keen on induction. She thinks, for instance, that if the Sun has risen on the last 13,000 mornings (that make up her life to date), that gives her a pretty good reason to think that the Sun will rise tomorrow. Riham then comes across some startling data. An awful lot of scientific theories have been superseded. Indeed, most of the theories that we currently hold will have replaced others, and those, in turn, will have replaced still other theories. Yet all of those theories – those superseded theories – *were*

> superseded. Riham then reasons as follows: if all of those other scientific theories (that we don't now hold) have been superseded, then that gives us an excellent inductive base from which to reason that all of our *current* scientific theories will be superseded, too. And, in turn, the theories that replace our current theories will be superseded, and so on and so on. Thus, she reasons, we shouldn't take *any* of our scientific theories (present or future) to be *true*. After all, a true theory will not be superseded.

Here Riham is engaging with the concerns at the heart of the pessimistic meta-induction. It can sometimes be tempting to see the pessimistic meta-induction as obviously flawed. After all, we're making progress aren't we? But the difficulty with this very intuitive response is that although we're capable of more and ever more *technical* achievement, the concerns Riham has are consistent with that. Riham will agree with us that we're now capable of many things that we weren't.

What she will also then want to say, though, is that if we use simple inductive reasoning, we see very straightforwardly that this improvement should not be taken to signal that we're now at the truth. It should, instead, be taken to signal that we have more work to do. Whilst one extreme option would then be to deny the power of induction, that too would seem to result in our having to give up on a good deal of scientific practice.

Questions

1. Write down a list of scientific theories that have been superseded, as well as the theories that superseded them.
2. Suppose that every time one of your friends told you that something was true, she later changed her mind. When she next told you something, would you think it was true?
3. Of what relevance is the last question to the pessimistic meta-induction?
4. What do you think that a professional scientist would say in response to Riham?
5. Suppose that science is false, but useful. Would that matter to you?
6. Suppose one of our scientific theories were true. The pessimistic meta-induction would have us believe it false. Is this possibility a problem?
7. Suppose that our scientific theories are all false, as suggested. What would be the most obvious source of error? Why are we getting it wrong? Is our language too imprecise? Are our instruments to blame?
8. Would the pessimistic meta-induction work in other domains, besides science? For instance, would the argument work in philosophy? In History?
9. How would you try to persuade Riham that she's wrong, if you had to?
10. Do you think that Riham is wrong?

Reading

Laudan, L. 1981. 'A Confutation of Convergent Realism', *Philosophy of Science*, 48(1), 19–48.

Magnus, P. D. and Callender, C. 2004. 'Realist Ennui and the Base Rate Fallacy', *Philosophy of Science*, 71(3), 320–38.

Nola, R. 2008. 'The Optimistic Meta-Induction and Ontological Continuity: The Case of the Electron', in L. Soler, H. Sankey and P. Hoyningen-Huene (eds), *Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities?*, Dordrecht: Springer.

Papineau, D. 2010. 'Realism, Ramsey Sentences and the Pessimistic Meta-Induction', *Studies in History and Philosophy of Science*, 41(4), 375–85.

---

48 Theories and evidence

When we're looking to test a scientific theory, we will look for evidence. We will look for evidence that either confirms or disconfirms our theory. Disconfirmation is easy. If your theory tells you that something should happen, and it doesn't, then so much the worse for your theory.

But what happens if your theory predicts that something will happen, and it does. Does that *confirm* your theory? We saw, when we were looking at falsificationism, that there are potential problems lurking here. But in this section we're going to look at another: can evidence *ever* really confirm a theory? Let's take a look:

Thought experiment

> Consider any two points on a graph. We can draw a straight line between them. We can also draw infinitely many *other* lines between them. These other lines will have a variety of kinks, twists and turns. The basic point is simply that for any two points, there are infinitely many ways to connect them.
>
> Now add points onto the graph. Suppose, indeed, that we add five hundred more points in such a way that we can draw a nice straight line *right through the middle* of all of them. We can draw infinitely many *other* lines through those points. The kinks, twists and turns will be tighter, but nonetheless it could be done. We can do the same if we add another five hundred points. Another thousand. It doesn't matter how many points we add, there remain an infinite number of ways of drawing lines that will go through all of the points. Now suppose that each point represents some observation that we've made. Each line running through them all represents a mathematical description of the state. There are infinitely many. No matter how many more points we have, we still have infinitely many mathematical descriptions of the whole state.

At this point, the thought experiment may seem fairly trivial. But don't forget that physical laws are mathematical descriptions of a given state. And what we've just said is that there are infinitely many ways in which we could describe that state.

Some people then find it tempting to think that this shows that the evidence that we have systematically underdetermines our choice of theory. That is, the evidence on its own doesn't help us decide which of the infinitely many theories we should

fix upon. After all, if the evidence is compatible with infinitely many states, then the evidence by itself isn't going to tell us which theory is correct. It is tempting, then, to see this as an argument for anti-realism: the evidence doesn't tell us which theory is correct, so why think that there's a correct theory?

But other people are inclined to see things differently. Presented with the view that the evidence systematically underdetermines our choice of theory, we could add to the tools that we use in allowing us to choose between options. For instance, we could say that we will opt for the *simplest* theory, or the most elegant theory. What is certain, though, is that the evidence by itself certainly won't confirm for us which theory is true. The best that it can do is rule out some false theories.

Questions

1. Draw a graph with ten points. Now draw ten different lines that each connect all of the points. Can you see how you would add more lines to the graph while still connecting all of the points?
2. Does the underdetermination of theory by evidence tempt you towards anti-realism? Why?
3. Are you convinced that the evidence really does underdetermine theory choice?
4. If we performed enough experiments, would the theory still be underdetermined?
5. Do you think that elegance in a theory is a guide to its truth? Why?
6. Do you think that simplicity in a theory is a guide to its truth? Why?
7. What other virtues might you look for in a theory? Why did you pick those particular ones?
8. Could we test whether or not the virtues listed are likely to lead us to truth? Does that matter?
9. Suppose that we have a theory that everyone agrees upon and that seems to get the right results. Does it matter whether it's underdetermined? Why?
10. Underdetermination looks a lot like a theory (that no theory is ever totally confirmed by evidence): is it, itself, underdetermined? Does it matter?

Reading

Bonk, T. 2008. *Underdetermination: An Essay on Evidence and the Limits of Natural Knowledge*, Dordrecht: Springer.

Earman, J. 1993. 'Underdetermination, Realism, and Reason', *Midwest Studies in Philosophy*, 18(1), 19–38.

Laudan, L. 1990. 'Demystifying Underdetermination', in C. Wade Savage (ed.), *Scientific Theories* (Minnesota Studies in the Philosophy of Science, vol. 14), Minneapolis, MN: University of Minnesota Press, 267–97.

Norton, J. 2008. 'Must Evidence Underdetermine Theory?', in M. Carrier, D. Howard and J. Kourany (eds), *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited*, Pittsburgh, PA: University of Pittsburgh Press, 17–44.

49 What are the laws of nature?

At the heart of just about any scientific study is a law. They're everywhere. Hubble's law of cosmic expansion; Kepler's laws of planetary motion; Newton's laws of gravitation and motion; the laws of thermodynamics.

But one question that we might be interested to consider is exactly what these laws *are*. We know many of their names, but in virtue of *what* shared feature is it right to call them *laws*? In order to get a sense of what's at stake, consider the following:

Thought experiment

> Connie notices that whenever x occurs, y follows afterwards. She finds this intriguing. She asks her friends to check, too. They too confirm that y follows x. Connie successfully applies for funding from a large scientific body and tests her claim that occurrences of y follow occurrences of x. She notes that, indeed, no matter how she tests for it, she cannot get away from cases in which occurrences of ys follow occurrences of xs. Connie is delighted. Talking with the team from her lab, she announces confidently that ys have followed the xs. Her lab-mates look bemused. They press: 'don't you mean that you've discovered that ys follow xs? Like, it's a law that ys follow xs?' 'No, says Connie, I just mean that ys have followed xs.'

It might well seem to us that Connie has gone wrong in her reasoning. It isn't merely that she's seen *some* ys follow *some* xs. She seems to have a good evidential base for the conclusion that there's a law that ys follow xs, such that when she performs her test again tomorrow, a y will follow an x.

Perhaps we could go for something stronger, then, that *all* ys are followed by xs. That's stronger. And so perhaps what a law *is* is a generalization that applies to all cases. It's something of the form: whenever x happens, y follows.

But are we prepared to stop even there? For instance, don't we want to say something a little bit stronger still? If there's a law involved, then do we not want to say that if x occurs, then y *must* follow? We don't think, after all, that it's optional for objects to partake of the law of gravity. If two objects are massive, *then* they've just *got* to be attracted to one another, haven't they?

In that case, it seems we must say that a law is some kind of necessitation that connects distinct entities. Thus, if it's a *law* that x follows y, then (we might think) the existence or occurrence of x *necessitates* the existence or occurrence of y.

Questions

1. Find out about each of the laws mentioned at the outset.
2. For those laws, do you think that the laws mean to say that if some state is realized, then some other state *must* follow? Use your examples to explain why.

3 Suppose that we do think that y *must* follow x. Would this require that there's something in the world (like a property or relation) that connects x to y, ensuring that the one follows the other?
4 If we did require that necessary connection, would that show that laws of nature are implausible?
5 Could it be a law of nature that sugar dissolves in water if sugar was never put in water? Explain your answer.
6 Suppose we say that all sugar dissolves in water. Could this itself be is a law? Do we need to say anything else?
7 How might we set about finding out what laws of nature are?
8 Suppose that there existed a universe with a lone particle travelling through it at a constant velocity. Would it be a law at that universe that the particle must travel as it does?
9 Do you think that scientists should care about what kinds of thing laws are?
10 Do you think that there might be a difference in kind between a physical law and (e.g.) a biological law? Explain your answer.

Reading

Armstrong, D. 1983. *What Is a Law of Nature?*, Cambridge: Cambridge University Press.
Lewis, D. 1983. 'New Work for a Theory of Universals', *Australasian Journal of Philosophy*, 61(4), 343–77.
Schaffer, J. 2008. 'Causation and Laws of Nature: Reductionism', in J. Hawthorne, T. Sider, and D. W. Zimmerman, (eds), *Contemporary Debates in Metaphysics*, Oxford: Blackwell, 82–107.
Tooley, M. 1977. 'The Nature of Laws', *Canadian Journal of Philosophy*, 7(4), 667–98.

---

50 Are there any laws of nature?

We have proceeded, so far, on the assumption that there are laws of nature, and that scientific practice reveals these to us (though see Chapter 41, 'Scientific realism'). But why think that's true? Why think that there are laws of nature? (And as to what these laws *are*, see Chapter 49, 'What are the laws of nature?')

The normal starting place would be observations. We see patterns and regularities in the world around us. Because we see these patterns and regularities, we think that there must be some explanation for this, and so we posit a law. But do we really see patterns and regularities? Do we see very exceptionless regularities of the sort that we might expect to see in a proper scientific law?

Thought experiment

> Bertrand is in his laboratory. He's testing the gravitational law governing the attraction between massive bodies: $F = Gmm'/r2$. He notices that every time he performs an experiment, he gets a result that approximates the relationship described by the equation, but never quite that relationship. It's always just a little bit out. Bertrand tries varying the experiment, but to no avail. Although he's able to make the experiment yield results that consistently track around the relationship $F = Gmm'/r2$, he's never able to, with complete accuracy, get that result. Nonetheless, thinks Bertrand, that's OK. After all, what's going on here is simply that other massive bodies that are in the vicinity are affecting the result. That's fine. He's still shown that the law is OK.

We might be tempted to take issue with Bertrand's line of reasoning. After all, he's not actually seen this law – this exceptionless regularity. What he's seen is something that approximates it, but that is something else. What he's seen, we might say, is that *in these specific circumstances, with these specific objects*, we get some results.

But that's uninteresting. That's just to state how a particular instance turned out! What we were interested in was the law. Note, though, that above we started off by saying that we should believe that there are laws because there are patterns and regularities. It now seems that this idea is under some pressure. *Sure*, there are rough and approximate patterns that folk like Bertrand are seeing. But in our scientific theories we're supposed to be after something more – something more precise. Something, like a law that is without exception.

What we're now seeing, and what we might well see in the examination of any putative laws, is that we don't have any such thing. All that we have available to us are a series of similar, though ultimately pretty different, goings on.

Questions

1. Can you think of any case in which someone might observe exceptionless regularity? Is it physically possible?
2. Do you think that laws (if there are any laws) have to be exceptionless? Could they be rough statistical patterns, for instance?
3. Might we expect there to be a difference between biological laws (for instance) and fundamental physical laws, with respect to whether or not they're exceptionless?
4. Would a law that admits of exception lose its explanatory power?
5. Would it matter if laws of nature were not explanatory?
6. What do you think of as the essential features of a law of nature? What must they do?
7. Write down a law of physics. Would it be contradicted if it failed in a single case?
8. Does that tell us anything about whether or not laws of nature are supposed to be exceptionless?

9 Would this question interest a scientist?
10 What view do you think a practising scientist would take on the question of whether or not she's seen exceptionless regularities? What position would she take on laws: are they exceptionless?

Reading

Cartwright, N. 1980. 'Do the Laws of Physics State the Facts?', *Pacific Philosophical Quarterly*, 61(1–2), 75–84.

Earman, J. and Roberts, J. 1999. '*Ceteris Paribus*, There Is No Problem of Provisos', *Synthese*, 118(3), 439–78.

Giere, R. 1999. *Science Without Laws*, Chicago, IL: Chicago University Press.

Lange, M. 1993. 'Natural Laws and the Problem of Provisos', *Erkenntnis*, 38(2), 233–48.

Part VI

APPLIED ETHICS

51 Judith Jarvis Thomson on abortion

Perhaps no other topic engenders such violent and impassioned responses as abortion. When discussing the morality of abortion, one of the central questions is whether the fetus is a *person*. A basic argument against abortion might run as follows: there is good reason to believe that the fetus is a person; abortion involves killing a fetus; abortion involves killing a person; killing a person is morally wrong; therefore, abortion is morally wrong.

However, Thomson (1971) argues that this is a mistake and that the preoccupation with the personhood of the fetus is not the way to move forward in this debate. Her strategy is to point out that there are some instances where we think that *killing a person might be morally acceptable*. If this is true, and if there are similarities between these instances and cases of abortion, then we can say that even if killing a fetus is killing a person, this doesn't mean that abortion is wrong.

Imagine a case of self-defence where, for example, the only way to save your family is to kill an assailant. We might think that in this case it is morally acceptable to kill the assailant, who is a person. Thus, if there are similarities between this sort of case and cases of abortion, then it might be morally acceptable to abort the fetus, even if the fetus is a person. Thomson uses a very striking thought experiment to make this point about killing people:

Thought experiment

> You are kidnapped in the night and the Society of Music Lovers has connected your kidneys to a famous but unconscious violinist. The violinist needs your kidneys as hers don't work. Her body will take nine months to recover, in which time she needs to remain connected to your kidneys at all times. Given this, your way of life will be restricted for nine months.

Will you have done something morally wrong if you disconnect yourself from the violinist, even though this will lead to her death? Thomson thinks that the answer is 'no'. If you 'walk away' you won't have done anything morally wrong. Importantly the question here isn't how would people actually behave, but rather, if you *do* detach yourself – despite whether you find it hard or easy to do so – have you done something wrong? Thomson thinks not.

Her point then is this. Disconnecting will lead to the death of a person, disconnecting is morally acceptable. So, sometimes it is morally acceptable to bring about the death of a *person*. The analogy with abortion, hopefully, is clear: the violinist who is attached to you for nine months represents the fetus; the tubes connecting

you to the kidneys of the violinist are the umbilical cord. Thomson argues that just as it is acceptable for you to 'walk away' from the violinist, so it is acceptable for a woman to 'walk away' from the fetus, even if the fetus is a person and will surely die.

To be clear, this thought experiment seems only relevant when the pregnancy is brought about without the mother's consent – in cases of pregnancy due to rape: after all, in the thought experiment you are 'kidnapped'. So, for example, we might think that if we had signed a contract agreeing to be connected to the kidneys and had met up with the violinist beforehand then it would *not* be morally acceptable to disconnect. Thomson is aware of this and brings in a number of other thought experiments to show that she thinks that even in less 'extreme' cases where, for example a woman is pregnant due to consensual sex and is in a stable economic and social situation, abortion can still be morally permissible.

Further, the thought experiment is focused on *you* detaching yourself, not someone else detaching you. So, the analogy with abortion seems wrong, as in most abortions a doctor will perform the abortion, not the woman. The correct question then seems to be, not is it OK for *you* to detach yourself, but is it morally acceptable for someone else to detach you from the violinist. Again, Thomson is aware of this and introduces further arguments to support the claim that it would indeed be morally permissible for someone else to detach you from the violinist; that is, for a doctor to perform the abortion.

Questions

1 Define 'abortion'.
2 'Abortion' is about terminating the pregnancy, not *killing* the fetus. Does that distinction make any significant difference to your thinking on the issue?
3 Consider the thought experiment. Do you think it would be morally acceptable to 'walk away'?
4 List the instances where you think it *might be* morally acceptable to kill a person. Are there any things which could be learned from these cases which we can apply to cases of abortion?
5 Why does the debate about abortion so frequently surround the personhood of the fetus?
6 Is the fact that the violinist is a stranger, and unconscious, important to Thomson's argument?
7 Does it make a difference if it is *you*, or someone else, who detaches you from the violinist? How might this relate to the abortion debate?
8 If you had *consented* to be connected to the violinist would it then be morally acceptable to 'walk away'?
9 Consider your answer to (8). What would count as 'consent'?
10 Should the definition of abortion limit abortion to *human* pregnancy?

Reading

English, J. 1975. 'Abortion and the Concept of a Person', *Canadian Journal of Philosophy*, 5(2), 233–43.
Marquis, D. 1989. 'Why Abortion Is Immoral', *The Journal of Philosophy*, 86(4), 183–202.
Thomson, J. J. 1971. 'A Defense of Abortion', *Philosophy & Public Affairs*, 1(1), 47–66.
Tooley, M. 1972. 'Abortion and Infanticide', *Philosophy & Public Affairs*, 2(1), 37–65.

---

52 Euthanasia

Rarely a month goes by that does not include a news story about euthanasia. Euthanasia can be thought of as ending the life of a person when it is thought to be in the person's best interest. Typically, this might involve turning off life support, or administering morphine to a terminally ill person leading to his or her death.

'Euthanasia' is in fact a catchall term for a host of different things. In particular, we will be concerned with *active* euthanasia, where the doctor actively brings about the death of the person – perhaps by injecting the person with a knowingly lethal dose of morphine; and *passive* euthanasia where the doctor decides not to intervene, or removes treatment – e.g. where a doctor stops a course of antibiotics for a terminally ill person – knowing that she will develop pneumonia and die.

The difference between active and passive euthanasia is particularly important because the medical profession often rules out active euthanasia but not passive euthanasia. This is based in part on a supposed moral distinction between *killing* (active euthanasia – wrong) and *letting die* (passive euthanasia – sometimes acceptable). In this chapter we will consider a thought experiment put forward by James Rachels (1997), that puts pressure on this distinction. If he is right then it is not true that killing is always morally worse than letting die:

Thought experiment

> Smith and Jones intend to kill their young cousin in order to inherit some money. They hear their cousin in the bath and each decides to drown him. In the first scenario, Smith enters the bathroom and drowns the boy. In the second, Jones, on entering the bathroom sees the boy slip and fall face down in the bath. Jones could extend his arm and help the boy out of the water, but he doesn't. He is prepared to push the boy under again if he regains consciousness, but he doesn't and the boy dies.

The question Rachels asks is this: is Smith's action morally worse than Jones's? Rachels thinks that our answer is going to be 'no'. He thinks that the actions of both men are equally as wrong. Do you agree?

On first encountering this thought experiment people often think that it clearly has nothing to do with euthanasia. After all, Smith and Jones want to kill their cousin and are nasty pieces of work and are motivated by corrupt reasons. Doctors, however, are motivated to help their patients, are not evil, and are moved by noble and praiseworthy reasons. However, this fundamentally misses the point. The way to see this is to ask what Rachels' thought experiment actually shows. Once we see this, then arguably it *is* directly relevant to the euthanasia debate. You should keep at the forefront of your mind that we are only asked to *compare* Smith with Jones, we *aren't* asked about whether Smith and Jones are good people, bad people, corrupt, etc.

Rachels thinks that there is only *one* difference between the Smith and Jones cases, namely, Smith *kills* his cousin whereas Jones *lets his cousin die*. To see why there is only one difference, think about the components of the event. We have (a) the intention of the men, (b) the act itself, and (c) the consequences. In the thought experiment the *intentions* of both men are the same – *to kill* their cousin; and the *consequences* are the same – their *cousin dies*. Thus, it is irrelevant that Smith and Jones are nasty and corrupt whereas doctors are not. We are asked to compare Smith and Jones and both have the *same* intentions. Rather the only difference between the Smith and Jones cases is the *act*. Smith *holds* the boy under, Jones *watches* the boy drown without intervening.

What this shows is that if we think that what Smith and Jones did was equally morally wrong, then it cannot be the case that there is a moral distinction between killing and letting die. But if this is so, and if the difference between active and passive euthanasia is just the difference between killing and letting die, then there isn't a moral distinction between active and passive euthanasia.

Finally, notice that it is consistent with Rachels' argument that there *could be* a moral distinction between active and passive euthanasia, but only if that difference isn't based on the difference between killing and letting die. For example, it might be the case that active euthanasia makes more people scared and unhappy than passive euthanasia and we might then think that this means there is a moral distinction between them. Crucially though, if Rachels is right, any such moral distinction can't be due to a distinction between killing and letting die.

Questions

1 What is 'euthanasia'?
2 Distinguish active and passive euthanasia.
3 Can you think of potential problems in demarcating these types of euthanasia?
4 Are there any circumstances in which you might think that killing is sometimes morally preferable to letting die?
5 Do you think that Smith and Jones are equally wrong? If you don't, what makes the difference?
6 Do you agree that Smith and Jones's action can be split into intention, action and consequence?

7 Think of your own example to illustrate Rachels' point.
8 Do you think the fact that Rachels' example is about killing a cousin makes a difference? For instance would it make any difference if the person in the bath was a friend, or was 80 years old?
9 If Rachels' argument succeeds, are there any other ways of showing that active euthanasia is morally worse than passive euthanasia?
10 What is it for someone to *bring about* someone's death?

Reading

Beauchamp, T. L. 1978. 'A Reply to Rachels on Active and Passive Euthanasia', in *Ethical Issues in Death and Dying*, Englewood Cliffs, NJ: Prentice-Hall, 253.

Lowe, E. J. 1980. 'Active and Passive Euthanasia: An Objection', *Philosophy*, 55(214), 550–1.

Garrard, E. 2005. 'Passive Euthanasia', *Journal of Medical Ethics*, 31(2), 65–8.

Rachels, J. 1997. 'Active and Passive Euthanasia', in N. A. S. Jecker, A. R. Jonsen and R. A. Pearlman (eds), *Bioethics: An Introduction to the History, Methods, and Practice*, Sudbury, MA: Jones & Bartlett Learning, 1–82.

---

53 Charity

If Jill gives money to a charity we think she has done something good and commendable, something 'above and beyond' what can (morally) be expected of her. Crucially, however, if Jill *doesn't* give money to charity we *don't* think that she has done something *morally wrong*.

Famously, Peter Singer (1972) has argued that this is a mistake. Contrary to what we may think, giving to charity isn't something that is 'above and beyond' what can be expected of us. Rather, giving to charity is a moral requirement on us. So, if Jill buys some music rather than giving money to charity then she has done something morally wrong. People have a *moral obligation* to give to the poor. How might Singer get us to this conclusion?

Thought experiment

> A child is drowning in a shallow river. You can easily save him. If you do, you will miss a meeting and wreck a new pair of shoes, but that is all. You don't help the child. Have you done something morally wrong?

Singer suggests that in such a case most people will say 'yes'. If you can bring about some good (saving the child) at little sacrifice to yourself (missing a meeting and wrecking shoes) then you *ought* to do it. Failing to act in that way means you have done something wrong.

What has this got to do with giving money to charity? Well, people can give money to charity with little sacrifice to themselves, which will bring about a good – the relieving of needless suffering. Thus, every time people don't give money to charity, this is equivalent to the person walking past the drowning child.

There are a number of responses that people give. We might grant that in the child/river case we have done something wrong, but try to show there is a non-trivial distinction between that case and the case of giving money to charity.

- First we might say that Singer's conclusion is simply impractical. However, let's put this worry to one side. Why? Well, the *truth* of a moral theory is not dependent on how *easy* it might be to live by.
- Second, whereas the drowning child is in front of us, the suffering brought about by poverty is often outside our direct experience. However, Singer says this cannot matter in this case. It may be true that proximity has a great effect on how we *feel* and what we might, as a matter of fact, do, but it doesn't change the obligations we have. Distance from someone is an arbitrary thing and it seems unfair to use it as a base for the strength of an obligation.
- Third, in the charity case there are millions of people who could help, whereas in the child case there is just you. So we might think that this allows us to claim that there are different obligations. However, this won't work either. It is true that because lots of other people fail to give money to charity you might find it easier to justify *why* you did nothing. This psychological reason, though, doesn't show that *you* have less moral obligation.
- Finally, we might worry that, whereas in the child case there is a direct causal relation between your action and the relief of suffering, in the charities case there isn't. We don't know whether the money we give will actually help someone: it might pay someone's wages or pay for a TV advert. Rather than undermining Singer's argument this point shows that we should be responsible with our money. We should research the most trustworthy charities. The fact that someone might squander our contribution doesn't in itself show us that we didn't have an obligation to give the money in the first place.

So it seems that if you don't give money away to charity you have done something morally wrong.

Questions

1 What is 'charity'?
2 Do you give money to charity?
3 If you answered 'yes', do you think you are doing something morally right? If you answered 'no', do you think you are doing something morally wrong?
4 Do you think the person who walks past the drowning child has done something wrong? If so, why?

5 How big must the sacrifice be – e.g. the loss of a limb vs the loss of a pair of shoes – for someone who fails to intervene not to be doing something morally wrong?
6 Do you think that the proximity to the suffering in the drowning case means it is different from the charity case?
7 Do you think the fact that lots of people don't give money to charity makes a difference to whether you have an obligation to give money to charity?
8 If giving money away in the way that Singer suggests is really hard to do, does that have any significance on whether Singer is right?
9 If someone argues that, instead of giving their money to the poor, she will invest millions in property which in the end might create more money which would mean she could give more money away, how would you respond?
10 Do you think that this argument generalizes to giving away 'time' as well as 'money'? If so what are the implications of the argument in terms of time?

Reading

Narveson, J. 2004. 'Is World Poverty a Moral Problem for the Wealthy?', *The Journal of Ethics*, 8(4), 397–408.

Pogge, T. 2005. 'World Poverty and Human Rights', *Ethics & International Affairs*, 19(1), 1–7.

Satz, D. 2005. 'What Do We Owe the Global Poor?', *Ethics & International Affairs*, 19(1), 47–54.

Singer, P. 1972. 'Famine, Affluence, and Morality', *Philosophy & Public Affairs*, 1(3), 229–43.

---

54 Environment

Philosophers make a distinction between *instrumental* and *non-instrumental* values. Bikes, teapots and shoes have instrumental value because they help us get stuff done. The value of *humans* on the other hand is non-instrumental. The fact that humans behave in certain ways – being charitable, stealing, voting, etc. – or the fact that they fall under certain categories – being a man, a Muslim, being old, etc. – is irrelevant to whether they have value. If humans have value then we think they have it irrespective of who they are and what they do.

Debates in environmental ethics ask the question: 'does the environment have instrumental or non-instrumental value?' The traditional view was that *only* humans had non-instrumental value. The value that, for example, trees, dogs and hens have, if they have any at all, lies in the fact they enable humans to get certain things: e.g. paper, eggs, companionship, etc.

Environmental philosophers have questioned this. They argue that some aspects, or all, of the environment have non-instrumental value. The value of trees, say, does not lie merely in the fact that they produce paper or oxygen *for* humans. They are valuable in and of themselves.

Thought experiment

> Things are going really badly for you. Humanity has died out and you are the last human on the planet; there are plants, trees, mountains, dolphins, etc., but no humans. You've done all that you want to do and are now bored.
>
> Before humanity died out you were given privileged information regarding a massive nuclear device that is so far underground that if it was detonated it would destroy *everything* in the planet. The destruction would be so complete and utter that there would be no hope of any life ever regenerating and evolving. You can arm it in such a way that when you die the bomb will go off. If you connect yourself to the bomb do you think you will have done something morally wrong?

Arguably most people's intuition is that 'yes' if you arrange for the destruction of the planet – its animal and plant life, its mountains and canyons, etc. – after your death, you have done something wrong. But why? Why think it is wrong to destroy the animals, trees, mountains, etc? It is not as if humans would somehow be affected, because there are no humans left. But if there is value in the planet which isn't due to the effect the planet has on humans, then that value must be non-instrumental.

This structure of thought experiment can be used to isolate our intuitions on the value of different things: all one has to do is vary what is being destroyed – e.g. imagine that it was just the plant life, or just the animal life which goes up in smoke.

If this is right and the environment does have non-instrumental value then this raises other questions about the correct moral theory we should have; for example, if non-human things on the planet have non-instrumental value then might they have *rights* which our moral theory has to recognize and respect?

Questions

1. How do you understand the term 'the environment'?
2. Explain the difference between instrumental and non-instrumental value.
3. List some things that you think have instrumental value and some things that you think have non-instrumental value.
4. Do you think that 'the environment' has non-instrumental value?
5. Do you think that if you set off the nuclear device in the thought experiment, you will have done something morally wrong?
6. Imagine you shared the intuition that the last person has done something morally wrong. Is there a way of explaining the intuition which doesn't rely on the environment having non-instrumental value?

7 If the environment has non-instrumental value, then how might this change how we lived our lives?
8 Could it be the case that only some aspects of the environment have non-instrumental value? If so, which parts?
9 If your intuition is that the last person did something wrong, what could you vary on the planet in such a way that what the last person did wasn't wrong? (Apart from not destroying the planet.)
10 How, if at all, do aesthetic values relate to environmental values? We might think that one reason that it would be wrong to destroy a painting is that it has aesthetic value. Could we say that one reason it would be wrong to destroy, say, a tree is because it has aesthetic value?

Reading

Fitzpatrick, W. J. 2004. 'Valuing Nature Non-Instrumentally', *The Journal of Value Inquiry*, 38(3), 315–32.

Lee, K. 1993. 'Instrumentalism and the Last Person Argument', *Environmental Ethics*, 15(4), 333–44.

Peterson, M. and Sandin, P. 2013. 'The Last Man Argument Revisited', *The Journal of Value Inquiry*, 47(1–2), 121–33.

Routley, R. and Routley, V. 1980. 'Human Chauvinism and Environmental Ethics', in D. Mannison, M. A. McRobbie and R. Routley (eds), *Environmental Philosophy*, Canberra: Australian National University, Research School of Social Sciences, 96–189.

---

55 Animals

Due to philosophers, non-human animals are thought of as having some moral significance. Society has become more sensitive to talk of 'animal rights' and 'animal welfare', and how, for example, morality might relate to whether we test medicine on animals. These thoughts have made us question the idea that humans are of special value simply by being part of the species *homo sapiens*. The shift away from an anthropocentric point of view can be illustrated by a thought experiment which looks at whether eating meat is morally acceptable:

Thought experiment

> Imagine that at some future time human cannibalism is widely practised. People are, without warning, plucked from their homes and served up in various meat dishes.
> Is this future society doing anything morally wrong? Is it morally wrong to take people against their will and eat them?

The obvious and compelling answer is 'yes'. One cannot take humans, kill them and eat them. However, notice that a large number of people believe that it *is* quite all right to pluck non-human animals – cows, sheep, pigs, etc. – from their homes and eat them.

This means that someone who eats meat (you?) is forced into the following choice. Either it *is* morally acceptable to eat both humans and non-humans or it is not acceptable because there is an important difference between humans and non-human animals.

Which option do you think is the most plausible?

Let's assume that most would opt for the latter option. What then is the difference between human and non-human animals? One thought is that animals have a very low IQ and/or are not sophisticated language users. So perhaps this is the difference that means that we are morally entitled to eat them? Yet this cannot be right because, if it were, then it would be morally acceptable to eat humans with a low IQ and/or who weren't sophisticated language users – e.g. babies. But we don't think we are morally entitled to do that.

Perhaps the difference is that humans don't *agree to* being served up as burgers, and it is this that means it would be wrong to kill and eat them. But this doesn't help either, as animals don't agree to be killed and eaten. Perhaps the difference is the suffering and grief amongst the humans who are left behind: parents, siblings, friends, etc.; which is not the case when we eat non-human animals. Yet again this isn't going to help because a consequence of this line of thought would be that if someone *wasn't missed*, it *would* be morally acceptable to eat them. But we don't think this is right.

We might cite *potentiality* as the difference. It seems that humans have a potential for a long and happy life and you are depriving them of this when you eat them. This means it is wrong to eat them. However, this isn't going to work either. It would license eating very old and terminally ill people, for they don't have a potential for a long and happy life.

After thinking hard about this problem many conclude that although we can't find a significant difference between humans and non-humans, we can't eat humans because they are *humans*. That is, the fact that a creature is from one species rather than another means we can treat it differently. Yet there is a significant problem here. Peter Singer (2012) points out that it is a principle of morality that we should not make different value judgements based solely on *arbitrary* features like, for example, being black, white, a man, a woman, etc. After all, this is what is morally objectionable about sexism, racism, etc.

But if we are going to say that eating humans is wrong because humans are from the species *homo sapiens*, then this does seem as if we *are* making a moral judgement based on an *arbitrary* feature. We wouldn't say it is OK to eat white people because they are white, but we are saying that it is OK to eat, say, a sheep, because it is a sheep and from a particular species. Singer coins the term *speciesism* for the view that humans have greater value than non-humans because humans are *homo sapiens*. He thinks there is no moral justification for being speciesist.

It seems then that if we want to continuing eating meat we are forced into a choice. Either be accused of speciesism – and this seems as objectionable as racism – or accept that it is morally acceptable to eat humans.

Questions

1 Are there any other ways of distinguishing humans from non-human animals, such that we can say that eating humans is wrong but eating non-humans isn't?
2 Is making a moral distinction based on race as morally wrong as making a moral distinction based on species?
3 What might a religious person put forward as a defence for eating meat? What might Singer say in response?
4 We could accept that it is morally acceptable to eat non-human animals and bite the bullet and claim that it is also acceptable to eat humans. What do you think about this line of reasoning?
5 If we want to avoid speciesism ought we to be vegans?
6 If we want to avoid speciesism could it be acceptable to test medication on animals?
7 If we want to avoid speciesism then would our moral domain also have to be extended to non-animals, e.g. to trees?
8 Is it right to save a Muslim over a Christian simply based on religion?
9 How does your answer to the last question relate to saving a human over a non-human animal? (See Chapter 65.)
10 If Singer is correct about speciesism then how would that change the way you lived your life?

Reading

Cohen, C. 1986. 'The Case for the Use of Animals in Biomedical Research', *The New England Journal of Medicine*, 315(14), 865–70.

Diamond, C. 1978. 'Eating Meat and Eating People', *Philosophy*, 53(206), 465–79.

LaFollette, H. and Shanks, N. 1996. 'The Origin of Speciesism', *Philosophy*, 71(275), 41–61.

Singer, P. 2012. 'All Animals Are Equal', in R. Shafer-Landau (ed.), *Ethical Theory: An Anthology*, Oxford: Blackwell, 14, 361.

---

56 Human cloning

There have been major advances in genetic engineering over the last twenty years. Scientists have successfully cloned animals – Dolly the sheep being the most famous example. In theory it is possible to clone humans. But should we?

Thought experiment

> Tragically you and your husband have recently lost your only child, Billy. Billy was the result of many years of IVF fertility treatment. You have been told that it is impossible that you could naturally conceive any other children. Heartbroken, you give up all hope. However, a phone call from the local hospital offers you a chance of having another child. They say they can clone your husband.
>
> Essentially they will take your husband's DNA and put it in one of your cells; then through some chemical intervention they will 'kickstart' the duplication process and implant it in your womb. Basically, in terms of the biology – you will be giving birth to your husband's identical twin. Apart from adoption or fostering, this is your only hope of more children. In deciding whether to proceed, what are the moral issues that you need to consider?

A few things need to be made clear in such a discussion. Fuelled by sci-fi films and books, there are many misconceptions when considering human cloning.

First as highlighted by this story there are already naturally occurring human clones; namely, identical twins. It is interesting that even those twins who are very hard to tell apart in physical terms have very different personalities, preferences, hopes and dreams. Just because Smith is a clone of Jones it doesn't follow that because Smith likes philosophy, fine wine and clay pigeon shooting, Jones will have the same preferences. That will depend on where Jones was brought up, who brought him up and all the other things that Jones will see and hear in his life.

Second, it is often forgotten that clones need to be implanted in a womb and a woman needs to give birth to them. It is not as if a human could be put into a giant photocopier and a life-sized version of that human would shoot out of the other side. This is important, because some people think that human cloning is wrong because it would allow an evil scientist to quickly create armies of followers. But there is no reason to think this would be any quicker or easier (in fact it would be slower and harder) than simply brainwashing current populations in the usual way through propaganda, etc. For the evil scientist would not only have to create, run and maintain the cloning programme, find and 'convince' women to grow and give birth to the clones, but also, when the clones are born, have to bring them up in such a way that they follow the evil scientist.

There is another issue here which needs to be separated from the issue of human cloning. We might think that it is morally wrong to destroy fertilized human eggs. Perhaps, for example, we think that human life is sacred and that as soon as a sperm and egg join we have something that we need to consider as having rights and a moral value. If this is true then we might be tempted to object to cloning on the basis that the amount of fertilized human eggs that would have to be destroyed in order to get the process right and create viable clones would make human cloning objectionable.

However, in order to be good philosophers, we should grant the defender of human cloning the claim that the technology works well and then ask – would human cloning be morally wrong? We might return to this worry about the destruction of many fertilized eggs later in an argument, but to be charitable we need to assume that human cloning has been perfected and then ask whether, if we were to clone a human, we would do something morally wrong.

Finally, probably the two most frequent responses given in this area are that human cloning is wrong because it is 'messing with nature', and/or because it is 'playing God'.

However, there are lots of things we think are morally acceptable but which count as 'messing with nature'. For example, antibiotics, life-support machines, anaesthesia, etc. So to defend the 'messing with nature' point we'd have to make some *non-arbitrary distinction* between *cloning*-messing-with-nature and *non-cloning*-messing-with-nature such, that we could say that one is wrong and the other is not.

The 'playing God' problem is similar. If it just means we shouldn't play God because this is 'messing with nature' then we have the same response as above. If it is something more specific to do with creating life, then clearly we need a story to tell about things such as contraception, IVF, fertility supplements, etc. At the least, the defender of the view that it's wrong because it is 'playing God' needs to tell us more in order for us to fully understand the problem.

Questions

1 What is cloning?
2 How does a clone differ from a 'normal' person?
3 Some people think that human cloning is (i) 'playing God' and (ii) 'messing with nature'. What might these ideas mean? What, if anything, would be wrong with 'playing God' or 'messing with nature'?
4 If someone had a large family and wanted a clone because it was the 'cool' thing to do, should that be allowed?
5 If someone wanted to have a clone for the primary reason that the clone would have the right bone marrow needed to cure a family illness, should that be allowed?
6 If you think there are good reasons for being allowed to clone and bad reasons for being allowed to clone, what makes those reasons good or bad?
7 Some countries provide a free healthcare system (e.g. the National Health Service in the UK). Assuming that is morally permissible to have free healthcare, should families be offered human cloning as part of that healthcare system?
8 Some people argue that cloning is wrong because in the wrong hands an 'army of psychopathic killers' could be created. Is this a good reason for banning cloning?

9 In the story, if the mother has a sexual relationship with the 'son', biologically speaking it would not be incest: after all it isn't biologically her son, it's her husband's identical twin. But would it be wrong?
10 Should science be allowed to discover and invent at will? Or should philosophers and others decide what parameters the scientists should work within? If you think the philosophers and others should have a role, how might that work in practice?

Reading

Burley, J. and Harris, J. 1999. 'Human Cloning and Child Welfare', *Journal of Medical Ethics*, 25(2), 108–13.
Harris, J. 1997. 'Goodbye Dolly? The Ethics of Human Cloning', *Journal of Medical Ethics*, 23(6), 353–60.
Häyry, M. 2003. 'Philosophical Arguments For and Against Human Reproductive Cloning', *Bioethics*, 17(5–6), 447–60.
Jonsen, Albert R. 2003. *The Birth of Bioethics*, Oxford: Oxford University Press, esp. chapter 9.

---

57 Jules Holroyd on blame and implicit bias

When is it acceptable to blame someone for their actions? This is a difficult question, and perhaps an easier question would be the reverse of this one. When are we prohibited from blaming someone? Well, it seems clear that it only makes sense to blame *agents*. Strictly speaking we can't blame the hailstones for the destruction of the crops or the mosquito for the spread of malaria. Of course in everyday conversation we might talk of blame, but this talk is just shorthand for 'x causes y'. After all, presumably we don't want to say to the mosquito: 'you really *ought not* to have done that'.

However, things regarding blame aren't as clear cut as they first seem. As Holroyd (2012) has argued, a good example illustrating this problem is what psychologists call 'implicit bias':

Thought experiment

> Judith is a well-respected academic who doesn't mark students' work anonymously. During one semester a student accuses her of favouring 'Western sounding' names in her marking practice and giving those essays higher marks.
>
> Judith is dismayed. She would never dream of this! She believes she treats everyone objectively and without prejudice. To clear her name she takes an Implicit

Association Test (IAT), which measures implicit bias. After taking this test it turns out that Judith does show a bias towards 'Western sounding' names. Upset she decides to mark anonymously in the future.

Later Judith is talking to her friends about this experience and they reassure her that she is not to blame for her marking bias, but Judith is unsure. Why isn't she to blame? After all it was she who read the work and marked it. What do you think? Do you think Judith is to blame?

Imagine that you think that Judith is *not* to blame for marking in a racist way, and that she shouldn't feel guilt and that she hasn't done anything morally wrong. You might, for example point out that the bias is *implicit*. You might think that Judith shouldn't feel guilty about her actions, and that people can't legitimately tell her she 'ought not' to have marked in that way. Perhaps Judith's actions are a bit like those of a dog that has been trained to attack humans. The animal isn't to blame for attacking humans; we don't think it should feel guilty; we don't think it makes sense to say that the dog 'ought not to have done that'. Rather, if anyone is to blame it is the people who trained the dog to attack humans. When Judith was marking, race wasn't something that she was conscious of, it wasn't an attitude that she had been cultivating and dwelling on over her years as a teacher.

But is this right? Is there more to say here? How might Judith's actions differ from Jim's? Jim was brought up in an explicitly racist household and holds racist views. Anyone who does anything 'non-Western' is treated in a negative way. These racist views have seeped into Jim's mind. Just as a daughter might like the football team her mother does, or think her dad's views on dance are the correct ones, so Jim hates the people his mother does and thinks his dad's view on immigration is the correct one.

Moreover, imagine that no one has ever challenged or educated Jim regarding race, that there has been no point which has ever caused him to reflect on his views. Is Jim to blame for being racist? If you think he is, what is the difference between Judith and Jim?

Questions

1 When is it acceptable to blame someone for her actions?
2 Can someone's upbringing mean that even though she is racist she is not blameworthy?
3 What is implicit bias?
4 Can we be mistaken about the mental states we have?
5 Take an implicit association test: https://implicit.harvard.edu/implicit/user/pimh/preliminaryinfo.html.
6 Read about Parks: http://www.psychologytoday.com/blog/the-superhuman-mind/201212/sleep-driving-and-sleep-killing. Do you think he is blameworthy for his actions?

7 Imagine that blaming people, even if they are not genuinely to blame, would eliminate implicit bias. Do you think we should blame people?
8 Could implicit bias ever be a good thing?
9 Could someone be blameworthy but not morally responsible?
10 How might this debate speak to the debate about free will?

Reading

Arpaly, Nomy and Schroeder, Timothy. 1999. 'Praise, Blame and the Whole Self', *Philosophical Studies*, 93(2), 161–88.

Holroyd, J. 2012. 'Responsibility for Implicit Bias', *Journal of Social Philosophy*, 43(3), 274–306.

Kelly, Daniel and Roedder, Erica. 2008. 'Racial Cognition and the Ethics of Implicit Bias', *Philosophy Compass*, 3(3), 522–40.

Talbert, Matthew. 2008. 'Blame and Responsiveness to Moral Reason: Are Psychopaths Blameworthy?', *Pacific Philosophical Quarterly*, 89(4), 516–35.

---

58 Catherine MacKinnon on gender and discrimination

If asked, most people would say that discrimination is wrong and that as reasonable enlightened individuals we should not discriminate against people based on such things as race, gender, sexual orientation, religion, etc.

However, Catherine MacKinnon (1987) argues that despite our good intentions we might unwittingly be perpetuating discrimination. Consider an example to illustrate her reasoning:

Thought experiment

> Bill is a boss of a small furniture sales business. He is educated and takes himself to be an enlightened person who doesn't discriminate. A vacancy has recently arisen in his company and he aims to be non-discriminatory in his hiring policies.
>
> Clearly his business has no legitimate grounds for discriminating on the basis of sex. After all both men and women can sell furniture! Bill writes what he takes to be a completely non-discriminatory job description. He asks for a second opinion from Human Resources and they agree it contains only gender-neutral language.
>
> As it turns out all the candidates that apply for the position are men, so unsurprisingly he hires a man. He reasons that he only hired a man because all the candidates were men and therefore he cannot have discriminated against women. How could he have done? He made sure that gender wasn't taken into account in making the hiring decision, and the language in the job description was gender-neutral. He is not in control of those deciding whether or not to apply for the job. Do you think that Bill could have discriminated against women?

Some think that it is possible that he is still discriminating. This is how: as a matter of fact, women take up the majority of care-giving roles for small children in society. This is how society expects women to behave and consequently it is how women are portrayed and discussed in the media – this 'moulding' is referred to as 'socialization'.

Imagine now that we look again at the job description and see that it contains phrases that suggest that there would be no flexibility surrounding child-care in the job (though this isn't explicit). For instance, the job description talks about how the successful candidate would need to travel extensively at short notice, and how the successful candidate should not expect any flexibility in their working hours.

If this is the case, and if it is women who are socialized into being the primary caregivers for small children, then it would obviously be unlikely that women would apply. So we might suggest that this is why the company ended up hiring a man. So even though Bill's intentions were good and he had not meant to discriminate, the fact that he had not recognized that women had care-giving roles made the job description discriminatory.

This line of thought is based on Catherine MacKinnon's (1987) distinction between 'difference discrimination' and 'dominance discrimination'. MacKinnon thinks that the *structures* of society can mean that women are discriminated against. So, for example, in a society where the best-paid jobs were incompatible with the primary care-giving roles, then *even if* gender-neutral language was always used and even if there was never *intent* to discriminate against women, this discrimination would still exist.

Questions

1 Do you think that the company discriminated against women?
2 Why do you think that women hold the majority of primary care-giving roles?
3 What is the difference between the 'difference' and the 'dominance' account of discrimination?
4 Do you think that the dominance account is genuine discrimination?
5 Assuming that dominance discrimination makes sense, can you think of any other examples of dominance discrimination?
6 What should Bill write so that his job description is non-discriminatory?
7 If MacKinnon is right and there is such a thing as dominance discrimination, how might we identify it?
8 Once we have identified it, then how might we make moves to counter it?
9 In the dominance account of discrimination, is anyone to *blame*? (See also Chapter 57.)
10 Imagine that Bill's business was a very small start-up business. After his accountant has done the sums, she tells Bill that it is not a viable option to sustain flexible working hours. Is Bill now justified in keeping the job advert as it is?

Reading

Kymlicka, W. *Contemporary Political Philosophy*, 2nd edn, New York: Oxford University Press, esp. chapter 9.

MacKinnon, C. 1987. 'Difference and Dominance', in *Feminism Unmodified*, Cambridge, MA: Harvard University Press.

Saul, J. 2003. *Feminism Issues and Arguments*, Oxford: Oxford University Press, esp. chapter 1.

Williams, J. 2000. *Unbending Gender: Why Family and Work Conflict and What to Do About It*, Oxford: Oxford University Press.

---

59 Potentiality and personhood

The issue of *potential* personhood is frequently discussed within applied ethics. For instance, it is sometimes argued that what is wrong with destroying a fertilized human egg is that it is a *potential person*; or that contraception is wrong because it stops people coming into existence.

However, philosophers have long been perplexed by this notion of 'potentiality'. To be slightly more formal, they have worried about the claim that an organism having certain *potentialities* is sufficient ground to ascribe to it *the right to life*. After all, we don't normally ascribe the same rights to potential things as we do to actual things. For example, we don't treat potential monarchs with the same rights as actual monarchs. Consider a thought experiment from Michael Tooley (1972) which challenges this claim about potentiality:

Thought experiment

> Imagine that there is a chemical called 'Potential', which if given to *a kitten* will mean that when it grows into a cat it will have the same psychology as a normal adult human. We would think that such *a cat* has *a right to life*. Moreover, we would think that it would be wrong to kill such a cat.
>
> Do you think that there is any *moral* difference between, on the one hand, failing to give a kitten 'Potential', and on the other hand, injecting a kitten which *has* been given 'Potential' with 'Neutralize'; where 'Neutralize' is a chemical which will neutralize the effects of 'Potential'?

Tooley thinks that you'll answer 'no' to this question. This is because he is relying on what he calls the moral symmetry principle. It claims that in two situations where the motivations are the same (stop the cat developing the right to life) and the effort involved is comparable (don't give an injection of 'Potential' vs give an

injection of 'Neutralize') then the fact that one is an *act* and the other is an *omission* cannot make a moral difference.

Where does this leave us? Tooley believes that we wouldn't think we'd done anything morally wrong if we failed to give a kitten 'Potential'. We don't have any obligation to give kittens 'Potential', simply because 'Potential' exists. If this is true, then given the moral symmetry principle, we *haven't* done anything wrong by injecting the 'Potential' kitten with 'Neutralize'.

Finally, one very simple way of neutralizing the effects of 'Potential' is to inject the kitten with a fast acting *poison*. So, given that it is not morally wrong to inject a 'Potential' kitten with 'Neutralize', then it is not wrong to kill a 'Potential' kitten.

Hopefully the conclusion is clear. The claim that if an organism has certain *potentialities* then that is sufficient ground to ascribe it the right to life is false, whether we are talking about a kitten injected with 'Potential' or a fertilized human egg. Therefore, if Tooley is right we can't object to the destruction of a fetus, or the disposal of many zygotes in an experiment, on the basis that they would have become people.

As we have seen from other chapters (Chapter 52) the moral symmetry principle is controversial. Moreover, even if Tooley is right it leaves many further questions in need of an answer. Here is one. If it *isn't* the potential for personhood that matters when ascribing the right to life, but rather it is actual personhood, then what properties matter for an organism to count as a person? It seems that if we think it is personhood that is morally significant, and not potential personhood, then we need an answer to this question.

Questions

1 What is it for something to have a 'potential'?
2 Can you give any examples of moral debates, other than abortion, where potential is used?
3 Do you think that is wrong to kill a human fetus?
4 If the kitten was injected with 'Potential' do you think it would be wrong to kill the kitten? If so why?
5 Do you think it would be wrong *not* to inject a kitten with 'Potential'?
6 What is the 'moral symmetry principle'?
7 Do you agree with Tooley that there is no moral difference between not giving a kitten 'Potential', and injecting a 'Potential' kitten with 'Neutralize'?
8 What do you think of the claim that there is no difference between giving a 'Potential' kitten 'Neutralize' and *killing* a 'Potential' kitten?
9 If it is personhood (rather than potential personhood) that matters, then what is it that makes someone a person?
10 Does your answer to (9) apply to *all* humans? Does it apply to any animals?

Reading

English, J. 1975. 'Abortion and the Concept of a Person', *Canadian Journal of Philosophy*, 5(2), 233–43.
Levin, D. S. 1985. 'Abortion, Personhood, and Vagueness', *The Journal of Value Inquiry*, 19(3), 197–209.
Marquis, D. 1989. 'Why Abortion Is Immoral', *The Journal of Philosophy*, 86(4), 183–202.
Tooley, M. 1972. 'Abortion and Infanticide', *Philosophy & Public Affairs*, 2(1), 37–65.

---

60 Capital punishment

Capital punishment is when the State tries someone for a crime, finds them guilty, and executes them as a form of punishment. It is still used across the world. The method of execution varies but common methods include firing squad, stoning, electric chair, beheading, hanging, gas chamber and lethal injection.

Some believe that this is the only fitting punishment for some crimes such as murder, rape, blasphemy or political dissent. However, others disagree. They argue that irrespective of what the law says there are never any morally justifiable grounds for the State killing an individual:

Thought experiment

> 'Did you hear about that murderer that was caught? He killed twenty people!'
>
> 'Yep, and he took a long time about it, torturing, raping and mutilating them.'
>
> 'He's confessed as well, they got him on video! It's an open-and-shut case.'
>
> 'He should be executed: jail is too good for him.'
>
> 'I sort of agree with you, but then what about that woman they caught a few weeks ago who confessed to killing her three severely disabled sons. Should she be executed?'
>
> 'No, that's different.'
>
> 'A murder is a murder isn't it? OK! What about the person in North Korea who published that blog ridiculing his leader. They executed him as a traitor. Was that morally acceptable?'
>
> 'Certainly not!'
>
> 'Ok, when is it morally justifiable?'

People's intuitions vary from case to case. Those who favour capital punishment for some crimes will claim that it is immoral in other cases. It is interesting to consider a number of different cases, such as the ones above, to try to find out what it is that people think is important and what is not important in these discussions.

The first thing to notice is that the question is about morality. This is different from the question 'when, if at all, should capital punishment be *legal*?' For example, we might think that it is morally acceptable under some conditions, but think that the law has no way of establishing when these conditions obtain; hence it should always remain illegal.

Some people rule out *all* forms of capital punishment, basing this on a belief in the *sanctity of life*, claiming that it is never morally acceptable for the State to kill someone, because life is sacred; it is only God who can take away life, and it is always morally unacceptable for the State to do so. Or a more secular version would be that everyone has the *right to life* which the State should recognize and protect. And importantly, given that it is *a right*, it isn't something that can be lost or forfeited.

Another approach to the discussion is via consequentialist considerations (see Chapter 65). That is, we might point to the consequences of allowing capital punishment, and hence rule out/justify capital punishment. Taking this line of thought it could be claimed that if the State legitimizes killing, this might have the effect that citizens would think life is less valuable. Also, if the State allows capital punishment, then it might follow that the State would start to kill 'unwanted' people; people with certain religious or political outlooks, for example.

Again, another consequentialist point – but *in favour* of capital punishment this time – might be that we should allow it because it is a *deterrent* and in fact overall would reduce the amount of serious crime. Notice that these responses are hostage to the empirical evidence. For instance, we might look to those countries which *do* have capital punishment and ask about crime figures, or about whether they have started to kill off 'unwanted' people, etc.

One of the most powerful arguments against capital punishment is as follows. It is clearly wrong to kill innocent people. There is never any moral justification for this. Moreover, there is no way to guarantee that if capital punishment is allowed only guilty people will be killed. So, if by having capital punishment we are allowing the possibility that the State might kill an innocent person, then we can never justify capital punishment.

There is a more general issue here which relates to a question about what are the morally acceptable grounds for punishment; the line of thought runs that we would find it morally objectionable for a mob to kill someone they believed to be a murderer (even if we were sure that they were right in their judgement) for this is just vengeance/revenge. But, the argument would run, capital punishment is just precisely that, a form of vengeance/revenge dressed up in laws and formal rituals. Hence capital punishment is morally wrong.

Questions

1 Find out which countries allow capital punishment. What forms do they use? How many people did they execute last year?

2 If you think that capital punishment is morally acceptable, then under what conditions? If you think that it is never morally acceptable, then why not? For example, does the form of execution matter? Does the age/gender of the executed person matter? Does the type of crime matter?
3 What do you think about the idea that capital punishment should not be allowed because it devalues life?
4 What do you think the link should be between what is morally right/wrong and what is legal/illegal?
5 Read Chapter 57 on implicit bias. How might this be relevant in this debate?
6 If they could make capital punishment instantaneous and painless would that change your view on whether it is acceptable or not? Explain your answer.
7 Is revenge ever a legitimate moral reason for punishing someone? If you do not think it is, then does that rule out capital punishment? If you think it is, then how might you respond to people who think revenge (think of a mob attacking a murderer) is morally unacceptable?
8 What do you think about the claim that punishment should 'fit the crime'? Is that a useful idea?
9 Do you think that if you commit a crime you should forfeit some of your rights? If you do, then do you think that you should ever forfeit your right to life? If not, why not? If you think you should do, under what conditions?
10 Do you think that locking someone up in prison for the rest of their life is morally acceptable?

Reading

Aspenson, Steve. 2013. 'The Rescue Defence of Capital Punishment', *Ratio*, 26 (1), 91–105.
Brooks, Thom. 2004. 'Retributivist Arguments Against Capital Punishment', *Journal of Social Philosophy*, 35(2), 188–97.
Cholbi, Michael. 2006. 'Race, Capital Punishment, and the Cost of Murder', *Philosophical Studies*, 127(2), 255–82.
Devine, Philip E. 2000. 'Capital Punishment and the Sanctity of Life', *Midwest Studies in Philosophy*, 24(1), 229–43.

Part VII

NORMATIVE ETHICS

61 The experience machine

What is it that makes life worthwhile? Well, life *isn't* going well if we are ill, or if someone has hurt us physically or emotionally; whereas when we experience pleasure in reading a good book or spending time with loved ones, we *do* think life is going well. This might lead us to think that what makes life worthwhile is pleasure, and the avoidance of pain. However, this view faces a serious challenge from Robert Nozick (1974):

Thought experiment

> Imagine that some very clever psychologists have designed an 'experience machine'. You are given an opportunity to be plugged into this machine and experience *anything* that you want. It is exactly as you would experience in real life but all the time you would be hooked up to a machine and floating in a tank.
>
> This machine is completely programmable. You can experience winning a gold in the Olympic 100 metres, or writing a best-selling novel, or marrying on a space station, etc. Nothing will ever go wrong with this machine, so there is no risk.
>
> The question Nozick asks is this. If you could choose to plug yourself into this machine *for the rest of your life* would you? Notice the question isn't 'would you like to try this machine out over the weekend' rather it is an all-or-nothing question.
>
> Nozick claims that our answer will be 'no' because he thinks most people, given the choice, would not want to be plugged into this machine. But what would your answer be?

If we think that life would not be as worthwhile being plugged into this machine then this arguably shows us something very important about the view that what makes life living is pleasure. Namely, that it is false. The person plugged in would be given every opportunity to experience any type of pleasure, and yet this offer is rejected. The person obviously thinks that there is something missing.

One response is to simply bite the bullet and claim that it *would be* a better life if we plugged ourselves into the machine. If we took this route we'd need to try and explain why people don't come to that conclusion themselves. We could, for example, try to show that people resist being plugged in because they are scared of the technology, or that someone might tamper with the machine.

Another response would be to accept the conclusion and develop an 'objective' view by arguing that the reason we would resist being plugged into the machine is that some things are independently valuable and constitute a good life, irrespective of whether they are pleasurable. Such things might include meaningful relationships, being healthy or engaging in meaningful projects. There are three reasons that Nozick gives for thinking we wouldn't want to be plugged into the machine.

- First, we might think that it is more important to participate in the making of the experiences rather than just the experiences themselves. So for example, it is the action of climbing a mountain that we think gives meaning to our lives rather than the experience we get from climbing a mountain.
- Second, life in the machine would turn us into an indeterminate blob and therefore, we might not be able to be the sort of person that we had aspired to.
- Third, we might think that one reason we don't want to be plugged in is that it gives us no contact with any *deeper reality*. Rather we are just experiencing a man-made construction of reality.

Whatever we say, this thought experiment is a fundamental problem for some accounts of what makes life worth living.

Questions

1 Would you plug yourself into the experience machine?
2 Why do you think that many people would resist being plugged into the machine?
3 What do you think makes life worth living?
4 If you were attracted to the 'objective' view, what sorts of things do you think should go on the list as 'objectively' valuable?
5 What else can matter to us, other than how our lives feel from the inside?
6 One response suggests that what we desire is to live in 'contact with reality'. What is so good with being in contact with 'reality'?
7 Imagine that a machine could not only give us any experience we want but also could transform us into any person that we might want to be. Would you plug yourself into this machine?
8 Do you think that it could be the case that the experience of *really* writing a novel might be different from the experience of writing a novel in the experience machine?
9 Felipe De Brigard (2010) has developed an 'inverted experience machine' which runs something like this. Imagine that we are approached and told that everything we have experienced thus far is a program and in fact you are a rich person who is floating in a tank somewhere. Would you unplug from this experience machine?
10 Considering all that has been said, how would you answer the question: 'what makes life worthwhile?'

Reading

De Brigard, Felipe. 2010. 'If You Like It, Does It Matter if It's Real?', *Philosophical Psychology*, 23(1), 43–57.

Feldman, F. 2012. 'What We Learn from the Experience Machine', in R. Bader and J. Meadowcroft (eds), *The Cambridge Companion to Nozick's Anarchy, State, and Utopia*, Cambridge: Cambridge University Press.

Nozick, R. 1974. *Anarchy, State, and Utopia*, New York: Basic Books, 42–5.

Weijers, D. and Schouten, V. 2013. 'An Assessment of Recent Responses to the Experience Machine Objection to Hedonism', *Journal of Value Inquiry*, 2(9), 461–82.

62 Psychological egoism

The reasons why people act in certain ways seem many and varied. The reason that Rosa Parks acted as she did has something to do with God, injustice, inequality, etc. The reason that people are recycling more has something to do with the environment, future generations, etc. However, some people have argued that this doesn't quite get to the heart of the matter. The psychological egoist thinks that ultimately we are all motivated by self-interest:

Thought experiment

> You and a friend are walking home and see someone asking for money on the street. You pull out some loose change and give it to him. Later your friend asks why you did this. 'Well he obviously was cold and wanted some change for a coffee.' 'OK,' your friend responds 'but why did you want to help him get warm and get a drink?' '. . . well, it seemed that it was the right thing to do.' 'OK but why do you do the right thing?'
>
> You shrug your shoulders. You friend continues 'did it make you feel good to do what you thought was the right thing?' You think about it and admit that it did. 'So,' your friend continues, 'the real reason that you gave money away was because it made *you* feel good . . . it was really for reasons of self-interest that you acted as you did.' The rest of the journey home involves you trying to give examples of actions which are genuinely altruistic.
>
> Could you give any examples? What would you say to your friend?

Psychological egoism is the view that, despite what people might think, everyone's ultimate aim is their own self-interest. It is claimed that this is true even if it isn't obvious to people. Hence for the psychological egoist, there are no altruistic actions. Psychological egoism isn't then an ethical theory. It doesn't tell people how they *ought* to behave. How could it? After all, it is an 'a priori' claim about how all people as a matter of fact are motivated. All people *will* behave in a self-interested way, so saying anything about how they *ought* to behave is redundant.

Initially we might think it odd to make an 'a priori' claim about everyone's psychology. We might reasonably think that if a theory is about psychology and is about people then we should run empirical tests. We can consider the case of a man who jumps in front of a car to save a child and claim that this is a time when it seems the action has been taken for patently non-self-interested reasons. Of course, in these cases the psychological egoist will claim that unbeknown to the agents in these types of cases, they really are actually still acting for self-interested reasons. Yet this raises another concern because if there is no conceivable way of showing psychological egoism to be false, then we might wonder whether the theory is saying anything meaningful about the world at all.

Another response runs as follows. One might think that in order to achieve what is in our self-interest it is best *not* to be motivated by self-interest. For example, I ought to *desire to be a good friend* because in doing so I will be a good friend; and it is this which will be in my self-interest. But if I enter my friendship being motivated by self-interest this will probably mean an unsatisfying and short-lived friendship, which would *not* be in my interest.

A final point is the following. In the above example the friend pointed out that we felt good after performing an action and took this as evidence that all our actions are motivated by self-interest. However, just because we have good feelings *as a result* of an action it does not mean that we were actually motivated by such a feeling. Moreover, we might think that a person acting selflessly is precisely the sort of person who will feel good *because* he performs selfless actions.

Questions

1 Explain psychological egoism.
2 Why isn't psychological egoism a moral theory?
3 What is self-interest?
4 What, if anything, is the difference between self-interest and selfishness?
5 Think about why you do the things you do and ask yourself how reasonable do you think psychological egoism is?
6 Can you think of any actions which seem to be altruistic? How might the psychological egoist explain such actions?
7 Does the fact that we can't falsify psychological egoism undermine it as a theory?
8 Some people have claimed that psychological egoism is true in *most* – not all – cases. How believable do you find this 'watered down' version of psychological egoism?
9 Do you think it is true that by not acting in a self-interested way we are more likely to fulfil our self-interest?
10 Imagine that psychological egoism was true . . . what should we do now?

Reading

Beardman, S. 2012. 'Altruism and the Experimental Data on Helping Behavior', *Ethical Theory and Moral Practice*, 15(4), 547–61.

Butler, J. 1983 [1726]. *Five Sermons, Preached at the Rolls Chapel* and *A Dissertation Upon the Nature of Virtue*, Indianapolis, IN: Hackett Publishing.

LaFollette, H. 1988. 'The Truth in Psychological Egoism', in J. Feinberg (ed.), *Reason and Responsibility*, 7th edn, Belmont, MA: Wadsworth.

McConnell, T. C. 1978. 'The Argument from Psychological Egoism to Ethical Egoism', *Australasian Journal of Philosophy*, 56(1), 41–7.

---

63 Philippa Foot's trolley problem

Consider a thought experiment inspired by the work of Philippa Foot (1967):

Thought experiment

> Imagine that you are walking home and you see a tram hurtling down the local hill. The driver is nowhere to be seen and it is picking up some serious speed. To your horror you can see that there are five children playing on the track. They are totally oblivious to the danger. You calculate that given the size of the tram and given the speed it is travelling it will surely kill the children.
>
> You feel helpless about this. You can't perform any heroics, it is going too fast. However, you see there is a lever that you can pull which will divert the tram from the track and hence away from the five children. However, to your dismay you see two children playing on this forking track. You could pull the lever and the five will survive, but the two will surely die.
>
> Is it morally acceptable to pull the lever?

Most people's intuition is that they would think it is morally acceptable to pull the lever. It is a simple decision; you have to make a choice between five children dying and two children dying. If there is no other option then the right action is the one that *saves the most people*. So you ought to pull the lever to divert the tram. This seems fine. However, consider a famous variation on this case.

> Imagine that you are walking home and see a tram hurtling down the local hill. The driver is nowhere to be seen and it is picking up some serious speed. To your horror you can see that there are five children playing on the track. They are totally oblivious to the danger. You calculate that given the size of the tram and given the speed it is travelling it will surely kill the children.

> You feel helpless about this. You can't perform any heroics as it is going too fast. However, you see that there are two children playing in a toy car near the track. You reason that you could push the toy car under the tram and consequently stop it before it hits the five children. You realize that this means that the two children will die but it will mean that the tram will be halted from killing the five children
>
> Is it morally acceptable to push the two children in the toy car in front of the tram?

Most people's intuition is no. It is not morally acceptable to push children under the wheels of a tram even if it means that five others are saved. Yet there seems to be a problem here, for when we think about it, the two cases are significantly similar.

In the first case, our action of pulling the lever means the death of the two children in order to save the five children. We thought that this was morally acceptable. In the second we are acting (pushing the toy car onto the track), which means that the two children will die in order to save the five children. But we thought this was morally unacceptable. What are we to think? Is there any difference?

Questions

1. Do you think that it is morally acceptable to pull the lever?
2. Do you think it is morally acceptable to push the toy car under the tram?
3. If you think that there *is not* a moral difference between the cases would you think they are *both wrong* or *both right*?
4. If you think that there *is* a moral difference between the cases what grounds do you have for thinking that?
5. Some people think that the there is a difference between acting and failing to act and this is a morally significant difference. Do you agree?
6. If you choose not to administer an antidote to someone who was poisoned, then are you acting or failing to act?
7. Some people think that we should *not* intervene in the first case and let the five die because if you acted by pulling the lever then you are responsible for the two deaths whereas if you didn't act then you are not responsible for the five deaths. Do you agree?
8. How far do you think the notion of 'naturally occurring' or 'letting nature run its course' helps in moral decision making?
9. How much do you think numbers matter in these cases? Would you change your intuitions if, in the first case, it was three children you would save if you pulled the lever?
10. Some experimental philosophers have shown that the intuitions we have in these sorts of cases can be manipulated by changing random features. For example, if we change the gender of the children, whether the story is told in

first person or third person, or even the order in which we hear the stories. If this is true what do you think it tells us about our intuitions in these cases?

Reading

Edmonds, D. 2013. *Would You Kill the Fat Man? The Trolley Problem and What Your Answer Tells Us about Right and Wrong*, Princeton, NJ: Princeton University Press.

Foot, Philippa. 1967. 'The Problem of Abortion and the Doctrine of Double Effect', *Oxford Review*, 5, 5–15.

Thomson, Judith Jarvis. 1976. 'Killing, Letting Die, and the Trolley Problem', *The Monist*, 59(2), 204–17.

Thomson, Judith Jarvis. 2008. 'Turning the Trolley', *Philosophy & Public Affairs*, 36(4), 359–74.

---

64 Moral luck

Consider four cases inspired by the philosophy of Thomas Nagel (1993):

Thought experiment

1. You get on your bike after a night out at the pub. You haven't been drinking alcohol, you are fully attentive, you aren't tired and you cycle home. The journey is uneventful and nothing happens.
2. You get on your bike after a night out at the pub. You haven't been drinking alcohol, you are fully attentive and you aren't tired. Unfortunately a young person steps out in front of you at the last minute and you crash into him and break one of his ribs and knock him unconscious.
3. You get on your bike after a night out at the pub. You *have* been drinking alcohol and are feeling quite drunk; your reactions are slower, you struggle home wobbling all the way. The journey is uneventful and nothing happens.
4. You get on your bike after a night out at the pub. You *have* been drinking alcohol and are feeling quite drunk; your reactions are slower, you struggle home wobbling all the way. Unfortunately a young person steps out in front of you at the last minute and you crash into him and break one of his ribs and knock him unconscious.

Consider whether you are more, or less, morally blameworthy in each of these cases.

Nagel suggests, and most agree, that we think that there is little difference between the moral blameworthiness in (1) and (2). Some may think that the mere fact that someone is hurt makes you blameworthy, but most think that given that the person stepped out at the last minute in (2), you aren't more morally blameworthy for your actions.

Now consider (3) and (4). Most people think that you would not be morally blameworthy in (3), in your first drunken cycle home. After all, nothing happened. However, most think you *are* morally blameworthy in (4) – the case where you are drunk and have run someone down. Nagel suggests that this reveals a puzzle regarding our moral intuitions about whether someone is morally blameworthy.

Why do we think that you, the cyclist, are *not* morally blameworthy in (2)? Well, the person's action was nothing to do with you and consequently it would seem strange to blame you. The broad idea is that if something is outside our control we can't be blamed for it. Nagel calls this 'the Control Principle'. All this seems fine and unproblematic. However, now consider cases (3) and (4). In (4) the person's action was nothing to do with you; however, it seems that we *do* think that if we have been drinking we are significantly more morally blameworthy.

Or consider another Nagel-inspired example. A mother leaves a baby in the bath to get a cup of coffee. In the first scenario she is lucky and the baby is playing and having fun when she returns. In the second case she is *unlucky*, as on her return she finds that the baby has drowned. We think that in the second scenario she is morally blameworthy whereas in the first she is not**.** It seems that due to luck we can be more or less morally blameworthy. It seems then that there is such a thing as *moral luck*. Now, if the Control Principle is correct then you ought *not* to be blameworthy in case (4); nor should the mother in the baby drowning in the bath case. After all, it was not *because* you had been drinking that the person stepped out, or because the mother wanted a coffee that the baby slipped. So according to the Control Principle in neither of these cases is the agent blameworthy. However, we do have very strong intuitions that they are, so what is going on?

There has been a lot written about moral luck, and there are two typical ways of responding. In the first we could claim we are simply mistaken to think that you and the mother are more morally blameworthy and that moral luck has nothing to do with moral evaluation. One way of expanding this is as follows. What we need is a way of explaining why we feel differently about cases (3) and (4) and if we can explain the difference without claiming that one action is morally worse than the other then it seems we've respected our intuitions without accepting there is such thing as moral luck.

One way of doing this (see Latus 2000) is to claim that the reason we feel differently is because you are the *type of person* who would cycle home drunk or that the mum is the type of mum who would endanger the baby's life. Importantly, this would not have been revealed to us if the baby hadn't drowned or the person had not been knocked down. So we have explained the difference in our feelings about the cases, without thinking that there is such thing as *moral* luck. We simply know more about the agent.

The second general way of responding is to say that what matters in moral evaluation are the consequences, that is all. So, the only question to ask when making a moral evaluation of an action is with regard to the consequences. If, in our cycling example, the person is hurt, then you *are* morally blameworthy, even if the fact that

he was hurt was outside your control. On this approach the Control Principle has nothing to do with moral evaluation.

Questions

1 What is moral luck?
2 Create your own examples of so-called 'moral luck'.
3 What is the Control Principle?
4 What do you think is the best way of responding to this conundrum?
5 What would our moral practice be like if we rejected the Control Principle?
6 How does moral luck fit with punishment?
7 How might the moral luck debate fit with the free will/determinism debate?
8 How might discussion of moral luck fit with the concept of fairness?
9 What is the link between causation and responsibility?
10 How might a theist respond to this puzzle regarding moral luck?

Reading

Browne, B. 1992. 'A Solution to the Problem of Moral Luck', *The Philosophical Quarterly*, 42(168), 345–56.

Latus, A. 2000. 'Moral and Epistemic Luck', *Journal of Philosophical Research*, 25, 149–72.

Nagel, T. 1993. 'Moral Luck', in Daniel Statman (ed.), *Moral Luck*, Albany, NY: State University of New York Press, 57–71.

Williams, B. 1981. *Moral Luck*, New York: Cambridge University Press.

---

65 Consequentialism and favouring others

When we ask the question 'why is this action wrong?' or 'why is that action right?' our reply naturally makes reference to consequences. If we think that pain is bad and pleasure is good, then that naturally gives us a way of working out what actions we should take. It seems that if we save ten people over two and we assume that being saved gives people pleasure then, all things being equal, we have done the right thing. Simply put, ten lots of pleasure are better than two.

Moreover, given that *anyone* can feel pleasure/pain there is no reason to favour any particular person. Very broadly, these two points make up the heart of the ethical position called consequentialism. The consequentialist claims that the right action is the one which brings about the best consequences and when weighing up the consequences we should treat all people equally.

There has been a great deal written about consequentialism and there are many things which *prima facie* look both right and wrong about it. However, let's consider one particular issue regarding impartiality in moral decision making:

Thought experiment

> You are walking along the canal and spot two people struggling in the water and shouting for help. You run to them and are just about to pull them out of the water when you hear a very familiar voice shouting in desperation a close distance away in the water.
>
> To your horror you realize that it is your mother shouting and worse still, it is obvious that you can't save these two strangers *and* save your mother. You make the decision to leave the strangers and run to save your mother. The strangers drown and your mother lives. You believe that you made the right moral decision given the terrible situation in which you found yourself.
>
> Years later you are recounting this to a consequentialist friend. She tells you that even though she can perfectly understand *why* you saved your mother, your action was in fact morally wrong.
>
> Taken aback you protest . . . surely you haven't done anything wrong! Moreover you think that if she says that favouring your mother was wrong then that shows she needs to think again about her theory. Unmoved, the consequentialist asks you to consider a different scenario, which she hopes will convince you otherwise.
>
> Imagine that all things were the same as in the previous case except it was a choice between two strangers and a life partner. Imagine you save your partner. She asks if you think you've done something wrong? You are slightly less sure but maintain that you haven't. She smiles and gives you another example.
>
> Imagine that you are in the Ku Klux Klan and are walking along the canal. You see two black strangers drowning and one white stranger. You decide that you'll save the one white stranger over the two black strangers. The consequentialist asks you 'Can such an action be wrong?' You are sure that this would be the morally wrong decision and that in this case you should simply save the most lives as possible – the two black strangers. Your consequentialist friend then asks you to give her a good reason why you treated the three cases differently. What are you going to say?

It is surprisingly hard to give an answer to this question. The issue here is very complex and gets to the heart of many normative ethical theories and controversies. We can vary these examples to illustrate how thorny this issue is. Imagine that you hate your mother in the above example; or that you've only met and married your partner the morning you see him fall in the canal; or imagine that you were adopted and that you arranged to meet your mother – basically a stranger to you – for the very first time at the canal and then she falls into the water? Or imagine you are a devout Muslim and it was a choice between saving two non-Muslims or one Muslim. If feelings such as love matter then how much do they matter? Why does simply being biologically linked to someone make a difference? If it is about closeness to a person, then why doesn't race or religion matter?

Maybe given this murkiness the only right way to proceed is *not to favour anyone* in moral decision making. Maybe in the end, your consequentialist friend is right and the right way to proceed is to accept consequentialism because our intuitions are unreliable and fickle and shouldn't be given any weight when constructing moral theories. But, of course, if this is true then you were wrong to save your mother.

Questions

1. What is consequentialism?
2. Do you find consequentialism intuitively plausible? If not, why not?
3. Do you think that favouring a family member over two strangers is morally acceptable?
4. If your answer to (3) is 'yes' then is there a point where the number of strangers is so great that it would be wrong to save a family member? If your answer is 'no' then what justifies your response?
5. If your answer to (4) is 'yes' then why did you pick the number that you did?
6. If you answered 'yes' to (3), then why?
7. *If* it matters whether the person is your mother, then do all family relations count? For example, if it was a cousin you saved over two strangers would that be acceptable?
8. Would it be OK to favour someone from your own religion? If not, why not?
9. Imagine that the choice was between your pet dog (which is your only friend in the world) and a stranger. If you saved your dog have you done anything wrong?
10. Consequentialism can seem intuitive, and saving a family member over two strangers seems intuitive. But what role, if any, should intuitions have in our moral theorizing?

Reading

Arneson, R. J. 2003. 'Consequentialism vs Special-Ties Partiality', *The Monist*, 86, 382–401.
Driver, Julia. 2005. 'Consequentialism and Feminist Ethics', *Hypatia*, 20(4), 183–99.
Hurka, T. 2006. 'Value and Friendship: A More Subtle View', *Utilitas*, 18(3), 232–42.
Railton, P. 1984. 'Alienation, Consequentialism, and the Demands of Morality, *Philosophy & Public Affairs*, 13(2), 134–71.

---

66 Higher and lower pleasures

The utilitarian Jeremy Bentham suggests that we can calculate which action is the right action to take by asking which brings about the greatest amount of *pleasure*. If action x gives three people pleasure, but action y gives ten people pleasure then

action y is the morally right action. Bentham's philosophy only makes a distinction based on the *quantity* of pleasure, which has led to some criticism of his work. After all, we might think that it matters *what* the pleasure is taken in. For example, consider a sadist trying to decide whether to kick a cat, which would give him lots of pleasure, or giving money to charity, which would give him far less pleasure. If quantity of pleasure is all that counts then the right action is to kick the cat; but this seems wrong.

It is for this reason that John Stuart Mill introduced a distinction not only in the quantity of pleasure, but also in the 'type' of pleasure. He thinks there is a distinction between what he calls 'higher' and 'lower' pleasures. But is there such a distinction?

Thought experiment

> With time to kill, you and your friend get into a discussion about what music you'd take to a desert island. You think you would take the complete works of ABBA but she would take the complete Wagnerian Ring Cycle. However, what starts out as a bit of fun quickly develops into something more serious, and your friend not only thinks you are misguided but actually wrong.
>
> 'Look,' your friend says, 'surely the pleasure you get from listening to Wagner is better because it is "higher." Although it may not be as intense and you might not enjoy it as much, you will, in actuality, get a far greater pleasure from Wagner.'
>
> You tell her she is a snob and listening to ABBA will be great and actually you do find this a 'higher' pleasure. She points to the fact that Wagner is cited as 'high culture' by many experts and these people spend lifetimes studying his work.
>
> You remain unmoved, but there is something nagging at you. Does she have a point? Is the pleasure the experts experience somehow 'higher'? Having thought about it, you realize that you scoff at the current 'pop' charts not merely as something you don't enjoy but as something that is somehow 'beneath you'. And, come to think of it, you do think that people are 'wasting their life' watching reality TV, even if those people experience great pleasure.

If you think there is a distinction between higher and lower pleasure, like Mill, then you might think that your friend has a point. For Mill there are some 'higher' pleasures and some 'lower' pleasures. These aren't about intensity of feeling but about worthiness. Maybe then, given the choice between ABBA and Wagner, the right action is to listen to Wagner even if ABBA gives you a more intense pleasure.

But why think that there is such a distinction? Mill suggests that we have good reason to think that there might be this distinction because, if we ask a competent judge what she would prefer, she would prefer the higher pleasure. Who are competent judges? Well, people who have experienced *both* types of pleasures.

Of course, we might simply deny this. But notice there does seem to be some convergence/agreement on what counts as higher pleasures. Notice, Mill isn't saying we should never eat fast food, watch cartoons or listen to ABBA. He is saying that, on reflection, if we experienced the pleasure gained from these, and the

pleasure gained from, say, watching a performance of Shakespeare, we would prefer the latter, and this shows that there is something more worthwhile in the pleasure we get from some experiences – and not simply because one lasts longer than another.

Questions

1 Would you prefer to watch a TV program or watch a Shakespeare play? Explain your answer.
2 Do you think that there are 'higher' and 'lower' pleasures?
3 Why might someone think there are 'higher' and 'lower' pleasures? What examples might she give?
4 Mill thinks that, given that there is a distinction between 'higher' and 'lower' pleasures, it is better to be a human experiencing a very small amount of higher pleasure than a pig experiencing loads of lower pleasure. Do you think this is right?
5 What do you think of Mill's 'proof' for the existence of higher and lower pleasures?
6 Could *anyone* count as a competent judge?
7 If there is no such distinction, why is there convergence in what people often take as 'worthwhile' pleasures?
8 Is it possible that out of all those people who experienced pleasure both from eating a burger and from watching Shakespeare that there may be one/some that would prefer eating the burger?
9 Are higher pleasures about cultural conditioning? If they are, does that mean that Mill is wrong?
10 Mill says that even though higher pleasures are less intense they are more pleasurable – can you make sense of that idea?

Reading

Bentham, J. and Mill, J. S. 2004. *Utilitarianism and Other Essays*, Harmondsworth: Penguin, esp. *Utilitarianism* II, 5.

Gibbs, B. 1986. 'Higher and Lower Pleasures', *Philosophy*, 61(235), 31–59.

Ryberg, J. 2002. 'Higher and Lower Pleasures: Doubts on Justification', *Ethical Theory and Moral Practice*, 5(4), 415–29.

Schaupp, K. 2013. 'Books Before Chocolate? The Insufficiency of Mill's Evidence for Higher Pleasures', *Utilitas*, 25(2), 266–76.

---

67 Ring of Gyges

Why do people act morally? Why is it that we keep our promises, tell the truth, do not steal, etc? One answer is that we act morally not because we are naturally

inclined to do so, but rather that we act morally because it will help us gain a good reputation, because it will help us get what we want and avoid punishment.

So, another way of tackling this is to ask the following question: 'is acting morally important only in as far as we are *seen to be* moral? Or is there something intrinsically valuable about acting morally?' In Plato's *Republic* (359d–360b) Glaucon uses a very famous thought experiment to suggest that morality is only valuable because of what results from behaving morally. It only has instrumental value rather than intrinsic value. Plato spends time in the Republic developing an argument in order to respond to Glaucon's challenge:

Thought experiment

> You are a shepherd, and during a terrible storm a cave opens up in the land. You descend into the cave. In the cave there is a hollow brazen horse and in the horse is a corpse wearing a ring. You decide to take the ring and wear it.
>
> You attend various meetings with the king of Lydia. Bored at one meeting you are messing with the ring on your finger and turn it inwards. In doing so something odd happens. The people around you act as if you have left the room. They talk about you as if you weren't there; they talk across you and don't ask for your views, etc.
>
> Over the next week you further experiment with the ring and realize that if you twist it inwards you disappear and if you twist it outwards you reappear.
>
> You hypothesize that with this power of invisibility you could probably seduce the king's wife, kill the king, become king and own the lands. You will not be caught because no one will know that it is you performing these deeds. Basically you could get away with anything you want. Would you put your plan into action?

Glaucon thinks that you would. He thinks that if you had this power, together with the assurance that you would not be caught, you would be freed from the shackles of morality and free to act the in way you want. He thinks that this proves that *why* we act morally is because we do not have the power to avoid getting caught. If we were certain of escaping detection then we would act immorally. We act morally only because if we don't, it will ruin our reputation, we will face punishment, etc. Essentially then, morality has value, but that value is purely practical, and measured in terms of how much it can benefit us.

Questions

1. Would you act immorally if you had the ring of Gyges?
2. Some people find this thought experiment unsettling – why might this be the case?
3. If you had the ring of Gyges would you still think some actions were 'off limits'? If so, why?
4. Would you act immorally if you knew that everyone else had access to the ring?

5 Do you think that the same point can be applied to the legal laws of the land, rather than moral laws? That is, would you be more or less likely to break the legal laws if you had the ring?
6 Irrespective of how we might actually act in the example, how do you think we *ought* to act in the example?
7 What, if anything, does this thought experiment tell us about human nature?
8 If Glaucon is right and we only act morally because we don't have the power to remain undetected, then does this help us in any practical way? For example, in designing education programmes, or about how to parent children, etc.
9 How might a theist respond to this thought experiment?
10 What might this thought experiment predict about our 'surveillance society'?

Reading

Batson, C. D. 2008. 'Moral Masquerades: Experimental Exploration of the Nature of Moral Motivation', *Phenomenology and the Cognitive Sciences*, 7(1), 51–66.

Copp, D. 1997. 'The Ring of Gyges: Overridingness and the Unity of Reason', *Social Philosophy and Policy*, 14(1), 86–106.

Matthews, G. B. 2000. 'The Ring of Gyges', *International Journal of Applied Philosophy*, 14(1), 3–11.

Plato. *Republic*, esp. 359c–360b.

Tiberius, V. 2013. 'Why Be Moral? Can the Psychological Literature on Well-Being Shed any Light?', *Res Philosophica*, 90(3), 347–64.

---

68 Deontology

Deontologists believe that when we are trying to work out which action to perform in a moral decision we need only concern ourselves with the *action itself*. That is, we ought not to be concerned with the consequences which may or may not follow. For instance, if 'telling the truth' is right, then for the deontologist no further thought about the rightness or wrongness is required, as the action is right.

That is, even if telling the truth brings about great sadness or suffering; or indeed, if telling the truth *never* brought about *any* benefit to *anyone*. This seems very odd indeed. How could it not matter *at all* what happens when we act, as long as the act itself is right action?

To understand this, we need to consider a number of problems we get if we focus on consequences. First, imagine that an action is right or wrong depending on the *consequences* of that action. If this is true, *and we don't have complete control over the consequences* of our actions, then it follows that we don't have control over whether our actions are right or wrong. So an action could be wrong, irrespective of how we behaved.

Second, related to Chapter 65, if consequences are what matter, then we can't give any special weight to actions even if they help people we love, family members and friends. For instance, if we could save our father or *two* strangers, the simple consequentialist has to say that we ought to save the strangers. This may seem strange as it doesn't seem morally objectionable to save a family member in this situation. But if we want to give more weight to certain people then obviously it cannot be just the consequences that matter.

In contrast the deontologist can say that we have a *duty* towards family members and therefore we ought to save our father in this case because the act of 'saving our father' is the right thing to do. However, even though there are benefits to be gained from deontology being blind to consequences, it has its own highly counterintuitive consequences. Consider an example:

Thought experiment

> Julie has suffered domestic abuse from her partner for five years. Her children have recently left home and she has decided that enough is enough. She has been told of a women's refuge locally and has ascertained there is a place for her and had assurances that she will be safe. So, she packs some things and, when her partner is out, she escapes.
>
> Julie's partner, however, tracks her phone and thus her location. He turns up at the door with a gun and demands to know whether she is inside.
>
> Would it be morally right to tell Julie's partner the truth that she is inside?

It would seem not. Arguably the right thing to do is to lie, to say that Julie isn't staying there, thus defusing a very serious situation which could escalate with devastating consequences.

However, simple deontologists are insensitive to these potential consequences, and for them 'telling the truth' is the right action, and so they would believe the right thing would be for the refuge to tell the truth – to tell the husband that Julie is indeed inside the house.

Comparing deontology to consequentialism, it seems that deontology has some things going for it. In particular, it allows us to give greater weight to friends and family in moral decision making and it doesn't allow the rightness and wrongness of an action to be hostage to consequences which are outside out control. However, as we have just seen, it leads to some highly counterintuitive conclusions.

Questions

1. What is a deontological moral theory?
2. If you were to adopt deontology, which actions might you class as right and wrong?
3. How did you arrive at the list in (2)?
4. Could telling the truth to the husband be the right thing to do?

5 How might a deontologist respond to the thought experiment?
6 How do you think a deontologist could respond to two morally correct actions which come into conflict? For example, imagine a case where we can *only* keep a promise *by* lying.
7 One way a deontologist might suggest is to claim that the right thing to do is not say anything. So *not lie*. What do you think of this response?
8 One supposed advantage of deontology is that it might give special weight to those actions which involve certain groups of people, e.g. family members. Is this a genuine advantage? What problems might this generate? (See Chapter 65.)
9 One answer a deontologist might give is that if you tell the truth you are not responsible for the things that happen to Julie? How plausible do you find this?
10 In the end aren't intuitions fickle and culturally dependent? So why should we develop a moral theory which respects them?

Reading

Freeman, S. 1994. 'Utilitarianism, Deontology, and the Priority of Right', *Philosophy & Public Affairs*, 23(4), 313–49.

Gaus, G. F. 2001. 'What Is Deontology? Part One: Orthodox Views', *The Journal of Value Inquiry*, 35(1), 27–42.

Isenberg, A. 1964. 'Deontology and the Ethics of Lying', *Philosophy and Phenomenological Research*, 24(4), 463–80.

Kant, I. 2005 [1785]. *The Moral Law: Groundwork of the Metaphysic of Morals*, Hove: Psychology Press.

---

69 Virtue theory

In Chapters 65 and 68 we considered moral theories which ask 'how ought we to behave?' The consequentialist says that we ought to fix what is good, and this will help us decide which actions are right: namely, those which maximize the amount of that good. The deontologist on the other hand says that we should decide which actions are right and wrong irrespective of consequences.

The other main theory in normative ethics claims that both of these approaches are wrong. What we should be thinking about is not *how* to work out which actions are right and wrong. We don't need a method or set of rules. Instead what we need to do is ask what *type of person* we should be. For, the thought continues, once this is sorted out we will – just as a matter of course – perform the right actions. The virtue theorist claims that the right 'type of person' is the virtuous one, and that we ought to strive to become virtuous. Some, most notably Elizabeth Anscombe (1958), have argued that we have lost sight of this focus and consequently been mistaken in our thinking about morality:

Thought experiment

> Imagine that you have been seriously ill in hospital. You are visited every week by two friends. They have been great; they have kept your spirits up and kept you entertained. A few months after you have been discharged you are having a coffee with them. In conversation one tells you that coming to visit was 'my duty'; the other shakes her head. 'I didn't visit because it was my duty! I visited because I worked out that if I visited, it would bring about more happiness than if I didn't.' You smile and pay the bill. However, you feel thoroughly deflated by their explanations. Why might you feel like this?

The virtue theorist would answer this in the following way. What matters in evaluating actions is the type of people we are. In our thought experiment we want our friends to visit us *because* they are our friends. The motivation feels like it should come from 'within' them, rather than through some sense of 'duty' or 'calculation'. It seems somehow less morally praiseworthy for them to visit for the reasons that they did.

Aristotle thought along these sorts of lines. For him, as we *develop* as moral agents we will probably ask 'what action ought I to do?' But this is only *on the way* to becoming a more fully developed moral agent. The developed moral agent, the virtuous agent, will act in the right way, at the right time, in the right proportion; in virtue of the type of agent she is. Consider an example to illustrate this: it seems better to be the *type of person* who never considers stealing, rather than the person who is tempted by stealing, works out it is wrong, and resists the temptation.

This 'virtue theory' approach seems to get something right about whom we think praiseworthy. We want our moral agents to be those whose praiseworthy actions naturally 'flow' from them. However, there are a number of things which might make us wonder about accepting virtue theory:

The first, and perhaps most obvious, is that it seems like we want a moral theory to tell us what to do! Imagine I'm trying to decide whether to go to war. Being told that I need to be a virtuous person doesn't really help!

This leads to the second point. If this approach is going to help at all, we will want to know what a virtuous person is like. If we are told to be a good sprinter then we've got some idea of what that means. It isn't really debatable what a good sprinter is. However, being told to develop the *right* character seems to already rely on some moral theory to judge *what is virtuous*: for instance, if I ask a Christian which virtues to develop she might say 'humbleness'; whereas if I ask someone else she might say I should try to think more highly of myself and my achievements. How then will I decide?

Finally, it seems that the development of the 'right character' will be, to some extent, outside my control. Presumably I will only be able to do this by watching the right people (whoever they are) and getting their guidance in my actions as I practise to develop my character. But this depends on me being in contact with those people, so they can guide and correct me. Without guidance, it seems

I'll be lost. But getting the correct guidance looks like it is going to be down to chance; being morally praiseworthy might then depend on things like class, socio-economic upbringing, etc. But surely this is not what we normally think when considering our moral thinking. We think that people – irrespective of location, class, economics, etc. – are capable of being moral.

Questions

1. In the thought experiment what would you think of your friends' responses?
2. Is there any instance when we think it is better for someone to respond 'well that was my duty,' rather than 'that is just the type of person I am'?
3. What do you think someone's *character* is?
4. This theory is often talked about as a 'virtue' theory. What do you think 'virtue' is? How does your answer differ – if at all – from your answer to (3)?
5. What might the benefits of a moral theory be which first and foremost asks 'what type of person ought we to be?'
6. If you think that the right moral theory is one which asks 'what type of person ought we to be' how might you answer that question?
7. Imagine if the answer to your question 'what sort of person should I be?' was 'be selfish, crush others and hide any weaknesses.' What might we say in response?
8. Virtue theory isn't a decision theory. It doesn't tell us how we ought to behave in a particular situation. Is this a problem for it as a moral theory?
9. If someone wanted to become more virtuous, how might they do this?
10. Does your answer to (9) rule some people out from becoming virtuous? If so, does that matter?

Reading

Anscombe, G. E. M. 1958. 'Modern Moral Philosophy', *Philosophy*, 33(124), 1–19.

Hursthouse, R. 1991. 'Virtue Theory and Abortion', *Philosophy & Public Affairs*, 20(3), 223–46.

Rowe, C. J. and Broadie, S. (eds) (2002). *Aristotle: Nicomachean Ethics; Translation, Introduction, Commentary*, Oxford: Oxford University Press.

Stohr, K. 2006. 'Contemporary Virtue Ethics', *Philosophy Compass*, 1(1), 22–7.

---

70 The Doctrine of Double Effect

Most people think that murdering someone is worse than killing them in self-defence. This is because we think that intentions are somehow important in working out whether an action is right or wrong. Murder is worse than self-defence

because with murder the intention is to kill and not to defend oneself; whereas in self-defence it is to defend oneself and not kill.

This thought process has come to be known as the Doctrine of Double Effect (DDE) and is used in many cases: for example euthanasia and abortion where the woman's life is in danger. Before making DDE clearer, consider a thought experiment to illustrate the idea:

Thought experiment

> You work for a counter-terrorist organization, and have secured intelligence on a terrorist bomb plot. The bombers' plan is to further their cause by blowing up part of the local university, which would result in the death of many students.
>
> Your reconnaissance has ascertained where the terrorists are making the explosives. If you blow up this factory then they won't be able to make bombs. You will be able to stop them, and the death-toll and the suffering will be prevented. Unfortunately the factory is built inside the chemical engineering department of the university. Although this is a safe distance away from the part of the university that the terrorists intend to attack, you know that it is inevitable that if you destroy the explosives factory, then despite your intentions many students nearby are sure to die.
>
> Do you think that if you decide to carry out your plan then you have done something as morally wrong as what the terrorists were planning?

Often in these types of cases people will say 'no'. Even if the consequences are the same – the death of students – the terrorists' actions are worse than yours. It is worth making some clarifications:

- First, the DDE is not a pass for people to commit any acts they want by simply claiming that nothing bad was intended. What we aim to do must always be good, not bad.
- Second, it can't be the case that the good that is brought about is a consequence of the morally wrong. For example, the DDE isn't a licence to torture a terrorist for information which may help save lives.
- Third, the foreseen but unintended bad consequences cannot far outweigh the good that is intended. Go back to our thought experiment. If we could stop the bombers by dropping a nuclear bomb on the factory, even though the death of millions of people wasn't intended, it still would not be permissible according to DDE.
- Finally, all possible options to stop the bombers must have been considered and rejected. For example, in the thought experiment, you had to consider all the alternatives before opting for bombing the factory.

Philosophers have been perplexed by the DDE. After all, we might worry about what 'intention' means, and how we can understand 'unintended but known consequences.' Or questions like: 'if we know that our action will bring about a consequence, then aren't we blameworthy for the consequence?', etc.

However we respond it is undeniable that the DDE plays a large role, not just in philosophical discussions but in many areas of public life, e.g. the justification of military intervention; the law; discussion about abortion, etc. It is then worth thinking hard about what this doctrine is and whether it can be defended.

Questions

1 Do you think that if you decide to carry out your attack then you have done something as morally wrong as the terrorists?
2 What is an intention?
3 Explain the DDE.
4 Can one intend more than one end when performing a single action? For example, when I shoot an attacker can I intend both that he be killed and that my life be saved?
5 If you focus only on the consequences of an action what will you think about the DDE?
6 How might the DDE be applied to the area of euthanasia?
7 How might the DDE be applied to cases of abortion where the mother's life is at risk?
8 If we know that our actions will bring about two effects, aren't we responsible for both effects?
9 Explain why the DDE doesn't allow someone to perform a bad action in order to bring about a good result.
10 What is the 'acts and omissions' doctrine? How does it differ from the DDE?

Reading

Bennett, J. 1981. 'Morality and Consequences', in Sterling McMurrin (ed.), *The Tanner Lectures on Human Value*, Cambridge: Cambridge University Press, vol. II.

Boyle, J. M. 1980. 'Toward Understanding the Principle of Double Effect', *Ethics*, 90(4), 527–38.

Foot, P. 1967. 'The Problem of Abortion and the Doctrine of Double Effect', *Oxford Review*, 5(1), 5–15.

Mangan, J. 1949. 'An Historical Analysis of the Principle of Double Effect', *Theological Studies*, 10(1), 41–61.

Part VIII

METAETHICS

71 Internalism versus externalism

It seems that a key feature of moral judgement is that it motivates us to act in certain ways. Metaethicists find this interesting and have developed various theories which describe and explain this phenomenon. Consider a thought experiment to get at the heart of two opposing theories in moral psychology:

Thought experiment

> You are talking with a friend about eating meat and she replies that she has not yet decided whether it is wrong or right to eat meat and would like your advice. You spend some time with her, setting out the arguments against eating meat, and eventually she agrees with your point of view. The next day you are out and see your friend at a fast food restaurant and you join her. You reach the front of the line and order a veggie burger, but she orders a triple pork burger with extra chicken wings.
>
> You remind your friend of the discussion you had yesterday, and that the last time you checked, pork and chicken were meat! She smiles and responds. 'Oh, I still agree with you, eating meat is wrong; but what bearing has that on what I actually do?'
>
> How might you explain your friend's response?

We need to make some clarifications:

- First, we are interested in *motivation* and not *action*. It is true, of course, that motivation often leads to action, but this is not always the case. For example, we might be motivated to get fit, but have a stronger motivation to be lazy.
- Second, what is important is whether a moral judgement motivates, not whether a moral judgement is true. A judgement might be about the rightness of genocide or the rightness of giving money to charity, but that is not important for this debate. The question is 'what is the link between judgement and motivation?'
- Third, we'll assume that all moral judgements are sincere, and therefore we can rule out any suspicion in the thought experiment that the friend was joking or being ironic, sarcastic, etc.

There are two broad ways we might respond to this thought experiment. We might decide that even though our friend was being sincere, the fact that she was not motivated to give up meat shows that she did not really *understand* what she was saying. In metaethics, this response is given by *internalists* who think that there is a *necessary link* between moral judgement and motivation, such that if a moral judgement does not motivate then it is not a genuine moral judgement.

The other response is that our friend *does* understand moral terms, and that her moral judgement is genuine. It is just that when she was in the line queuing for food she ceased to care about what was right and wrong. This means that she was not motivated by the judgement that eating meat is wrong. This would be the type of response that the *externalists* make in metaethics. They think that the link between judgement and motivation – though a very common psychological

feature of humans – is not a necessary link. The link is *contingent* and depends on something external to the judgement – such as the desire to do what is right.

Questions

1 What is it to be motivated to do something?
2 Are you more attracted to internalism or externalism?
3 What is the link between a person's motivation and her action?
4 If someone says she is motivated to do x, but never ends up doing x, can we say that, despite what she thinks, she isn't really motivated to do x?
5 *Weak* internalists say that the link between judgement and motivation is necessary but *defeasible*. They claim that *if* an agent is psychologically normal – e.g. she is not depressed – then if she makes a moral judgement, that will necessarily motivate her. What do you think about weak internalism?
6 Would using empirical data – e.g., looking at how sociopaths behave, or scanning people's brains using an MRI scanner – be helpful in this debate?
7 Putting motivation to one side, what do you think the link is between moral judgements and *reasons* for action? For example, if I judge that it is right to give money to charity then will I necessarily have a reason to give money to charity? Or is the link to reason contingent?
8 Do you think there is any link between someone having a reason to do something and their being motivated to do that thing? If you do think there is a link, what do you think the link is? If you don't think there is a link, what is it to have a reason to do something?
9 Find out what cognitivists and non-cognitivists believe.
10 Why is this debate in moral psychology relevant to the debate between cognitivists and non-cognitivists?

Reading

Dreier, J. 2000. 'Dispositions and Fetishes: Externalist Models of Moral Motivation', *Philosophical and Phenomenological Research*, 61(3), 619–38.

Lenman, J. 1999. 'The Externalist and the Amoralist', *Philosophia*, 27(3), 441–57.

Sadler, B. J. 2003. 'The Possibility of Amoralism: A Defence Against Internalism', *Philosophy*, 78(1), 63–78.

Smith, M. 1994. *The Moral Problem*, Oxford: Blackwell, chapter 3.

---

72 Moral motivation

What is it to be motivated to do something? As discussed in Chapter 71, there is a connection between judgement and motivation, but describing the nature of this connection does not answer the question regarding what it is to be motivated.

One highly influential view claims that, to be motivated to do something, we need two distinct mental states: beliefs and desires. This seems an intuitive idea. I am motivated to check the football score if I *believe* there is a football game on and if I *desire* to know the score. If, on the other hand, I don't like football, then my belief that there is a football game on won't motivate me to check the score. Conversely, if I really desire to know the football results but don't believe the game has started, I won't be motivated to check the scores.

Holders of this view, which has come to be known as 'the Humean view of motivation', claim to know 'a priori' that in all cases of motivation there is a belief and an appropriately related desire, *and* that beliefs and desires are distinct mental states. However, we might doubt whether this could be true for all motivation, especially when it comes to moral motivation:

Thought experiment

> Imagine that a father promises to play football in the garden with his daughter. However, as it gets closer to going out he ceases to feel like it. He is tired, it is cold, his favourite programme is on TV, he's just eaten a big lunch, etc. In fact, there is nothing he wants to do less than kick a ball around in the garden.
>
> However, he remembers his promise and also believes that it is right to keep promises. So after much huffing and puffing he goes outside.
>
> What motivated the father to go outside?

Recall that the Humean thinks that in all cases of motivation there is not only a belief and an appropriately related desire but that the belief *cannot entail* a desire, as they are distinct states. What this means is that the anti-Humeans can refer to desires when it comes to talking about motivation. For example, they might say that although a desire is present when someone is motivated, it is not actually required for motivation.

The anti-Humean would say that the thought experiment illustrates the inadequacies of the Humean account as it seems natural to explain the above example like this: the reason that the father was motivated was that he *believed* that keeping the promise was right. That is, despite the very strong desire to sit in front of the TV and keep warm he was motivated to play football with his daughter. The anti-Humean could continue that moral belief is special in that simply believing that something is right/wrong/good/bad is enough to motivate.

The Humeans will be unimpressed, for they can always explain cases like this in terms of belief *and* desire. Perhaps, for example, the father has a general desire to do whatever is right. So even though he might say that he didn't want to go outside with his daughter, he has a general desire to do what is right, so there is a desire present.

Do moral beliefs motivate without the need for an appropriately related desire? Or is there always a need for there to be a desire to be present, even if the agent is not aware of one?

Questions

1 How do you think we should describe the thought experiment? Was the father motivated by belief alone?
2 What is it to have a desire for something?
3 Why think that it is only *moral* beliefs that motivate without desires? Could other beliefs? Can you give an example?
4 Some philosophers postulate the existence of other mental states which aren't beliefs or desires but are a belief/desire mix. What do you think about this sort of approach?
5 Beliefs and desires are sometimes distinguished by the 'direction of fit' metaphor. What is this metaphor? Does this metaphor help in advancing the debate between the Humean and the anti-Humean?
6 Read Chapter 71. How does that discussion relate to this one?
7 Do you think there could be 'a priori' knowledge regarding human psychology?
8 If the Humean and the anti-Humean can adequately explain any example that they are presented with then how might the debate between them progress?
9 Could the Humean ever be entitled to ascribe desires to the father if the father does not recognize a desire in himself?
10 Does it make sense to talk about human action without motivation?

Reading

McDowell, J. 1979. 'Virtue and Reason', *The Monist*, 62(3), 331–50.
Pettit, P. 1987. 'Humeans, Anti-Humeans, and Motivation', *Mind*, 96(384), 530–3.
Smith, M. 1987. 'The Humean Theory of Motivation', *Mind*, 96(381), 36–61.
Smith, M. 1988. 'On Humeans, Anti-Humeans, and Motivation: A Reply to Pettit', *Mind*, 97(388), 589–95.

---

73 Error-theory and beyond

Metaethicists are interested in whether our moral talk can be thought of as true or false. Some, most notably John Mackie (1977), argue that *all* moral claims are systematically and uniformly *false*:

Thought experiment

> After studying metaethics you have come to the following conclusion. When we make moral judgements we are describing the world as having certain properties. For example, when we claim 'kicking cats is wrong' we are describing the act of kicking cats as having the property of wrongness.

> You have also come to the conclusion that moral properties do not exist. This means that every time we make moral judgements – i.e. ascribe moral properties to the world – we are speaking falsely. After all, we have decided there are not any moral properties. Strictly speaking 'kicking cats is wrong' is false, 'giving to charity is right' is false, 'war is wrong' is false, etc., for *all* and *every* moral claim.
>
> Given your conclusion, what do you think people should do? Should they, for example, abandon moral talk altogether?

To clarify: just because we think that 'kicking cats is wrong' is false, it doesn't mean that 'kicking cats is right' is true. That also is a false statement because it is a moral claim, which to be true would need moral properties, but there are none. Second, we can use moral terms to create true sentences. For example, the claim 'killing is wrong is false' is itself true. Rather what this position is claiming is that if we *use* moral terms – for example, if we actually claim that 'killing is wrong' – then we are speaking falsely. The position crudely described in the thought experiment is the position of *error-theory*. There seem to be a few responses we could give in light of the truth of error-theory:

- *Abolitionism* claims that moral practice should go the same way as other false talk – such as fairy- and witch-talk – we should abandon it. If when we say 'killing is wrong' we are saying something false, then we should stop saying 'killing is wrong', 'giving to charity is right', etc. One problem with abolitionism is that in giving up morality we will be giving up some of its benefits; would, for example, the civil rights movement have been as successful without moral language?
- *Conservationism* claims that even if we know that moral talk is false we should simply carry on *believing* in morality. This would mean that we would not have to give up moral talk. However, it does seem rather an odd stand to take. We might even think that it is not possible to believe x as well as, and at the same time as, believing x is false. At best, conservationism would require some quite impressive mental gymnastics.
- We might instead try and forge a middle ground. With a *fictionalist* account we don't genuinely believe that, say, killing is wrong, or giving money to charity is right, rather we *make-believe* it. This allows us to keep morality whilst avoiding the oddity of believing something that we believe at the same time is false.

Putting conservationism to one side for now, why might we go for fictionalism rather than abolitionism? Well, surely it is better to keep morality so that we can, for instance, coordinate action, forge meaningful relationships, build trust and institutions, resist temptations, etc. Essentially keeping morality is useful despite it being false.

However, we need to be careful here. Recall that fictionalism is not conservationism. It says we should *make-believe* morality and not believe it. The question then arises whether the usefulness of morality can be retained if we talk in terms of make-believe. So, it may be true that we can get people to respect others, form meaningful relationships, coordinate action, etc. when we actually *believe* in morality, but would the same be true with *make-believing* morality? Is make-believe strong enough to secure the supposed pragmatic benefits of morality?

Questions

1 Do you think that moral claims can be true?
2 If you think the answer to (1) is 'yes', can you name a few? If you think the answer is 'no', what sorts of claims do you think people would typically think are true?
3 How would you feel if all moral claims turned out to be false?
4 Do you think Mackie might be right and that all moral judgements are systematically and uniformly false?
5 What is the difference between believing in x and *make-believing* x?
6 What are the potential benefits of morality?
7 What are the potential problems with morality?
8 Why, if at all, do you think that truth is important in ethics?
9 Do you think it makes sense both to believe x, and believe not-x, at the same time?
10 Do you think we could be error-theorists in other areas, such as maths, aesthetics, religious language, etc?

Reading

Garner, R. 2010. 'Abolishing Morality', in Richard Joyce and Simon Kirchin (eds), *A World Without Values: Essays on John Mackie's Moral Error Theory*, Dordrecht: Springer Press.

Lenman, J. 2008. 'Against Moral Fictionalism', *Philosophical Books*, 49(1), 23–32.

Mackie, J. 1977. *Ethics: Inventing Right and Wrong*, Harmondsworth: Penguin, esp. part 1.

Olson, J. 2011. 'Getting Real About Moral Fictionalism', in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics*, Oxford: Oxford University Press, vol. VI, 181.

---

74 Divine Command Theory

Imagine that you believe in God and that you believe God is interested in how we live our lives. We might then have a view about how right and wrong are related to God's commands. Divine Command Theory (DCT) is a very common view

amongst theists. It is the view that respecting others, giving to charity and loving others are all thought to be right because *God* commands us to do these things. However, there is a problem, first put forward in Plato's *Euthyphro*, which challenges this DCT:

Thought experiment

> You are committed to the DCT, so when asked why it is morally right to keep your promises you answer 'because God commands us to keep our promises.' However, your friend challenges you with the following question: 'But why does God command us to keep our promises?' What do you say? What answer would you give to this question?

It seems that you have two possible answers, both of which lead to some unpalatable conclusions. Hence what follows is known as the *Euthyphro dilemma*.

The first answer you could give is that God commands us to keep our promises because keeping promises is right. But this just leaves unanswered the original question regarding what makes something right in the first place. Your friend could legitimately ask – 'OK, but what makes the things that God commands right?'

The second answer – the one most common amongst theists – is that keeping promises is right *because* God commands it. This makes God the explanation, but it now creates another problem. For if God commands what is right, then God could command that we kill children at 9 a.m. and not kill children at 10 p.m., commit adultery every other weekend, kill all left-handed people, etc. The problem is that what is morally right and wrong now looks completely arbitrary.

Notice that the discussion isn't about how we come to *know* what is right and wrong. This is a genuinely interesting question, but it isn't what is at issue. Furthermore, there is no reason why we have to accept the DCT if we believe in God. There are other moral theories which rely on God but aren't DCT, e.g. Natural Law Theory.

The dilemma illustrated above has troubled theists for years. Most try and accept the claim that something is right/wrong because God commands it. So we'll assume this is the way ahead and see what responses we can give.

One natural response people give is that God's nature is such that He just wouldn't command us to kill children, kill left-handed people, etc. However, this is more problematic than it first seems. After all, *which* part of God's nature would restrict Him? One might think that it is His *moral* nature. But if He himself dictates what is morally right and wrong, there could be no external limitations on him from His nature.

Another response would involve 'biting the bullet' and claiming that *if* God did /had commanded that killing children is right, then it would be. But as far as we know He hasn't, so it is not.

Finally, taking a suggestion from Joyce (2002) we might think that God's commands are *identical* to right and wrong. Perhaps they are identical in the same way that water and H_2O are? If this is true, then there may not be a Euthyphro dilemma at all, for we don't have a dilemma for the water–H_2O identity. We don't worry whether something is water *because* it is H_2O, or whether it is H_2O *because* it is water. These sorts of questions simply don't make sense.

Questions

1 If there is a God do you think that He would have anything to do with morality?
2 If we don't adopt the DCT then how might God be related to morality?
3 Why do you think that the theist is more likely to opt for 'x is right because God commands it' rather than 'God commands x because it is right'? Would she be right to do so?
4 Do you think there are any *necessary* claims when it comes to morality? For example, do you think it is necessarily true that killing children is wrong?
5 Do you think we could respond to the Euthyphro dilemma by citing God's nature?
6 Of the responses discussed, which do you think is the most reasonable?
7 Do you think questions about how we could come to *know* God's commands really are irrelevant to the debate?
8 What do you think of the 'biting the bullet' response?
9 Some identity claims don't allow us to ask 'because' questions: for example, 'water is H_2O'. Why? And can you use this answer when defending the DCT?
10 If someone claims that x is right because *society* believes it is right, can we develop a Euthyphro dilemma for this position?

Reading

Adams, R. M. 1979. 'Divine Command Metaethics Modified Again', *The Journal of Religious Ethics*, 7(1), 66–79.

Allen, R. E. (trans. and ed.) 1984. *The Dialogues of Plato*, vol. I: *Euthyphro; Apology; Crito; Meno; Gorgias; Menexenus*, New Haven, CT: Yale University Press.

Joyce, R. 2002. 'Theistic Ethics and the Euthyphro Dilemma', *Journal of Religious Ethics*, 30(1), 49–75.

Peppers-Bates, S. 2008. 'Divine Simplicity and Divine Command Ethics', *International Philosophical Quarterly*, 48(3), 361–9.

75 Convergence and truth in ethics

It seems that in many cases we take convergence in belief as a sign of truth. Imagine that there is an investigation into a car accident. If the twenty witnesses interviewed said that the driver was wearing a black cap then that seems to be really good evidence that the driver was indeed wearing a black cap. There may be other explanations for this, such as that it was just luck that they all agreed, or that all the witnesses saw a film about a black-capped car driver, or that there was mass collusion. But all things considered we take the best explanation to be that it is actually *true* that the driver was wearing a black cap.

Of course, convergence in belief doesn't *guarantee* truth. If it did, then it would have been true that the Earth is flat, that people can't run a mile faster than four minutes or that we couldn't land humans on the moon. But if there is persistent convergence amongst the majority of people then we take this as a sign of truth, and, more importantly, we believe that if anyone thinks that the beliefs are false, then the burden of proof lies with them.

If this is a plausible general principle, then we might look to various areas of life to gain an insight into truth. Consider the following example:

Thought experiment

> Imagine that we pick twenty key people representing twenty different cultures across the planet. We don't let these people confer with one another and we isolate them from research methods, such as the use of the Internet, and then we ask them to write down the ten most important *moral* rules. It seems reasonable to assume that among those twenty lists there will be quite a lot of convergence. Perhaps all would have 'we shouldn't kill the young' or 'we shouldn't cause pain for fun'. Of course, there will be differences, but let us assume for the sake of argument that there is a large amount of overlap.
>
> Given what we said above, do you think that such convergence in moral beliefs is a sign that those beliefs are indeed true?

It is worth clarifying a point here. We might think that the idea of convergence in this example is not very probable. However, we may be able to counter such a view in the following way: imagine two communities. One community allows old people to starve to death whilst the other does not. Although at first sight it may seem that they have different moral values, this isn't necessarily so. Here is why: imagine that although both communities think that life is of moral value, one community lives in a very harsh environment where food is limited and therefore decides that it is better to let the older people starve rather than the young. This shows that it doesn't obviously follow from the fact that people act in different ways that they have different moral values.

If there can be truth in moral belief then some think that this means there are genuine moral *facts*. The thought is that truth is somehow linked to what is *real*.

In our first example, if it is true that the driver was wearing a black cap then this means that as a *matter of fact* the driver was wearing a black cap. If this line of thought is correct, then not only might we have an argument for the idea that certain moral beliefs can be true, but also an argument for the existence of moral facts. That is, convergence in moral belief might support moral realism.

Questions

1 Can you think of areas in life where there is convergence in belief?
2 Do you think that, in general, convergence is a sign of truth?
3 Do you think that there would be convergence in the moral beliefs of the twenty different people from different cultures around the world?
4 If you think that answer to (3) is 'no' what is an alternative explanation?
5 If you think there would be convergence do you think that this is a sign that those moral beliefs are true?
6 What do you think of the point about people having the same values but acting in different ways due to different environments?
7 Imagine that we agree that moral beliefs can be true. Does that mean that moral facts have to exist?
8 In some areas we think that convergence is important but only amongst the right type of people, e.g. convergence on what constitutes a proton matters if the convergence is between scientists, but not amongst non-scientists. How might this relate, if at all, to the moral case?
9 Can we use this line of argument in other areas of philosophy? For example, convergence in aesthetic judgements?
10 In general what do you think of the philosophical approach of using the views of people as evidence when developing a theory?

Reading

Finlay, S. 2007. 'Four Faces of Moral Realism', *Philosophy Compass*, 2(6), 820–49.
Sayre-McCord, G. 1991. 'Being a Realist About Relativism (in Ethics)', *Philosophical Studies*, 61(1–2), 155–76.
Skorupski, J. 1985. 'Objectivity and Convergence', *Proceedings of the Aristotelian Society*, 86(1), 235–50.
Williams, B. 1995. 'Truth in Ethics', *Ratio*, 8(3), 227–36.

---

76 The Frege-Geach problem

The non-cognitivist believes that when we make moral judgements we express non-cognitive states – for example, emotions. So when people claim that 'torture

is wrong' they are not describing torture as having certain features but rather they are expressing a negative emotion towards torture. When someone says 'torture is wrong' this might amount to 'Boo! Torture!' However, there is a problem that has continually dogged non-cognitivism: the 'Frege-Geach problem':

Thought experiment

> Imagine you think that when people say that 'torture is wrong' they are really expressing an emotion towards torture; when your friend says 'torture is wrong' you think that what she is really saying is: 'Boo! Torture!'
>
> Now imagine that your friend asks you: 'is torture wrong?' and you respond 'yes, torture is wrong!' You believe, quite reasonably, that this response fully answers your friend's question.
>
> However, a metaethicist would be forced to say that you *haven't answered the question.* Why? Well the question '*is* torture wrong?' *isn't* an expression of an emotion towards torture. This is because even though the question contains the words 'wrong' and 'torture' it *is not an assertion*. The question your friend is asking *isn't actually making a moral claim*. Just as my question 'is the moon made of cheese?' isn't making any claim about the composition of the moon.
>
> This is in contrast to your answer: 'torture is wrong'. This *is* a moral claim and hence you *are* expressing an emotion. The problem is that as a non-cognitivist you think that the *meaning* of a moral claim is dependent on whether you are expressing an emotion or not. Given that the answer 'torture is wrong' *does* express an emotion and the question 'is torture wrong?' *doesn't* express an emotion, the question and answer have different meanings. But if they have different meanings, your response 'torture is wrong!' isn't a genuine answer to your friend's question; but, prima facie, that looks very odd.

The point about meaning varying between asserted and unasserted contexts and the counterintuitive results this generates is the Frege-Geach problem. This problem does not only apply to questions and answers but is a worry *mutatis mutandis* for *any* situation in which there is a *mix* of asserted (e.g. torture is wrong) and unasserted (e.g. 'is torture wrong?') contexts.

So, for example, it seems like the following *modus ponens* argument is invalid:

(1) *If* killing is wrong then getting your friend to kill is wrong;
(2) killing is wrong;
(3) *therefore*, getting your friend to kill is wrong.

For our non-cognitivist this is invalid because the word 'wrong' has different meanings across (1) and (2). This is because in (1) there is no emotion expressed, because there is no moral claim being asserted; whereas in (2) there is a clear assertion and an expression of emotion. But this is a very odd position indeed! Any position which shows a *modus ponens* argument is invalid seems to have a lot of explaining to do.

There are two very broad strategies which the non-cognitivist could take, depending on how much the non-cognitivist wants to respect how we think our everyday moral language operates:

- On the one hand the non-cognitivist might say that 'killing is wrong' has to be an answer to the question 'is killing wrong?', and the moral modus ponens just has to be valid. In this case the non-cognitivist will explain how non-cognitive expressions *can* operate so that we reach those conclusions. For example, perhaps there is a logic of emotions, which once articulated explains how validity works when working with morality?
- On the other hand the non-cognitivist might 'bite the bullet' and say that despite how things might appear 'on the surface' moral language doesn't allow us to say the things we want to say. So, for example, although the moral modus ponens above looks valid, it isn't. Or, despite appearances, 'killing is wrong' is not an answer to the question. If we take this line of thinking we are committing people to massive error. We will then need to justify such a claim and also explain how it is that people have got it wrong.

Questions

1 What is non-cognitivism?
2 Why might someone be a non-cognitivist?
3 Give some examples of asserted and unasserted contexts.
4 Why does it seem that the non-cognitivist is committed to the meaning of moral terms varying across asserted and unasserted contexts?
5 Put the Frege-Geach problem in your own words.
6 What do you think that the non-cognitivist should say in response to the Frege-Geach problem?
7 Why isn't there a Frege-Geach problem for cognitivists – i.e. for those who think that moral judgements are expressions of belief?
8 What do you think the non-cognitivist should say in response to the Frege-Geach problem?
9 What do you think of the response to the Frege-Geach problem that says that the way we think moral language operates is just wrong? So, for example, the fact that the moral *modus ponens* is not valid isn't a problem, because it is just a useful way of making plain what most people are unaware of.
10 Non-cognitivist accounts have become hugely complex in light of this problem. Do you think that this should count against non-cognitivism or not?

Reading

Marino, P. 2006. 'Expressivism, Logic, Consistency, and Moral Dilemmas', *Ethical Theory and Moral Practice*, 9(5), 517–33.

Schroeder, M. 2008. 'What Is the Frege-Geach Problem?', *Philosophy Compass*, 3(4), 703–20.

Skorupski, J. 2012. 'The Frege-Geach Objection to Expressivism: Still Unanswered', *Analysis*, 72(1), 9–18.
Zangwill, N. 1992. 'Moral Modus Ponens', *Ratio*, 5(2), 177–93.

77 Moral Twin Earth

In 'How to be a Moral Realist' Richard Boyd suggests that the best way to be a moral realist is to draw a very close parallel with scientific realism (see Chapter 41). In particular, he thinks that he can establish that moral properties are identical to sets of natural properties by *empirical* means, rather than reflecting on the meanings of moral terms.

For Boyd, moral terms are rigid designators – they pick out the same natural properties across all possible worlds. So if 'good' picks out a set of natural properties, then it does so at any world. Furthermore, Boyd claims that the meaning of moral terms is determined by what they refer to. So, if Boyd is correct, the *meaning* of moral terms doesn't vary across possible worlds.

How might we establish these claims? Putnam's famous Twin Earth thought experiment, see Chapter 36, is meant to establish this for *scientific* terms such as 'water'. However, Horgan and Timmons set out a very influential challenge to Boyd. Their basic line of argument is that whereas Twin Earth thought experiments work to support scientific realism, they do in fact, fail to support moral realism. Consider then their 'Moral Twin Earth' thought experiment:

Thought experiment

> Imagine that on Earth 'good' picks out a set of natural properties identified by a consequentialist theory. We say something is 'good' if it causes us to pursue it and praise it, and given our consequentialist commitment, we think that maximizing happiness is what causes us to pursue and praise. So, whenever we talk about 'good', we are in fact referring to the consequentialist property.
>
> Now consider Moral Twin Earth (MTE). On MTE everything is just like Earth. People are praised for doing what is good and people pursue what they judge to be good, etc. However, on MTE the property which causes the pursuit and the praise is signified by a duty-based ethic, which means that the property causing the pursuit and the praise is in accord with what people reason is their duty.
>
> So when someone on Earth says 'giving to charity is good' she is referring to the consequentialist property, whereas when someone on Twin Earth says 'giving to charity is good' she is referring to the deontological property.

Remember that for all intents and purposes there is no big difference between how the person on Earth lives and talks and how the person on Twin Earth lives and talks. The things that Eartheans call 'good' are very much the same things that the Twin Eartheans call good. For example, the killing of children would not be

people's duty and would fail to maximize consequences, so both the Earthean and the Twin Earthean would judge the killing of children to be wrong.

The question that Horgan and Timmons ask us to consider is this: imagine that an Earthean says that some action is 'good' whereas a Twin Earthean says it is 'not good'. Are they having a *genuine* disagreement, or are they in fact, simply talking at cross-purposes?

Horgan and Timmons suggest that we would think that there *was* a genuine disagreement between the Earthean and Twin Earthean. After all, as in our example, if both Earthean and Twin Earthean use 'good' to refer to what is praised and pursued, then both *mean* the same thing by 'good'. But if Boyd is right we would, on reflection, judge that they didn't disagree. This is because in the mouth of the Earthean 'good' would mean one thing, whereas in the mouth of the Twin Earthean it would mean something else. So it seems that Boyd's linking of meaning to underlying properties – be that to do with consequences or duty – fails.

Questions

1 Read Putnam's Twin Earth experiment (Chapter 36). Are you convinced by his original thought experiment?
2 Without becoming involved in the complexities of language and metaphysics . . . do you think that there are similarities between scientific practice and moral practice?
3 Do you think that the meaning of moral terms is known 'a priori', or 'a posteriori'?
4 What does it mean for a term to be a rigid designator? Give some examples.
5 Do you have the same intuitions as Horgan and Timmons regarding whether Eartheans and Moral Twin Eartheans are genuinely disagreeing?
6 Richard Boyd's theory allows for the possibility of moral disagreement *between Eartheans* who come from different cultures and communities. Why might this be useful for the realist?
7 Can you think of any other benefits of drawing a parallel between moral and scientific realism?
8 One way to respond to the MTE thought experiment is by being relativist about the meaning of 'good'. What do you think of such relativism?
9 What is your reaction to this very theoretical approach to ethics?
10 Given your answer to (9) what, if anything, do you think the value is of metaethics?

Reading

Boyd, R. N. 1988. 'How to Be a Moral Realist', in P. K. Moser and J. D. Trout (eds), *Contemporary Materialism: A Reader*, London: Routledge, 2002.

Copp, D. 2000. 'Milk, Honey, and the Good Life on Moral Twin Earth', *Synthese*, 124(1), 113–37.

Horgan, T. and Timmons, M. 1991. 'New Wave Moral Realism Meets Moral Twin Earth', *Journal of Philosophical Research*, 16(1), 447–65.
Sayre-McCord, G. 1997. '"Good" On Twin Earth', *Philosophical Issues*, 8(1), 267–92.

---

78 Moral realism and the real

Moral realists think that when we make moral judgements we are making claims which can be true or false; that what make them true are moral properties and that these moral properties are *real*. This then begs the question – what justifies the realist in claiming that moral properties are real? Some realists – the Cornell realists – answer this question by talking about the seemingly indispensable role they play in our explanations:

Thought experiment

> You and some friends are out for a walk and as you come round a corner, you see a group of young people laughing and shouting. As you get closer you can see they have a cat pinned to the ground and have just set fire to its fur. The smell of burning flesh and petrol is easily identifiable in the air. You are horrified and come to the immediate judgement that their actions are wrong.
>
> Why is it that you come to the conclusion that their actions are wrong?

This Harman-inspired (1977) example has been at the centre of a debate regarding whether moral properties are real. One very plausible way of answering the question of why we judged the act of setting fire to the cat to be wrong is that setting fire to cats has the *property* of wrongness. So, then, our everyday way of talking prima facie supports moral realism.

However, merely citing moral properties in an explanation will not lead us to the conclusion that moral properties are real. If this principle is correct then it would follow that the explanation we give to children about how presents appear at Christmas would mean that Father Christmas is real.

So, just citing moral properties in an explanation isn't enough. This is because there are other ways we might try to explain why we believe setting fire to the cat is wrong, *without* talking about moral properties. For instance, we might claim that we came to the view that setting fire to the cat was wrong because we love cats, we hate suffering, we could see the cat was in pain, we could hear the squeal from the cat, etc. So it is for these reasons, those concerning our psychology and the other non-moral facts of the situation, that we judged the action to be wrong; not for the reason that the act of setting fire to the cat had the property of wrongness.

Therefore, the realist needs to show that her explanation which cites moral properties for the reason why we believe that the action is wrong, is not just *an* explanation, but the *best* explanation. To put it in another way, we might think that

the realist is entitled to talk about the reality of moral properties if, and only if, moral properties figure ineliminably in the best explanation of experience.

The realists have some reasons to think that this is a viable approach. Just consider how science works. We typically think that protons, electrons, quarks, etc. are real precisely because they figure ineliminably in the best explanation of our experience. The question then is whether the fact that just because setting fire to cats has the property of wrongness, is it the best explanation for why we judged that setting fire to cats is wrong? If it is, then the realist might say that moral properties are real, just as the scientific realist is entitled to say that protons, neutrons, etc. are real.

Questions

1 Do you think that moral facts are real?
2 Consider these things and categorize them as real/unreal: the equator; red; time; society; pain; love; concepts; quarks.
3 In the thought experiment how would you explain why we came to judge that the action was wrong?
4 What is a causal explanation?
5 What is a non-causal explanation?
6 What does it mean for an explanation to be the 'best' explanation?
7 Some people talk about a 'counterfactual' test for whether something is explanatory relevant what does this mean?
8 If you ran a counterfactual test in the thought experiment above what does this show?
9 What do you think about the approach of comparing scientific realism to moral realism as outlined above?
10 Why do you think people care so much about whether something is real or not?

Reading

Harman, G. 1977. *The Nature of Morality*, Oxford: Oxford University Press, esp. chapters 1 and 2.

Korsgaard, C. 1996. *The Sources of Normativity*, Cambridge: Cambridge University Press.

Leiter, B. 2001. 'Moral Facts and Best Explanations', *Social Philosophy and Policy*, 18(2), 79–101.

Murdoch, I. 1970. *The Sovereignty of Good*, New York: Schocken Books.

---

79 Internal and external reasons

In everyday talk we say that people have reasons. We might have reason to lose weight, or recycle or find a new job. Philosophers are interested in what it might

mean to say that people have reasons to do certain things. This is important because we believe that, in some cases, if people don't do what they have reason to do, then they have done something wrong.

Internalists about reasons claim that people have a reason to do something if *and only* if they can, from their current set of beliefs and desires, come to be motivated to do that thing. For the internalist, we can say that Joan has a reason to study for her exam if, and only if, Joan *could be* motivated to study for her exam.

Conversely, *externalists* about reasons think that it makes sense to say that people could have a reason to do something even if they cannot be persuaded to do that thing. So, for the externalist, Joan could have a reason to study for her exam, even if there is no way to motivate her to do so from her current set of beliefs and desires.

Internalists think that such a view is incoherent. They ask, how can there be a reason for a person to act in a certain way if that person cannot be moved or persuaded to act in that way? Consider an example which highlights this issue:

Thought experiment

> Judy is shy. In fact, she is so shy that she never goes out of the house. She has no friends and has no interaction with other people. When the phone rings or the doorbell sounds she hides away. She hates the person she is. The mere thought of meeting people fills her with terror.
>
> However, *if* Judy could go out, there are many things that would enrich her life: friends, fresh air, exercise, etc. But she is unable to do this because of the type of person she is.
>
> Do you think that Judy has a reason to leave the house?

The internalist would, in such a case, claim that if it is really true that Judy cannot be persuaded to go outside, perhaps because her shyness is too acute, then it makes no sense to say she has a reason to go out. The externalist, on the other hand, would maintain that Judy does have a reason to go out even if – because of the beliefs, desires and mental attitudes she has – she cannot be persuaded to go out. After all, we can for example imagine that if Judy found herself outside she might be happier than she has ever been, she might make friends, have more opportunities in her life, etc.

This is the choice then: either internalism or externalism about reasons.

Questions

1. Name some things that you have reason to do.
2. What does it mean to say that you have a reason to do something?
3. Could you have a reason to act in a particular way, but not know that you do?
4. If you think the answer to (3) is 'yes', in what sense do you have that reason? If you think the answer is 'no', then how do you explain the fact that we talk as

if we do? For example, we might say 'if only I'd *known* that he was ill I would have . . . '.

5 What is the distinction between internalists and externalists about reasons?

6 Do you think that Judy has a reason to leave the house?

7 If you do, then do you think that we can, in principle, persuade her to leave the house?

8 If there is *nothing* we can do to persuade her to leave the house, then do you think she has a reason to leave the house?

9 If externalism about reasons is right, how might we explain reasons people have that aren't linked to them in the way that internalists think?

10 Why might the debate between the internalists and externalists about reasons be important?

Reading

McDowell, J. 1995. 'Might There Be External Reasons?', in J. E. J. Altham and R. Harrison (eds), *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, Cambridge: Cambridge University Press, 68–85.

Smith, M. 1995. 'Internal Reasons', *Philosophy and Phenomenological Research*, 55(1), 109–31.

Williams, B. 1979. 'Internal and External Reasons', reprinted in *Moral Luck*, Cambridge: Cambridge University Press, 1981, 101–13.

Wong, D. 2006. 'Moral Reasons: Internal and External', *Philosophy and Phenomenological Research*, 72(3), 536–58.

---

80 Moral progress and truth

What is it for something to 'progress'? Well, when we talk of 'making progress', it is often linked to talk of 'success'. So, for example, schools are deemed 'successful' 'because they are 'moving along', they are progressing, progress seems to bring with it some *standard* by which the thing can be judged.

By contrast, you would not class the fact that you now, as an adult, like strawberry ice cream as progress. Rather it is simply that at one point in your life you didn't like strawberry ice cream and at another point you did; contrast this with, say, someone 'progressing in their wine tasting course' where the novice wine taster can *progress* toward the advanced wine taster.

So the general idea seems to be that if we can talk of beliefs progressing then there must be a way of judging those beliefs *independently* of the beliefs themselves. There has to be something by which we can judge whether there has been progress. This raises some interesting metaethical issues. Consider an example:

Thought experiment

> It used to be common practice to employ small children – usually boys – to climb up chimneys in order to clean out the soot. The boys were paid next to nothing, and were often beaten into staying with their master. The boys regularly broke bones, scraped and bruised themselves and had respiratory problems. On some occasions boys became trapped in chimneys and died.
>
> This practice was debated, discussed and argued about until an agreement was finally reached which stated that the use of child chimney sweeps was cruel and morally wrong. As a society it was agreed that this practice should be outlawed. Not because it didn't make economic sense but for the reason that it was morally wrong.
>
> Would you categorize the move to abandon the use of chimney sweeps as *moral* progress?

We suspect – hope – that most people would claim that abandoning the use of child chimney sweeps *is* a good example of moral progress. We can look back at this 'bad old time' and feel satisfied that we are now more aware of things like equality, fairness and human rights. It is not hard to think of other examples such as universal suffrage, or outlawing of slavery, or the move to make female genital mutilation illegal, etc.

Whatever example you pick, it does seem natural to talk about moral progress. Of course, this is not to say that all the changes we make are about progress: some might be classed as regress. All that we are asking in this chapter is whether 'moral progress' makes sense. But if it does, then how are we to understand *moral* progress? We make sense of progress in maths, or weight loss or crime reduction because there are benchmarks independent of what people hope and want. So, if there is moral progress it seems like we will need to agree there are some standards – something we are getting better at – which are independent of how we think.

Moral realists theorize that if there are moral facts then they are independent from our moral judgements. For instance, if a society judges that killing a certain race of people is right, this doesn't decide the matter. For what makes it right or wrong is whether as a matter of fact it *is* right or wrong. Now if this is true then we have a neat way of explaining progress. Moral progress is simply being able to recognize what as a matter of fact is right and wrong. It seems then that if talk of moral progress is meaningful, then it would support moral realism.

Questions

1. Name some areas where we happily talk about progress.
2. Name some areas where talk of 'progress' seems nonsensical.
3. What makes the difference between the areas listed in (1) and (2)?
4. What is it for something to 'progress'?
5. Do you think that there is such thing as moral progress? Can you give an example?

6 Why might moral realists think that talk of moral progress supports their position?
7 If you think that moral progress makes sense, but you don't think there are moral facts, then how might you explain moral progress?
8 How could someone deny that there is moral progress? In particular, how might she explain the common intuition that there is moral progress?
9 Although progress doesn't necessitate an end point, do you think it makes sense to talk about an 'end point' in ethics?
10 How might this debate be helped by considering the debate about moral epistemology?

Reading

Jamieson, D. 2002. 'Is There Progress in Morality?', *Utilitas*, 14(3), 318–38.
Macklin, R. 1977. 'Moral Progress', *Ethics*, 87(4), 370–82.
Moody-Adams, M. 1999. 'The Idea of Moral Progress', *Metaphilosophy*, 30(3), 168–85.
Wilson, C. 2011. 'Moral Progress Without Moral Realism', *Philosophical Papers*, 39(1), 97–116.

Part IX

AESTHETICS

81 What is art?

If you visit an art gallery it isn't long before the question is asked: 'Yes, but is it art?' People have very entrenched views and so get very animated about this question and even the media periodically question why large sums of money are exchanged and awards are given to what many would claim is not art. Philosophers have been interested in this question, although more and more of them argue that it's a question that isn't worth trying to answer:

Thought experiment

> Your cat tips a paint tin over your rug and rolls around on it. You notice that it is quite pretty, and cut the rug into the right size and frame it. A visiting friend notices it and comments on how great it is – 'I love that sort of modern art.' Being a prankster at heart you decide to see how far you can take this. You carry your framed 'cat painted rug' to the local art gallery. They are impressed and display it. A couple of weeks later some art critics write about it in the local art magazine as having 'depth' and 'insight'. Things get even more surreal when you receive a phone call offering you ten thousand pounds for your work. You start thinking about this – perhaps you were too quick, perhaps it is art? Perhaps you were wrong to dismiss it as a 'prank'; after all it appeared in a gallery, people discussed it as art and someone is willing to pay a vast amount for it. But you still have a nagging feeling that it is not art. But if it isn't, what is missing?

We might think that a work is a work of art if it has the right sort of aesthetic *properties*. Perhaps our 'cat painting' is/isn't art because it has/doesn't have the property of being, say, beautiful. However, taking this approach is problematic. There are works which we think are art but which are ugly or horrifying – just consider the work of Francis Bacon for example. If we think of the variety of things we might judge to be art – paintings, sculptures, music, dance, etc., then we might think it is hard to find any common properties.

However, there are other problems. Let's say that 'beauty' is what we judge as the key aesthetic property. There are things which we judge as beautiful but we don't typically judge as art; for example, a sunset, people or an antique – and some even talk about maths as being beautiful.

Or, take for example the works of Duchamp or Andy Warhol. Their works of art are often indistinguishable from non-art works (consider, for example, Duchamp's 'readymades'). Given these sorts of worries a number of philosophers abandon this 'property' approach. Rather, they say that a work is a work of art if it bears the *right sort of relation* to the/an art world.

We might think that an object is an art work if, and only if, it is something upon which someone – acting on behalf of the art world – confers on it the status of

being a contender for appreciation. So our 'cat rug' *is* art given that art critics have conferred on it the status of being a candidate for appreciation. But this of course begs an important question. *Who* is it that can do the conferring? If for example, one critic loves our 'cat rug' but two do not, does that count? What if every member of the public – the non-critics – thinks it is great art but the critics disagree, then what?

Moreover, if we look back at the history of art there are things which we now consider as classic works of art, for example Duchamp's 'fountain', that at the time were not considered to be art by many critics. Of course, we might take all of these worries as reasons for thinking that a search for a necessary and sufficient definition of what makes something 'art' is a lost cause. Some have taken precisely this route, maintaining that 'art' is like Wittgenstein's notion of 'family resemblance'. He suggested that although it might be considered that all the members of one family obviously belong together, there is not one property, or set of properties, which is common to all members of the family, but rather there are enough overlapping features to class them together. So perhaps then we can apply this thought to 'art'. Namely, that there is some obvious group which we are happy to call 'art', but there is not one property, or set of properties, which is common to all members of that group.

Questions

1 Do you think the 'cat rug' is art?
2 Would you have a different opinion if a 10-year-old had painted it?
3 Do you think there is a property/set of properties that all art works share?
4 How important do you think the intention of the artist is?
5 How important do you think the art critic is?
6 If everyone who saw it judged that the 'cat rug' was art, is that sufficient for it to be art?
7 Do you think that cave paintings could be art?
8 Do you think that the latest mobile phone design could be art? If not, why not?
9 Do you think that we can give a definition of art?
10 Why do you think that people care so much about the question 'is it art?'?

Reading

Beardsley, M. 1982. *The Aesthetic Point of View*, Ithaca, NY: Cornell University Press.
Dickie, G. 1984. *The Art Circle*, New York: Haven.
Gaut, B. 2000. 'The Cluster Account of Art', in Noel Carroll (ed.), *Theories of Art Today*, Madison, WI: University of Wisconsin Press.
Walton, K. 2007. 'Aesthetics: What? Why? and Wherefore?', *Journal of Aesthetics and Art Criticism*, 65, 147–62.

82 Moral value and aesthetic value

We all make moral judgements: for example, that it is right to give money to charity, it is wrong to drown babies, peace is a good thing, etc. We all make aesthetic judgements: for example, that Michelangelo's *David* is beautiful. This then prompts the question, 'What, if anything, is the link between the aesthetic and the moral?'

Thought experiment

> In 1935 Leni Riefenstahl's film *Triumph of the Will* was released and it is now widely recognized to be a great piece of film making. The techniques and style were groundbreaking and the cinematography was ahead of its time. If you watch the film you might judge that it had aesthetic value, and indeed people do think that the aesthetic qualities are outstanding.
>
> Yet the film is a glorification of the Nazi war machine. It portrays Hitler as godlike, with his followers mesmerized and hanging on his every word. Clearly the glorification of all that the Nazis stood for is a moral flaw.

Should we take this moral flaw into consideration when we evaluate the aesthetic value of the film?

- First, note that the question isn't whether we should ban such films, as the morality of censorship is not what concerns us here.
- Second, the question isn't whether in destroying works of art we have done something morally right or wrong.
- Third, the question isn't whether works of art are good or bad in virtue of the *effects* they have. *Triumph of the Will* is a piece of propaganda, and presumably it roused people to be more devoted to the Nazi cause. However, this isn't important here. If it helps, we can imagine the film having been discovered last year, and that therefore no one has ever been influenced by its contents.
- Finally, the question isn't about whether Riefenstahl was a Nazi. She was, but the moral outlook of the artist isn't what is in question.

Berys Gaut (2005) and others think there are three possible responses to this type of thought experiment:

- First, *autonomism*, which sees moral flaws as having *no relation* to aesthetic value; so, for the autonomist there is no difference to the film's aesthetic value just because it is about the Nazis. This might be initially plausible in this case. After all, we held both the judgement that *The Triumph of the Will* is beautiful

and that it is morally flawed. However, to show that there is *no* relationship between the moral and aesthetic value it isn't sufficient to say that a piece of work has both types of value. Rather, it has to be demonstrated that the work would not have changed aesthetic value if it had a different moral value.

- Alternatively we might hold *immoralism*, namely that the moral flaws in a work give the work *greater* aesthetic value. We might cite, for example, *The Silence of the Lambs*. Part of the aesthetic value of this film is the portrayal of evil by Anthony Hopkins as Hannibal Lecter. However, this response has a problem. These types of example about portraying evil characters do not support claims about the aesthetic value of *the work as a whole*. For example, we might think that the film *The Silence of the Lambs* is about courage and overcoming diversity – presumably valuable moral qualities. So, even though certain characters are morally flawed, the work itself isn't.
- Finally, we might hold *moralism*, that's to say that the moral flaws in a work of art *decrease* the aesthetic value of the work. So, in *The Triumph of the Will* the fact that the film has a moral flaw means that it isn't as aesthetically valuable as it could have been. However, there are, as a matter of fact, truly great pieces of art which *are* morally flawed which, importantly, we can't conceive as having greater aesthetic value.

Questions

1. Do you prefer autonomism, immoralism or moralism? Explain your answer.
2. Can you think of any other works of art which are arguably morally flawed?
3. Does this debate presuppose a moral and/or aesthetic realism?
4. Does the moral outlook of the creator of the art work – e.g. Wagner's anti-Semitism – make a difference to its aesthetic value?
5. Could you remove the moral flaws of a piece of work whilst claiming it is the same piece of work? For example, imagine removing Hitler from *The Triumph of the Will*. If you can't, then how does this relate to immoralism?
6. Do you think that a piece of art with moderate aesthetic value can be improved by improving its moral worth?
7. Do you think mathematics could have aesthetic value?
8. Do you think that something being inaccurate detracts from its aesthetic value?
9. Do you think that some moral flaws are so great that there could never be any aesthetic value in the work, e.g. 'snuff' movies?
10. If works of art have moral flaws then are artists morally blameworthy for their work?

Reading

Anderson, J. C. and Dean, J. T. 1998. 'Moderate Autonomism', *The British Journal of Aesthetics*, 38(2), 150–66.

Eaton, A. W. 2012. 'Robust Immoralism', *The Journal of Aesthetics and Art Criticism*, 70(3), 281–92.

Gaut, B. and Lopes, D. (eds) 2005. *The Routledge Companion to Aesthetics*, New York: Routledge.
Jacobson, D. 1997. 'In Praise of Immoral Art', *Philosophical Topics*, 25(1), 155–99.

83 Forgery

The issue of art forgery is of legal, social, anthropological and psychological interest, but is it of aesthetic interest? Philosophers of art have been divided in their answers to this question as they have considered the wider issues relating to the nature of aesthetic value.

Thought experiment

> In the local gallery there is one of your favourite paintings by Picasso. It has been in the gallery for twenty years and has helped attract countless visitors. The curator and other art experts have written about the aesthetic qualities of the painting and how it fits in with Picasso's cubism. It has toured other galleries and been valued for insurance at many millions of pounds.
>
> Imagine then that a woman comes forward and can demonstrate that she in fact painted the 'Picasso'. It turns out that she is an exceptionally brilliant forger and this is an excellent forgery.
>
> On discovering that this 'Picasso' is a forgery, do you think it has less value?

- First, we will put aside questions about how much things cost. Clearly the painting is devalued financially and may even be deemed worthless. Yet this doesn't get to the heart of the issue, which is why people are not willing to pay the same for forgeries.
- Second, we must try not to focus on *copies* of works of art but rather on the artists who are creating new paintings *in the style of* other artists. So in our thought experiment, the 'Picasso' wasn't copied from an original Picasso but was painted in Picasso's style and passed off as a Picasso.

Most philosophers of art would think that this 'Picasso' has less value. The tough question is trying to explain why. One natural response is to say that it has less *aesthetic* value; but what does this mean? After all, in our thought experiment, people, including art experts, have written about the painting's wonderful aesthetic qualities and many thousands of people have enjoyed looking at it. We can imagine the experts talking about the brush work, the colours, etc., yet presumably none of this will cease to be true when we discover it is a forgery. It isn't as if the painting suddenly fades when it is revealed as a forgery.

Of course, we could simply respond that the art experts weren't that expert after all and that they were mistaken; perhaps they were snobs and simply judged the painting of great worth because they saw the name 'Picasso'. However, we can

get over this by saying that in our example every expert there is, and ever will be, judges it to have great aesthetic value.

So what has happened? How can the features of the painting remain the same but the aesthetic value change? Of course, we might try and debunk our intuitions and claim they are unreliable; this would mean that we could claim that the forgery does indeed have *the same aesthetic value*. However, perhaps this should be only a last choice, given that the majority of people think differently. If we ditch this requirement then other – non-observable – things might count as being important for aesthetic value. But what?

One answer is the *intention of the artist*. Thus, because the forger *intended to deceive* then her work has less aesthetic value. However, presumably in our example we could say that the painting was stolen from the artist's studio where she was simply training herself in art and had never intended it to be shown in public. Or we might think that the reason that the forgery is of less aesthetic value is that it is less *authentic*. But what does 'authentic' mean? Arguably the forger is certainly being *true to herself* in our example: this is what she is good at, she wants to do it and identifies herself as a forger.

One influential response put forward by Dutton (1979) suggests that we should think of the painting as a kind of *performance*. If a painting is a performance then this, Dutton argues, involves grasping what sort of achievement the work represents. He thinks that forgeries have less aesthetic value because they 'misrepresent artistic achievement'.

Questions

1 Do you think that the forged 'Picasso' painting is of less *aesthetic* value? Explain your answer.
2 What does 'original' mean?
3 Does the notion of forgery make sense with all art forms? For example, what about forgery in the performing arts?
4 What is it for an artist to be 'authentic'?
5 Could knowledge of a work as being a forgery change the way you *observe* the work?
6 Can there be unobservable features of works of art which are important to their aesthetic value?
7 Does being a copy – rather than a forgery – make any difference to the aesthetic value? That is, if it was found that the woman in our example had an original Picasso in her studio and had copied that, would that change your thoughts?
8 Imagine in our thought experiment that it was found that she had programmed a computer to produce the painting and hadn't painted it herself. Would that change the aesthetic value?

9 If a conceptual artist forged works *as part of her art* then how does that relate to aesthetic value?
10 How does uniqueness relate to aesthetic value?

Reading

Dutton, D. 1979. 'Artistic Crimes: The Problem of Forgery in the Arts', *The British Journal of Aesthetics*, 19(4), 302–14.
Dutton, D. (ed.) 1983. *The Forger's Art: Forgery and the Philosophy of Art*, Berkeley: University of California Press, 131.
Goodman, N. 1968. *Languages of Art*, Indianapolis, IN: Bobbs-Merrill, esp. chapter 3.
Hick, D. H. 2010. 'Forgery and Appropriation in Art', *Philosophy Compass*, 5(12), 1047–56.
Lessing, A. 1965. 'What Is Wrong with a Forgery?', *Journal of Aesthetics and Art Criticism*, 23(4), 461–71.

---

84 What is music?

It is clear that music is valued in society. Millions, if not billions, of pounds are spent on the purchase of music, musical instruments, tickets to gigs, etc. People get married because of music; film producers spend vast amounts of time picking the right music tracks; software designers commission particular artists to compose music for their computer games; people turn to music in times of grief, etc.

The philosophy of music seems to generate some questions which are peculiar to it, as distinct from more general aesthetic questions. For example, we might ask what is the ontological status of music? What is a musical performance? What is the link between music and emotions? What is musical improvisation?

What we are interested in here is a much more basic question. Namely, if music is so important, do we have a sense of what music is? That is, is there a clear way of demarcating music from non-music. As it turns out this is harder to answer than we might first think:

Thought experiment

> You are irritated with your friend: 'but it's not music!' You are discussing his passion for electronically generated music. His 'Gabba Techno' involves a distorted drum beat running at 220 bpm, and various other noises. To you, it sounds more like the pneumatic drill used by nearby workers than 'music'. There are no lyrics, no melody to speak of. To be honest it sounds a bit like your car alarm and that, you are sure, isn't music.
>
> You think to yourself, 'Mozart's work, now that is music; so also is Mahler, the Beatles and many more.' But now your friend puts you on the spot: 'OK, what is

> music? What makes you so sure that you are right in saying that Gabba – or come to think of it your car alarm – isn't music but the work of Mozart is?'

Of course, you don't need to be a fan of classical music or have a hatred of Gabba to empathize with this sort of debate. Some of the enlivened discussions we have whilst growing up are with people who are convinced that their taste in music is correct and that what we listen to isn't really music at all. Yet it is very difficult to try and decide what, in fact, 'music' really is.

Perhaps one way to go is to claim that something is music if it includes *vocals*. But this seems overly restrictive. True, Gregorian chant, opera and the songs of the Rolling Stones would normally be classed as music, but so would many classical and jazz pieces that have no vocal element. We wouldn't, for example, say that Mozart's 'Jupiter' symphony isn't music because it doesn't include voices.

Maybe then something is music if there is variance in the notes that are generated. This seems to capture the feeling about why a single note played continuously might not be music. However, it doesn't seem sufficient. For police sirens aren't normally classed as music but they *are* more than one note; and we can think of a number of different examples of variance in notes which aren't music: car alarms, someone drilling a wall, squeaky wheels on a bike, etc. But in fact the type and variance of notes might be the wrong way of proceeding anyway. For it is often thought that John Cage's '4′ 33″' is music. And Cage's '4′ 33″' involves a musician sitting in silence at a piano for 4 minutes and 33 seconds. If this is music, then it is music which contains *no* notes at all!

But putting '4′ 33″' to one side, we might think that it isn't just about a variance in notes, but is about some form of *melody*. It isn't enough to have more than one note, but rather for something to be music there needs to be a *tune*. However, there are apparent counterexamples to this. Consider for example, 'You Suffer' by the grindcore band Napalm Death which has a running time of 1.316 seconds. If we are happy to call this 'music' then music can't be about a melody, for 'You Suffer' doesn't have one.

Perhaps then music is about *how* it is generated rather than what it does (or doesn't) contain. So, the thought would run, if the noise was *created with the intention* of it being classed as music, then it is. This would capture the Napalm Death and John Cage examples as music. However, there are other examples which look like they weren't created to be classed as music at all. For example, consider Yoko Ono's 'Toilet Piece/Unknown' which is the sound of a toilet flushing. The person who designed and built the toilet wasn't intending it to be taken as music. More worrying is that this definition of music – i.e. that it depends on the intention of the artist – seems too liberal. For on this account *anything* could count as music. We can of course dig in our heels and claim that '4′ 33″', 'Toilet Piece/Unknown', 'You Suffer' and others aren't music. But in doing so it seems that we need to have at least some sort of answer to a potential challenge to us that the music *we* like isn't really music.

Questions

1 Listen to Cage's '4′ 33″'; Ono's 'Toilet Piece/Unknown' and Napalm Death's 'You Suffer'. Do you think that they are music?
2 Do you think that bird song counts as music?
3 Imagine a famous composer 'composes' a piece which is one note generated by a keyboard lasting for an hour, could this be classed as 'music'?
4 If a chimp hit a xylophone and accidently played 'Happy Birthday to You', would this be music?
5 Do you think the spoken word can be classed as 'music'? Imagine, for example, someone with a strong regional accent reading out parts of a poem?
6 Can something be both music and art?
7 Do you think we can give necessary and sufficient conditions for what makes something 'music'?
8 Why do you think it matters to people whether something is 'music'?
9 Do you think people are right to worry about whether something is 'music'?
10 Are there any wider philosophical implications to this discussion?

Reading

Alperson, P. (ed.) 2010. *What Is Music? An Introduction to the Philosophy of Music*, University Park: Penn State University Press.
Amevans, M. 1967. 'What Is Music?', *The Journal of Aesthetics and Art Criticism*, 26(2), 241–9.
Kania, A. 2010. 'Silent Music', *The Journal of Aesthetics and Art Criticism*, 68(4), 343–53.
Varzi, A. 2013. 'Cover to Cover', *Current Musicology*, 95, 177–91.

---

85 The ontology of music

The question in this chapter is not what allows us to say that something is a piece of music – we dealt with that in Chapter 84. Rather we are interested in an ontological question. If a piece of music exists, what is it? What for example, is a symphony? Of course, there is no mystery if we ask about the nature of a particular *performance*. In this instance we can talk about the duration, the notes, the musical instruments, etc. But that is not the question. The question asked is 'what is the thing that they are performing?' What *is* Beethoven's Fifth? What is a symphony?

Thought experiment

> Andy and Tim have just watched the band Slayer play 'Reign in Blood' live. They start to discuss what 'Reign in Blood' *is*. What is confusing them is this. They have heard it many times in a variety of different venues, they have heard an acoustic

> version, they have heard different bands cover it, they have heard it given an 'artistic interpretation' on xylophone; they have heard it played in record quick time, and mind-numbingly slow time. They have heard it with words, instrumental, live (with a break in the middle for singing from the crowd) and on CD. Some were great, they agree, and some were less so.
>
> However, they think that they were all genuine performances of 'Reign in Blood'. Given the variety of different performances, it can't be true that the actual work 'Reign in Blood' is identical to any particular given performance. But what then is the work?

One suggestion might be that a performable work – at least in the case of music – is the collection of notes written down by the original composer. This certainly has some resonance with how we talk; we might say, for example, that we have just taken Beethoven's Fifth out of the library, where that means we have borrowed the musical score. This, though, can't be right. If we reflect on the last example, more correctly we might think that we haven't actually got *the* fifth symphony from the library, but rather an instance of it. Moreover, in some cases there might be no *score* – or at least no original score – at all. For example, traditional blues or folk songs which are simply passed down by being played and repeated.

We might think that performable works are identical not to an original score, but to a really good 'exemplar' performance. So, perhaps what the performable work 'Reign in Blood' *is*, is just the *best* performance of it by the band Slayer? But again, presumably we think it is possible for a piece of music to have *never been* performed. Just imagine that a new Beethoven symphony is discovered tomorrow, which has *never* been performed. It is still a symphony, but it has never been performed, there is no discernible 'exemplar' of the symphony at all. It seems then that it is a genuine question to ask what a piece of music is. What, for example, *is* Beethoven's Fifth, or Slayer's 'Reign in Blood'?

Questions

1. Give an example of a performance which you have recently seen.
2. What is a performable work?
3. What is the difference between performable works and performances?
4. Is there a difference between performable works in music and literature?
5. Is there a difference between performable works in music and film?
6. Is there a difference between performable works in music and drama?
7. When does a performance fail to be a performance of a performable work?
8. If a performable work is never performed do you think it is still a performable work?
9. Could it be possible for a performable work – perhaps a new symphony – to be written in such a way that it is not even possible to perform it – for example, it includes contradictory instructions?

10 How do you think this question about the nature of performable works is best answered?

Reading

Davies, D. 2010. 'Multiple Instances and Multiple "Instances"', *The British Journal of Aesthetics*, 50(4), 411–26.

Davies, S. 1991. 'The Ontology of Musical Works and the Authenticity of Their Performances', *Noûs*, 25(1), 21–41.

Dodd, J. 2004. 'Types, Continuants, and the Ontology of Music', *The British Journal of Aesthetics*, 44(4), 342–60.

Wolterstorff, N. 1975. 'Toward an Ontology of Art Works', *Noûs*, 9(2), 115–42.

---

86 Film and emotion

The philosophy of film is a recent and developing field, and there are many aspects of film which are philosophically interesting. However, one of the most discussed – and one which is applicable to other aesthetic areas such as literature – is our *emotional engagement* with film. It seems to be a trivial truth that people who watch films become emotionally engaged. People cry, laugh, are happy, sad, etc. In fact, one recent advertising technique is to show the audience's emotional reaction to films, to encourage others to share in such an emotion by watching that film.

But the question that interests us in this chapter is *why* is there such an emotional engagement? After all, if you ask a cinema audience if they *really* believed that – say – ghosts were going to attack them, or that aliens were attacking their hometown, they would say 'no'. So why do people have such emotional reactions if they don't really believe what they are seeing? On the one hand, when we watch a film we think what we are watching doesn't really exist; but on the other hand we certainly seem to care very much what happens to the people in these films. This seems strange:

Thought experiment

> Imagine that you watch a 'Rom-Com'. You leave the cinema very happy that the main character – Jonny – found out that he could be in a happy relationship despite his eccentric personality. You start to discuss this with a friend. Your friend asks you why the fate of Jonny matters to you so much that when his life is going well you are happy. As he points out, Jonny doesn't exist!
>
> After a bit of reflection, you suggest that you see some of yourself in the plight of Jonny, and in some way you see your own life's struggles being played out in this film. It is this which you think might be causing you to be happy.

> This initially seems OK, but then your friend points out that this was not the case in the film *Casper, the Friendly Ghost*, where you were extremely upset at the fate of Casper. So, he asks, in what sense did you identify yourself with a young ghost?
>
> On reflection, you think that maybe the emotional reaction is not because of identifying with the fictional character but because we have certain *beliefs* which themselves bring about emotions. That sort of explanation makes sense in our everyday experience. After all we think it is our belief that the bridge might collapse which generates the emotion of fear. However, there is a problem here. For how can we have reflective *beliefs* about characters who we know don't exist?
>
> What do you think? Why is it that we are so emotionally invested in the films we watch?

Philosophers of film have developed each of the thoughts above. For instance, perhaps we were too quick to dismiss the 'identification' theory and perhaps in some ways we *can* be said to identify with the fictional characters we are watching. Or perhaps we can develop the idea that it is the thoughts we have that cause emotions in us; or that we can have genuine beliefs about fictional characters?

Another popular response is 'simulation theory'. The idea here is that normal emotions have a certain feel and active element; so my fear of a tiger means that not only will I feel fear but I will actively run away. The simulation theory holds that when we experience fictional characters in film we can run a simulation of the emotion and detach the active part from the 'feeling' part. So we can feel scared, we can feel happy, etc., without running from the cinema.

However, it is not clear how we detach the action from the emotion. Moreover, we might be concerned that we can't really make sense of an emotion like fear if we remain completely unmoved by it. We might think that precisely what it means to be fearful is that we will run.

It is then an open question and a mystery why it is that we are so emotionally invested in films when it is quite possible that we don't believe that any of the characters we are watching are real.

Questions

1. Name the last fictional film that invoked an emotional reaction in you.
2. Describe that emotional reaction.
3. Reflecting on this, do you think this emotional reaction was genuine?
4. Can you explain why you reacted the way you did to the film?
5. Do you think that you identified with the character(s)?
6. Do you think that what caused the emotions were the thoughts that occurred to you when you were watching the film?
7. Do you think there is a difference between thought and belief?
8. Do you think we can have genuine beliefs about characters we believe are fictional?

9 Do you think that it makes sense to remove the 'active' part of emotions and still talk about the emotions? For example, do you think that you could be said to be in genuine fear without being compelled to do something?
10 Could you run this line of argument with literature? Are there any important philosophical differences between film and literature?

Reading

Friend, S. 2010. 'Getting Carried Away: Evaluating the Emotional Influence of Fiction Film', *Midwest Studies in Philosophy*, 34(1), 77–105.
Harold, J. 2010. 'Mixed Feelings: Conflicts in Emotional Responses to Film', *Midwest Studies in Philosophy*, 34(1), 280–94.
Hopkins, R. 2010. 'Moving Because Pictures? Illusion and the Emotional Power of Film', *Midwest Studies in Philosophy*, 34(1), 200–18.
Smith, M. 1995. *Engaging Characters: Fiction, Emotion, and the Cinema*, Oxford: Oxford University Press.

---

87 Aesthetic testimony

The topic of *aesthetic testimony* within the philosophy of art is complex, as it is interwoven with, amongst other things, sophisticated epistemological issues. However we can get a sense of the general concern. For very many things, we take the testimony of others as sufficient for knowledge. If you asked me how I know that there is no milk in the fridge, or how I know what colour your friend's new carpet is, then the answer 'because my brother told me' seems adequate. But in other areas such as aesthetics, the testimony of others doesn't seem to be sufficient grounds for knowledge:

Thought experiment

'It is a great film; it was subtly shot, the lighting was beautiful and the script was harrowing.'

'Oh, so you finally got around to seeing the film did you?'

'No, never seen it.'

'But how do you know all about these qualities in the film?'

'Well, my brother has seen it and he told me.'

'But in that case you can't really *know* that the film has these aesthetic qualities. You need to part with the cash and actually go and watch it.'

Do you think this is true? Can we know about the aesthetic qualities of the film without actually seeing it? Philosophers disagree as to the extent to which we can gain knowledge through aesthetic testimony. At one extreme the *pessimists* think that there is *no* chance of aesthetic knowledge through testimony; at the other extreme are the *optimists* who think that we *can* gain aesthetic knowledge through testimony.

It is important to notice that the debate between these two positions *isn't* about whether there is such thing as aesthetic knowledge: this is taken for granted, but the question is as to how it is that we can acquire such knowledge. Philosophers who write on this topic often hold a view somewhere between two extremes. They think that we ought not to believe that just *any* form of aesthetic testimony is sufficient for aesthetic knowledge, but it still may be true that *some* forms of testimony are sufficient. For instance, perhaps testimony about nature is one such type of testimony. Perhaps I could know that the view from your roof-top garden is beautiful because you told me?

There is often very little defence for pessimism, because it just seems so obviously true. The idea being that the best explanation for why people think that we can't gain aesthetic knowledge based on aesthetic testimony – as exemplified in our exchange above – is that as a matter of fact we can't. Some pessimists, though, think this is not enough, and try to defend more sophisticated and principled differences between aesthetic and ordinary uncontroversial forms of testimony. Perhaps there is a difference in how beliefs operate or how knowledge is transmitted?

A way to respond to the apparent obviousness of the pessimists' view is to explain *why* people – mistakenly, as it turns out – find it obvious. One thought is that we can't know whether something is beautiful via testimony because aesthetic judgements are *subjective*; that is, if *I think* a painting is beautiful then *it is*. If this is true then I can only know something is beautiful once I've made my own mind up on the matter. However, it is unlikely that you'd find a philosopher of art subscribing to such a radical subjectivism.

On the more optimistic side we might try and point to features of everyday practice which seem to suggest that we *can* gain aesthetic knowledge through testimony. Imagine that in our thought experiment above the person decides to go and see the film. If testimony is unjustified and doesn't give us knowledge, this action would be mysterious and surprising; whereas, if we *did* gain knowledge through testimony it wouldn't be. This can be generalized. We are moved to do various things based on aesthetic testimony – choosing a holiday destination or a book to read – and the best explanation for this is that aesthetic testimony gives us aesthetic knowledge.

Yet a pessimist can respond that this practice can be explained with an account which doesn't require *knowledge*. For example, in these sorts of cases it seems that all that is required is that we *trust*, or give *sufficient weight* to, aesthetic testimony.

And perhaps this is possible without going the whole way to granting that aesthetic testimony does give us *knowledge*.

Questions

1. What is testimony?
2. If you tell someone about the aesthetic qualities of a film, song or book that you have experienced but which she hasn't, do you think she now *knows* that those qualities exist in the film, song or book?
3. Do you think that there really is an asymmetry between everyday testimony and aesthetic testimony?
4. Consider your answer to (3). Are you an 'optimist' or a 'pessimist'?
5. Even if you think that most aesthetic testimony isn't sufficient for knowledge, do you think there are some special cases that are? For example, testimony about lost art works or testimony about nature.
6. Why do you think that the majority of philosophers of art would reject radical subjectivism as a response to the question of aesthetic testimony?
7. What do you think about the optimists' claim that the evidence that we *can* gain knowledge through aesthetic testimony is that people change their behaviour in response to aesthetic testimony?
8. What do you think of the approach that although we *can* gain aesthetic knowledge through reliable testimony, *most* aesthetic testimony is unreliable? For example, perhaps people don't want to admit that they hate the Mona Lisa or find Wagner deadly boring, and hence their testimony about these things is unreliable.
9. Another thought in this area is that we can only base our aesthetic judgements on *first-hand experience*. Explain why this isn't exactly the same as the issue of aesthetic testimony. Do you think this principle is right?
10. Have a think about other types of testimony: do you think there are parallel worries in other cases of testimony? For example, testimony about what is morally right and wrong, or religious testimony?

Reading

Goldman, A. H. 2006. 'The Experiential Account of Aesthetic Value', *The Journal of Aesthetics and Art Criticism*, 64(3), 333–42.

Hopkins, R. 2011. 'How to Be a Pessimist about Aesthetic Testimony', *The Journal of Philosophy*, 108(3), 138–57.

Meskin, A. 2004. 'Aesthetic Testimony: What Can We Learn from Others about Beauty and Art?', *Philosophy and Phenomenological Research*, 69(1), 65–91.

Robson, J. 2012. 'Aesthetic Testimony', *Philosophy Compass*, 7(1), 1–10.

88 Is photography art?

The art world is ever changing. It has always been the case that at any point in time various groups of people have been enraged by what they see being classed as 'art'. Just consider Duchamp's signed urinal, for example. However, as time passes the piece(s) at the heart of the controversy will become simply 'everyday' and part of the artistic canon.

That said, it does seem to matter greatly to people whether something is art (as we saw in Chapter 81). And this question is hotly debated when it relates to photographs. Some photographers class themselves as 'artists' and create artistic installations of their photographs. It matters to them, because the word 'art' seems to give their photos some 'gravitas' and kudos amongst the cultural elite. This then is what we'll consider in this chapter: namely, is the photographer entitled to say that photos can be classed as art?

Thought experiment

> You are wandering around the local art gallery and come to a room called 'Photos of War'. In this room there are horrific, poignant and very moving images taken from various conflicts around the world. You are in front of an image of a child soldier standing over a dead man.
>
> 'This is very moving and it certainly makes you think!' you say to your friend.
>
> 'Sure, but of course it isn't art.'
>
> 'Why not? It is moving me and it is making me think. That is what all those paintings in the last room did, and you class them as art don't you?'
>
> 'Yes', your friend continues 'but, if you think this is art then let's look at the shape that the boy's body makes against the sunset, or let's consider the placement of the gun and how the colour of the blood is changing as it is mingling with the tattered clothes . . . '
>
> 'Don't be disgusting, that isn't very respectful!'
>
> 'That is my point, we say precisely these sorts of things about art – in fact we did, when considering the paintings in the last room. We point out the colours, the structure of the shapes, etc. But here it doesn't seem right. This makes me think that photography isn't art.'
>
> Is your friend right? Is it right to think that the photos of war could be art?

The line of thought here then is as follows: maybe there are a number of differences between photography and traditional art forms, such that we are entitled to claim that photography isn't art. There does seem to be a difference between how we look at, and talk about, photographs and how we talk about painting. When looking at a painting we are expected to consider the features of the *art itself* but in contrast we don't consider the features of a photo itself. In fact, as in the case of the child soldier, it seems that such a focus seems in terribly bad taste.

There are other seeming differences as well. There seems to be an aspect of control in the creation of art works which doesn't seem present in the case of photographs. A photographer has very little control. In our example we can imagine that the photographer stumbled on the horrific scene and took the shot. Of course, she chose *where* to point her camera, but the control is limited. This is different from a painting which is, in its entirety, under the control of the painter. The painter can decide to do whatever she wants.

Finally, in photography we seem always drawn to, or made aware of, a *real thing*; the thing that is photographed. There really is a child soldier and we can't get away from that fact. In paintings there is not this connection to what is real. So, the thought runs, photographs aren't art. Some philosophers remain unimpressed by this line of argument and challenge this claim. The most obvious response is to deny that just because these features aren't present in photographs but are present in paintings it means that photographs aren't art. In fact we might take the difference as a way of individuating *types* of art rather than marking a boundary between art and non-art forms.

For example, there are features that seem unique to photographs, such as *inquisitiveness*. When faced with a photo of a Victorian slum, we might ask 'did they survive?' 'did they earn any money?' 'what age did they live to?' Yet when looking at a painting we don't have this sort of curiosity. Perhaps then, this feature, together with the other differences pointed out above, shows not that photographs aren't art, but that they are a particular and distinct form of art.

Questions

1 Do you think that photographs can be counted as art?
2 If you think the answer to (1) is 'yes' then why?
3 Do you think your answer to (2) is different to why you might count painting as 'art'?
4 What do you think about the claim that we do not aesthetically contemplate the features of the photo itself? For example, we don't say 'wow, look at the angle the boy soldier is making in relation to his gun.'
5 What do you think about the claim that photographers don't have the same control as painters, and that this might support the claim that photographs are not art?
6 Does the fact that photographs are connected to real things affect our thinking as to whether we consider photographs art?
7 If a photograph was (a) intentionally abstract or (b) partially manipulated by a computer programme, does this make you more inclined to say it is art? If so, why?
8 Imagine that a painter copied a photograph so that it wasn't immediately obvious that it was a photo. Could such a painting be classed as a piece of art?

9 Photos are connected to feelings of nostalgia in ways that other art isn't. What does this mean? Do you think it is correct?
10 In this chapter we have considered photographs and painting. Try thinking about how these issues might play out with other art forms, e.g., sculpture, literature, dance, film.

Reading

Costello, D. and Phillips, D. M. 2009. 'Automatism, Causality and Realism: Foundational Problems in the Philosophy of Photography', *Philosophy Compass*, 4(1), 1–21.
Freeland, C. 2007. 'Portraits in Painting and Photography', *Philosophical Studies*, 135(1), 95–109.
Maynard, P. 2007. 'Portraits as Displays', *Philosophical Studies*, 135(1), 111–21.
Scruton, R. 1981. 'Photography and Representation', *Critical Inquiry*, 7(3), 577–603.
Walton, K. L. 1984. 'Transparent Pictures: On the Nature of Photographic Realism', *Critical Inquiry*, 11(2), 246–77.

---

89 The paradox of tragedy

Tragedy and the value of its art have fascinated philosophers for as long as there has been tragedy. For example, Plato thought that it was corrupting (see for example, the *Republic*, books 2, 3 and 10) and anti-philosophical, whilst Aristotle (see for example, the *Poetics*) had a more sympathetic view. Eighteenth-century writers such as David Hume (*Of Tragedy*) and Edmund Burke (*A Philosophical Enquiry into the Origin of Our Ideas of the Sublime and Beautiful*; section 15) also wrote about the philosophical problems concerning tragedy. One of their central focuses was on what has been called the 'paradox' of tragedy.

Thought experiment

'I have just watched the most amazing production of *King Lear*!' your friend tells you. 'It was so very enjoyable that we have just bought another set of tickets.' Not knowing the play you ask for a general idea of what it is like: 'well basically it is a tragedy . . . you know, lots of suffering, betrayal, death, that sort of thing . . . '

'And you find pleasure in this misery! Isn't that, well, a bit weird! Surely if we witness someone being murdered, betrayed, tortured or whatever, I hope we wouldn't find *that* pleasurable. So how come you find representations of such things in tragedies so pleasurable?'

Your friend is looking less happy that he mentioned his weekend trip to the theatre. You continue, 'What is wrong with you – and all of us for that matter – that we take pleasure in watching such suffering?'

If your friend wants to try and respond to you, what should he say?

So in tragedy we seem to find pleasure proportional to the ability it has to accurately capture pain and suffering. For instance, we wouldn't think *Romeo and Juliet* was nearly as good a play if they had survived rather than taking their own lives. So, why is it that in one context – real life – we find the suffering of others awful and distressing, but in another context – tragedy – we find it pleasurable and valuable?

One of the things that is worrying here is not just the strangeness of our psychology, but that tragedy seems to force us to be *inconsistent* and *irrational* in our emotional responses, and this, we may think, is not just strange but *undesirable*; it seems that usually we strive in our lives to be consistent and rational in our responses and the decisions we make.

Moreover, there also seems to be a moral concern here as well. If someone took pleasure in the suffering of others we would naturally think that this was a *moral failing* on her part. However, we don't think that someone who enjoys *King Lear*, *Romeo and Juliet* or *Macbeth* has a *moral* failing. One way of dealing with the inconsistency is to *deny* that any inconsistency exists. After all, although people don't wish suffering and tragedy on others, they certainly do *seem* to take pleasure in watching it. Just think of the number of people who watch awful events unfold on news channels or watch people hurting themselves in YouTube videos.

Others, such as David Hume, think that we do in fact feel both pleasure and negative emotions such as terror, anxiety and sorrow when watching tragedy. But he thinks that they are in proportion to one another and that sometimes the portrayal of the suffering and misery is *too great*. This in turns means that the negative emotions won't be in balance with the pleasure we feel and it will, over all, reduce the aesthetic value of the play. Perhaps this rings true with our experience that there is normally some kind of 'tipping point' where audiences can't 'stomach' the suffering and harrowing nature of the tragedy, and cease to find it enjoyable, or, at least, people judge the work to be of less aesthetic value?

Questions

1. Do you like watching tragedy? Explain why.
2. How would you define tragedy? Can you give both a classic and a contemporary example?
3. Can you explain the apparent 'paradox of tragedy'?
4. Do you think that we find something *valuable* if we find pleasure in that thing?
5. Do people enjoy watching real life suffering and misery?
6. Do you think that there is anything *morally* wrong with taking pleasure in another's suffering?
7. Do you think that the representation of suffering can reach a point where it *is* wrong to take enjoyment in it?
8. Do you think that the representation of suffering can reach a point where it makes the tragedy of less aesthetic value?

9 Do you think that progress can be made in this debate through neuroscience and experimental psychology? If so, how might they help? If you don't, why not?
10 If you think there is a difference between contexts in which we enjoy watching suffering and contexts in which we find it painful and uncomfortable to watch someone suffer, then what are the reasons for this difference?

Reading

Currie, G. 2010. 'Tragedy', *Analysis*, 70(4), 632–8.
Friend, S. 2007. 'The Pleasures of Documentary Tragedy', *The British Journal of Aesthetics*, 47(2), 184–98.
Hume, D. 1826. 'Of Tragedy', in *The Philosophical Works of David Hume*, Edinburgh: Adam Black and William Tait.
Lamarque, P. 1995. 'Tragedy and Moral Value', *Australasian Journal of Philosophy*, 73(2), 239–49.

---

90 What is humour?

Humour plays a central role in our lives. Just take a quick glance through the current television listings, top films, best-selling books, etc. However, the philosophy of humour hasn't gained much attention until now, but this is starting to change. There are many philosophical questions we could ask about humour. What, for example, is it? What is the link between humour and laughter? Are there some subject matters that can never be humorous – for example, the gas chambers in Auschwitz? What is the relationship between gender and humour, etc? We will briefly focus on the first question: what is it for something to be humorous?

Thought experiment

> Your friend tells you a joke: 'Father Christmas, the tooth fairy, a child and a trustworthy lawyer all spot a ten pound note on the floor. Who gets it? . . . The child of course, all the others aren't real.'
>
> You find this funny. You aren't a lawyer but you know enough of them to think that the joke rings true. You tell the joke to a lawyer friend but she does not find it funny. In fact, she says 'it should have been "an interesting philosopher", not "a trustworthy lawyer"'.
>
> You wonder why it is that she doesn't find it funny. You think that perhaps something is only humorous if it gives one a chance to *feel superior* to others. She doesn't find the lawyer joke funny because she is a lawyer, you do find it funny because it makes you feel proud that you are in a profession that isn't based on twisting the truth.

Perhaps then something is humorous if it gives one a chance to feel pride at one's qualities by pointing out the flaws, limitations, or simply idiotic features of others? However, let us clarify that although it is an interesting question to ask about the link between morality and humour, this *isn't* what we are interested in here. If you read Chapter 82, there we discuss the link between aesthetic and moral value, and presumably we could run some of the same debates in the area of humour as well.

Furthermore, we are interested in humour rather than laughter. Presumably these aren't the same thing, as we can imagine cases where we might laugh but wouldn't say that something is humorous. Being tickled is a good example of this. So, we need to keep the issue of laughter distinct from the question of humour. First, let's see why the superiority view – the view exemplified by the thought experiment above – can't be right. Just look at what people find humorous; for example, what if your friend slips in the street, we might find this humorous; or consider Monty Python's *Holy Grail* which includes 'the knights that say "Ni"'. Slapstick and surreal comedy are often taken to be humorous, yet we don't experience any feelings of pride or superiority when finding them funny. So if humour isn't about feeling superior, what is it about?

Another theory is the idea that humour is a form of *relief*. Perhaps the reason that we found the joke above funny is because we were anticipating a punch line, there was a build up, and when it was delivered we could release mental energy. This seems particularly pertinent to 'knock knock' jokes, for example. However, such a theory seems at best to apply only to a very narrow set of things that we judge to be humorous; consider a pun given by Shaw (2010: 115): 'The man who fell into the upholstery machine is now fully recovered', or consider finding all the contents of your room set out exactly as they were in your room, but on the roof of your house.

What these last examples suggest is that perhaps something is humorous if it is *incongruous*. Going back to the Monty Python example, we have certain expectations of medieval knights, and for them to be going around saying 'Ni' and looking for a shrubbery isn't one of them. This incongruity view is neat because 'incongruity' can be applied to many different conventions: linguistic, social, religious, political, etc. For example, the reason people find 'dirty jokes' humorous is because they go against social conventions, and puns are humorous because they go against linguistic conventions. However, this of course leaves the question 'what is it for something to be incongruous'?

Questions

1 Tell your favourite joke. Why is it funny?
2 List three questions about humour – apart from 'what is humour?' – that philosophers might be interested in.
3 Give an example of something that is humorous because it makes us feel smug and reminds us of our own superiority?
4 Give an example of something which would support the idea that something is humorous because it is a release of mental energy.

5 Give an example of something which is humorous because it is incongruous.
6 Can you think of anything that is a counterexample to the idea that something is humorous because it is incongruous?
7 Can you think of times when we are presented with incongruous instances but we would not judge them as humorous? If so, what is the difference between these and the ones that are humorous?
8 Is the right approach to this issue to try and find necessary and sufficient conditions for something to be humorous?
9 What, if anything, do you think is the link between humour and laughter?
10 Could someone be wrong if she claimed that something is humorous?

Reading

Clark, M. 1970. 'Humour and Incongruity', *Philosophy*, 45(171), 20–32.

Gaut, B. N. 1998. 'Just Joking: The Ethics and Aesthetics of Humor', *Philosophy and Literature*, 22(1), 51–68.

Kivy, P. 2003. 'Jokes Are a Laughing Matter', *The Journal of Aesthetics and Art Criticism*, 61(1), 5–15.

Shaw, J. 2010. 'Philosophy of Humor', *Philosophy Compass*, 5(2), 112–26.

Part X

PHILOSOPHY OF POLITICS

91 Rawls and the veil of ignorance

A recent Oxfam report claims that half the global wealth is held by 1 per cent of the population. On the face of it this seems remarkably unfair. Why do we live in a world where some people are worried about the colour of their yacht whilst others are struggling to feed their family? This question has caused political philosophers to try to answer the question: 'what is the fairest way to distribute goods in society?'

One of the most recent and influential responses to this question is given by John Rawls. In his book *A Theory of Justice* he develops an account of justice based on the distribution of resources which, importantly, is not based on merit or consequences:

Thought experiment

> Imagine that you have recently been given a powerful position in the government, and it is your job to distribute goods across the population. You are trying to decide the fairest way to proceed. You are offered an 'ignorance' pill by your friend who claims that this will be the best way to distribute the goods fairly.
>
> This pill temporarily makes you unaware of your own particular qualities. You won't know whether you are rich or poor; you won't know about your race, gender, health or religion; whether you have a disability, your educational background, your IQ, etc. If you take the pill will this help you distribute the goods fairly?

Rawls's (2009) answer is 'yes'. Rawls writes that we need to be behind what he calls 'the veil of ignorance' in order to work out what would be the fairest way to distribute goods. We can see why other ways of distribution might not work. For instance, we might think no distribution should be random and arbitrary. This is what some people find so unfair about the rich inheriting vast wealth simply because of who their parents were. But neither should we select 'merit' as a way of deciding who gets what. Apart from the fact that the concept of 'merit' is vague, 'merit' is also distributed in an arbitrary way because some people just happen to be cleverer, faster, stronger than others.

Another way which seems unfair is to distribute goods so as to make the most people happy. If we took this line, we can imagine a case where a large number of people would be very happy if, say, disability benefits were stopped. We might worry that going for the happiness of the majority in this way would be legitimizing the tyranny of the majority.

So Rawls claims that we should think about goods distribution from behind this 'veil of ignorance'. From this position we would be forced to consider the distribution of resources in a way that it wouldn't unfairly favour us. Rawls thinks

that if we did this, then we would put ourselves in the shoes of the most adversely affected person – because that could be us. From this 'veil of ignorance' he derives two main principles:

- The first is the principle of *equal fundamental liberties*: freedom of speech, freedom of association, freedom of religious expression, etc. This principle of taking individual needs seriously means that he rejects the 'maximizing happiness' principle. It wouldn't be possible for the minority to have their freedoms curtailed even if it meant that the overall majority benefitted in some way.
- The second is what is called the *difference principle*. This states that resources should be distributed equally, *unless* unequal distribution would benefit everyone. Just to clarify this last point: for Rawls, fairness isn't about *strict equality*. A fair society for Rawls might be one in which some people earn much more than others; that is, a society where there is an unequal distribution of wealth. But he says this could be fair if, and only if, this society was the one that benefitted *everyone overall*. He thinks that these principles are evident from behind the veil of ignorance.

However, some disagree. Interestingly Rawls has attracted criticism from feminist philosophers (see Okin 2003) who worry, amongst other things, about the coherence of the original position.

Questions

1 'A recent Oxfam report claims that half the global wealth is held by 1 per cent of the population.' What do you think when you read this?
2 In the thought experiment above, would you take the ignorance pill?
3 Do you think that it is unfair to distribute goods based on merit: such as fitness, athletic ability or IQ?
4 Do you think that it is unfair to distribute goods to maximize happiness?
5 Do you think that equality is a good?
6 Do you think you could make a decision at all, if you removed all the personal facts about yourself, e.g. that you are religious, or a woman?
7 How would you feel if the grades for your next essay were decided from behind the veil of ignorance? If you think this would be wrong explain why.
8 How might you argue that the current huge wage differences are fair, using Rawls's 'difference principle'?
9 Does it really make sense to claim that we can consider things from a position of complete ignorance of our own particular circumstances?
10 What is wrong with rewarding people for things that are out of their control? (How might this relate to Chapter 64 on 'moral luck'?)

Reading

Cohen, G.A. 2000. *If You're an Egalitarian, How Come You're so Rich?*, Cambridge, MA: Harvard University Press.

Nozick, R. 1974. *Anarchy, State, and Utopia*, New York: Basic Books.

Okin, S. M. 2003. 'Justice and Gender: An Unfinished Debate', *Fordham Law Review*, 72, 1537.

Rawls, J. 2009. *A Theory of Justice*, Cambridge, MA: Harvard University Press.

---

92 Nozick and Wilt Chamberlain

There is a finite amount of goods in the world, but there should be enough for everyone to be comfortable. However, this isn't the world we find ourselves in. Some people have vast sums of money, some people live hand-to-mouth, going hungry or dying from poverty-related disease. We might legitimately think that any such pattern of distribution of wealth – where some people have millions and others have very little – is simply unjust.

Robert Nozick (1974) argues that this isn't necessarily true. Famously, he argues that we can't prejudge what a just society would have to look like in order for it to be just. He uses his famous 'Wilt Chamberlain' thought experiment to show why:

Thought experiment

> Pick any pattern of goods distribution in society which you think *is just*. You wouldn't go for something as simple as this but imagine you think the just pattern of distribution is to take all the goods and divide them up equally amongst everyone.
>
> Ok, imagine that you've done this. Everyone has the same amount of money, the same size house, the same income, etc. Call this pattern of distribution (P1).
>
> Starting from P1, now imagine that many people are basketball fans and want to see the great player Wilt Chamberlain play basketball. Wilt agrees to play for team Philosophy, but he wants a quarter of the ticket price. The people agree and think this is a good idea, and so when they come to watch team Philosophy play they put one quarter of the ticket price into a box and all the money in the box goes to Wilt. This way Wilt gets very rich and people are happy watching some great games of basketball.
>
> Now though, we have a *different* pattern of distribution of goods (P2). In P2, Wilt has loads of money and some people have less. So, if a just society is one which follows pattern P1, some injustice must have occurred – but what?

Robert Nozick argues that the final pattern P2 is *not unjust* because no injustice has occurred in arriving at that pattern. What the example shows is that we can't

prejudge what a just society would have to be like in order to be just. We started with the ideal pattern – P1. People *freely* gave their money away – it wasn't as if there were bouncers frisking them at the gate. No one was hurt, and no one's freedom seems to have been violated. People freely put one quarter of the ticket price in the box and suddenly there was a different pattern. Nozick thinks this line of argument can be applied to *any* theory that says that a just society has to follow a certain *pattern*.

This argument has been much discussed and is very influential. However, there are a number of things we might think need further defence and consideration. Here is one: Nozick thinks that — roughly speaking – as long as the process is fair, the consequences of that process will be fair. Furthermore, he thinks that not limiting people's freedoms is part of the process of being fair. This seems correct in the Wilt example where no freedoms have been violated by putting one quarter of the ticket price in the box. But we might worry that most transactions aren't like this. We can imagine a situation where our actions through free transactions directly affect the freedom of others. If we are very rich and fairly acquire considerable wealth and consequently pay the best surgeons to work for us, then this will of course directly affect the health care available to people from poorer families. The best surgeons won't be available to them.

The general point here is that it seems that transactions between people actually do have implications on others outside the transaction. If this is the case, we might think there is a need to limit such transactions, or at least to be aware that Nozick might have oversimplified the situation.

Questions

1 If you were to pick a 'pattern' of distribution of goods in society, what would it be?
2 Do you think that just distributions can be captured in a particular pattern?
3 What do you think about Nozick's 'Wilt Chamberlain' example?
4 Do you think that just distribution is whatever we arrive at if the processes in the system are just?
5 How might Nozick's argument relate to taxation?
6 Do you think it could ever be unjust for people to freely transfer their goods to another?
7 What do you think about the problem that free transfer might threaten the freedom of others, as in the example of private healthcare?
8 If Nozick is right, then how should a government operate with regard to people's lives?
9 When discussing these issues people often use the word 'entitlement'. In the example, Wilt is *entitled* to the vast sums of money. What do you think 'entitlement' means?
10 Is it right to think that respecting people's freedom is important?

Reading

Cohen, G. A. 1977. 'Robert Nozick and Wilt Chamberlain: How Patterns Preserve Liberty', *Erkenntnis*, 11(1), 5–23.

Fried, B. 1995. 'Wilt Chamberlain Revisited: Nozick's "Justice in Transfer" and the Problem of Market-Based Distribution', *Philosophy & Public Affairs*, 24(3), 226–45.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*, New York: Basic Books, esp. 160–4.

Wolff, J. 1991. *Robert Nozick: Property, Justice, and the Minimal State*, Stanford, CA: Stanford University Press, esp. 73–118.

---

93 Group and individual reasons

In chapters 13–14 of *Leviathan*, Thomas Hobbes gives a justification for why it is rational to agree upon a social contract. Hobbes thinks that a social contract is the, or a, way of ensuring that people in society are better off. He thinks this is the best way out of a dilemma concerning how to coordinate action. This is a bit abstract so let's consider an example:

Thought experiment

You are the mayor of a city – Hobbesville. Over the valley there is another city – Platoville. You are trying to decide whether to go to war with Platoville. Thinking this through, you recognize there are four outcomes.

- The first outcome is that you might attack Platoville but they remain peaceful. This is the *best outcome*, for then you can steal lots of great resources from Platoville, and the people of Hobbesville will be very happy and be pleased with you.
- The *second best* outcome would be for both cities to choose peace. Of course, Hobbesville would not acquire extra riches and spoils, as if it had taken the first option, but the citizens would not have to worry about attack. If both cities were enjoying peace, then there could be harmony, and people could put their energies into developing resources and having fun.
- The *third best* outcome is if Hobbesville attacked Platoville and Platoville responded in kind. This would mean that Hobbesville wouldn't be a sitting duck because they would already be prepared to defend themselves against attack. But on the other hand Hobbesville couldn't spend time on improving the lot of the inhabitants and having fun. This is because they would be spending their time fighting Platoville.
- The final outcome is the *worst*. This is when it is decided to pursue peace with Platoville, but Platoville chooses to attack you. If this was the outcome you would be sitting ducks!

So, as mayor of Hobbesville and *knowing nothing about what Platoville will do*, what do you think the most rational thing to do is? Go to war or go for peace?

Most people would suggest that for the mayor the *rational* decision is to go on the offensive. If he does this then he might get the first and best outcome – the one where Platoville is the sitting duck and where Hobbesville takes all its resources. But more importantly you, the mayor, wouldn't then get the worst outcome, the one where you go for peace and Platoville attacks you.

However, we now start to see why philosophers have been so interested in this, because this line of reasoning will be precisely the same line of reasoning that the mayor of Platoville would take. She will think that an offensive is the best option. So this means that if each individual mayor acts rationally, both will end up with the third best outcome – the one where they are both at war. But surely the best for both is where both are at peace – the second outcome.

It seems then that there is a *group reason* to be peaceful but an *individual reason* to go to war. But then, how do we get our second best outcome? How do we get out of this dilemma? Well the solution seems obvious. Why not wander over to Platoville and have a chat with the mayor about maintaining the peace? This though would be very risky for how do you know you can trust the other mayor? You might end up with the worst outcome if your trust is misplaced.

So to solve this dilemma something with more 'teeth' is needed. One solution may be to enter a contract with Platoville which would be an agreement not to go to war. Importantly, the punishment for breaking the contract needs to be very severe indeed, requiring a person – a sovereign – to monitor and enforce the contract, so that all elements of reliance on trust are removed. If both mayors sign this contract then they can be reasonably confident they won't arrive at the worst outcome – the one where one city is at peace but the other city is at war. We will escape from that dilemma and also from the third outcome.

Hopefully you can see how this structure leads to the following conclusion: it is best for each individual in a state to give up some of her freedoms (for example, the freedom to steal from her neighbour) because if there is a contract amongst everyone where some freedoms are sacrificed, and there is some policing of this through a sovereign, then everyone will get the best result. For Hobbes this is why it is rational to form a social contract and be governed by a sovereign who can police such a contract, even if it means that individuals do not always get the best outcome.

Questions

1 What would you do as mayor of Hobbesville?
2 Create your own version of the dilemma.
3 Put in your own words why Hobbes thinks the dilemma justifies a social contract.
4 What do you think of the idea of a 'social contract'?
5 Why might we need a 'sovereign' as well as the contract? What powers might a sovereign need in order to get us out of the dilemma?

6 Is there another solution to the dilemma?
7 What do you think would happen if we kept repeating the 'game'?
8 What account of human nature does this dilemma rely on? Is it correct?
9 How might this relate to our *moral* actions?
10 What if Platoville was full of weak and starving people and had no army? Would it then be rational for us to take advantage of them? How does this relate to your answer to (9)?

Reading

Gauthier, D. P. 1986. *Morals by Agreement*, Oxford: Oxford University Press.
Hampton, J. 1988. *Hobbes and the Social Contract Tradition*, Cambridge: Cambridge University Press.
Hobbes, T. 1968 [1651]. *Leviathan*, edited with an introduction by C. B. Macpherson, Harmondsworth: Penguin.
Woodard, C. 2003. 'Group-Based Reasons for Action', *Ethical Theory and Moral Practice*, 6(2), 215–29.

---

94 The state of nature

The state of nature is a key ingredient to many great political philosophers' works. For instance, Hobbes, Locke and Rousseau use the 'state of nature' as a way of starting to think about many issues. Although each thinker varies in the details, the general idea is the same: namely, we start our political philosophy from the state of nature. So what is this 'state of nature'?

Thought experiment

> Imagine that you and one hundred others are on a cruise around the Bermuda triangle. Unfortunately, the weather starts to worsen and the ship gets into trouble. Your ship starts taking in water and keels over. Luckily there are enough lifeboats and you all make it to a remote island. On this island there is no evidence of civilization. There is no shelter, no electricity, no phone coverage, etc.; all the evidence points to the conclusion that you are the first human beings ever to set foot on this island. What are you going to do?

There are numerous books and films based on this sort of thought experiment. Some are utopian, some are dystopian. We might think that we would all sit down together and chat about the best way to proceed. We might decide the jobs that need to be done? Perhaps we'd elect a leader, organize provisions, organize hunting groups, etc. Perhaps we would decide amongst ourselves that some people will be farmers, some will be look-outs, some will maintain the shelters, etc.

Furthermore, we might think that with this sort of mutual agreement regarding the organization and running of the island we could maintain a reasonably comfortable existence indefinitely. Sure, there might be a few people who didn't 'play ball', but these could be dealt with in a way agreed by the group.

In contrast we might think that it would be absolute chaos. Maybe at first people would be nice to one another, perhaps helping each other to build shelters, or sharing provisions. Maybe some would volunteer to fish for the others, collect firewood, etc. However, this would deteriorate quickly. Most people would have the same sorts of skills and be able to 'offer' the same sorts of things. There would then be no sense of everyone being able to contribute to the group. People would start stealing from one another, and exploiting people's good will and trust. The physically strong would rise up to crush others, and make sure they had the full stomachs and best and warmest shelter. Perhaps there would be splits, and vigilantes? It would be a nasty place to be.

Of course, what you think will happen on the island reveals a lot about what you think about human nature. Are people inherently selfish and out for themselves? Or are they altruistic and caring, etc? This 'island' is pre-political. There is no voting or government, no agreed laws, etc. The island is what political philosophers would call a 'state of nature'. It is a state (not to be confused with a governed 'State') where the trappings of modern living are peeled back to reveal the true nature of people. One question we might ask is whether it matters if there ever actually was such a state? Was there a time in the history of humanity where people lived in a pre-political state? Does it matter in terms of our political philosophy if we think not? Maybe all we need for such a state to be a useful tool is that it is a *possibility* that humans were once in such a state?

However you answer this, it is true that such a state *has* helped, and continues to help, political philosophers think about various issues. For example, what might be the best way out of such a state? Or what is the best way to organize people? How do we decide ownership of property, or how sanction and punishment would work?

Questions

1 What would you do if you were washed up on the island?
2 What would you do if you were washed up on the island and twenty people were very ill and needed 24/7 care?
3 Imagine that you were made leader, what would you do then?
4 What would be your biggest fear on the island? How might you deal with that fear?
5 Do you think that your answer to (1) tells you anything about human nature?

6 If you think the answer to (5) is 'no', what do you think it tells you? If you don't believe there is such thing as human nature, can you make any general claims about how humans behave?
7 Do you think that there ever was such a 'state of nature'?
8 If you think there never was a 'state of nature' then is there any use thinking about it?
9 Why do you think *political* philosophers talk about the state of nature?
10 What sorts of questions might the state of nature help political philosophers to answer?

Reading

Dodds, Graham G. and Shoemaker, David W. 2002. 'Why We Can't All Just Get Along: Human Variety and Game Theory in Hobbes's State of Nature', *Southern Journal of Philosophy*, 40(3), 345–74.
Hamilton, J. 2013. 'The Origins of Hobbes's State of Nature', *Hobbes Studies*, 26(2), 152–70.
Simmons, A. 1989. 'Locke's State of Nature', *Political Theory*, 17(3), 449–70.
Zwolinski, M. 2011. 'States of Nature', *Journal of Value Inquiry*, 45(1), 27–36.

---

95 Tacit consent

How can it be that in some instances someone *ought* to obey the State? Consider a case where someone gains nothing from the State, no benefits at all. Why might it be the case that that person ought to obey the state?

Thought experiment

Imagine you are talking to your politically radical friend who is living 'off the grid'. She doesn't have a passport, she supports herself by growing food, she self-medicates, educates her children, etc.

She is getting particularly animated because it turns out that the council have found her tent on the local common land and are trying to move her out. She isn't doing any harm to anyone, she is hidden away and no one knows she is there. In fact, it was only happenstance that she was discovered at all. She clears up after herself and isn't a nuisance. Why should she move?

In your chat she tells you that she hates all forms of local and national government and she doesn't agree with anything they do and therefore doesn't think she ought to move simply because they want her to. You are unconvinced by her argument and point out to her that just because we don't want to follow rules it does not mean we can just ignore them. She agrees, but says that this isn't her point. She didn't vote

for the government; she doesn't reap any benefits from the State, she didn't sign any forms consenting to obeying the rules. If she had voted or was benefitting from the State, then she readily agrees that she ought to follow the rules.

You are a bit less sure now of why she ought to obey the rules of the government. You suggest that perhaps she ought to obey them, not because she's explicitly consented to anything but because she is living on land maintained by government in a country which is run by the government. But she points out that she hasn't chosen to live there and she isn't rich and can't simply leave. Why then should she be bound by the rules?

You are now unsure. On the one hand you feel that she really ought to follow the rules. But on the other hand you can't work out why. Which is the right response here? Does she have an obligation to obey the rules? And if so, why?

Discussion of the authority of the State, political obligation and consent are central to political philosophy. There are different notions, each with a rich research literature. However, we can get a rough and ready understanding of some of the issues by considering this thought experiment. One way to explain why the State has authority over the individual is that individuals have *consented* to being part of the State. But it is hard to make sense of the citizens of the country consenting in this way – this is the point made by John Locke in *An Essay Concerning Human Understanding*. It was for this reason that Locke introduced the notion of *tacit consent*.

Tacit consent is, very roughly speaking, consenting without actually actively signalling consent through writing, verbal agreement or in any other way – even by just nodding one's head. So, for example, although we might not say 'I agree' or write down 'I agree' to a set of rules before we use a swimming pool, by jumping into the water we are agreeing to abide by the rules of the pool. We are tacitly consenting to allowing the life guards to have authority over what we do. Locke thought that if people *used* the footpaths, roads, etc., in a country and gained some (however minimal) benefit, then they were tacitly consenting to obey the rules.

However, as outlined in the example above, this is odd if we think about how consent normally works. Normally, consent only makes sense if we *freely* give it. Consent which isn't chosen seems like an oxymoron. So, tacit consent to obey the State would only make sense if we freely choose to stay in a country. But for most people it isn't a simple choice. Most people did not decide to live in that particular country, but are living there probably because they haven't got enough money to do otherwise. So if they don't explicitly consent, and they don't tacitly consent, the question is why should they obey the State at all?

Questions

1 Do you think that the woman in the thought experiment ought to move?
2 Do you think that as a citizen you *ought* to obey the State?
3 If the State is corrupt and unjust, then do its citizens have an obligation to obey it?

4 How do we gauge this injustice and corruption? If for example, 25 per cent of the population surveyed would judge the State as corrupt does that exempt them from obeying the State?
5 Give an example of something to which you might consent?
6 What things might be important in (4) in order for your consent to be genuine?
7 Do you think the notion of tacit consent makes sense?
8 How might the notion of tacit consent be abused?
9 If the State doesn't acquire authority through its citizens consenting (either explicitly or tacitly) then how does it get its authority?
10 Do we need a state and government at all?

Reading

Bennett, J. 1979. 'A Note on Locke's Theory of Tacit Consent', *The Philosophical Review*, 88(2), 224–34.
Carr, Craig L. 1990. 'Tacit Consent', *Public Affairs Quarterly*, 4(4), 335–45.
Green, L. 1990. *The Authority of the State*, Oxford: Clarendon Press.
Simmons, A. John. 1976. 'Tacit Consent and Political Obligation', *Philosophy and Public Affairs*, 5(3), 274–91.

---

96 Freedom and the State

How much power should the State have? Is there any justification *at all* for the State to have power? The anarchists would say not, as they think that the State should have no power at all. Or maybe, as Hobbes thinks, the State should have absolute power over anything it wants? Political philosophers have long struggled with the question about *how much* power the State should have:

Thought experiment

Imagine you think that the State should keep its nose out of people's business. For example, you think that it has no right to tell people who they can or cannot marry; or if people want to engage in same-sex relationships, then what has that got to do with the State? It really isn't the State's place to intervene.

However, on the other hand you do think that the State should have *some* say in what goes on. The State should stop drunken people from driving, and it should stop drug companies putting products on the market which they haven't tested correctly.

The 'personal/public' distinction doesn't make a difference here either. For you do think that the State should get involved in 'personal' matters. For example, if a husband is beating his wife, or if a couple are failing to feed their children, then the State should move to intervene.

These are quite common intuitions but the problem is how you can identify these cases? Why might the State be allowed to intervene in some instances but not others? How much power do you think the State should have? In 1859, John Stuart Mill (2008) put forward a very famous principle, the Harm Principle, which attempted to answer this question. In deciding how the State should be involved, he put forward the idea that unless our actions are *harming* others, the State has *no right* to intervene.

So if we want to worship many gods, or vote a particular way, or join a union, or eat or drink ourselves to death, then – according to Mill – we should be allowed to. Given that we are not harming anyone, the State has no right to stop us. This principle then means that we can agree that child abuse, domestic abuse or drink-driving should remain illegal, because they harm others. It may of course be unclear what *does* harm others, but just because it is an open question that does not mean that we should not see it as the correct principle to guide our actions. At least, that is what Mill thought.

This principle perhaps seems obvious to us, but it is quite radical. Especially if we consider how it relates to freedom of speech. Mill thought that censorship is an anathema because the occasions on which *speech* harms others are few and far between. Notice that Mill thinks there is a clear distinction between *offence* and *harm*, and believes that it is never right to stop someone causing offence. So, Mill thinks that if someone is expressing bigoted or homophobic views that may cause offence, there is no reason to stop them. Of course if a racist is encouraging and instructing other racists to set fire to a church then their speech should be stopped because that is leading to harm and isn't simply about offence. So the Harm Principle allows us to do what we want as long as people aren't harmed, and it allows freedom of speech and the right to offend.

People have worried about this very liberal account. Some think that 'harm' is actually not as clear cut as we might think. Hitting someone clearly causes harm. But if someone is bullied at school simply through spoken insults, is this harm? Or if you owned a bakery and another bakery set up business next door this would harm you in some sense. Or if you were racist you might think influx of foreigners was harming your identity, etc.

But if the Harm Principle can be so easily manipulated because the notion of 'harm' is so flexible, hasn't something has gone wrong?

Questions

1 What do you think of anarchism – the view that the State has *no* legitimate power?
2 What do you think of Hobbes's view that the State should have *absolute* and unlimited power?

3 If you think that (1) and (2) are wrong approaches, how might you demarcate those areas in which the State has legitimate powers and those in which it doesn't?
4 Some think that the State ought to have no say in our 'personal lives' – is that right?
5 Do you think that the State has a right to tell companies how to advertise their products?
6 What do you think of Mill's Harm Principle?
7 Do you think that people have the right to offend others?
8 What is the difference between causing 'offence' and causing 'harm'?
9 A Nazi might say that the existence of a Jewish state causes great harm. Are they wrong? How might you stop Mill's Harm Principle being used to pernicious and repugnant ends?
10 Do you think that there should be censorship on the Internet? If so, how would you go about deciding what should be censored and what shouldn't?

Reading

Holtug, N. 2002. 'The Harm Principle', *Ethical Theory and Moral Practice*, 5(4), 357–89.
Mill, J. S. 2008. *On Liberty and Other Essays*, Oxford: Oxford University Press.
Ripstein, A. 2006. 'Beyond the Harm Principle', *Philosophy & Public Affairs*, 34(3), 215–45.
Turner, P. 2014. '"Harm" and Mill's Harm Principle', *Ethics*, 124(2), 299–326.

---

97 Negative and positive freedom

Most people are in favour of freedom. In Chapter 96 we dealt with the question: 'to what extent is it legitimate for the State to limit people's freedom?' However, political philosophers, most notably Isaiah Berlin, have argued that 'freedom' is in fact ambiguous. It can mean *positive* freedom or it can mean *negative* freedom. Moreover, these are not simply two sides of the same coin but are different and competing accounts of freedom.

Negative freedom is the absence of constraint, which means that someone's negative freedom is increased in proportion to how far the constraints are removed from what she might want to do. Positive freedom is the freedom to fulfil one's potential, to realize one's purposes.

You might think that positive freedom has got nothing to do with the State and politics. But many philosophers, e.g. Jean Jacques Rousseau, have disagreed. Some think that in order for people to be truly free they need to have restrictions

put on them by society or be forced by society to enter into some types of activities or forced to live certain types of lives. This leads to the paradoxical-sounding claim that people can be *forced to be free*:

Thought experiment

> You have recently come to think that people can be forced to be free. This was brought home to you at the local kids' football match. The match was an important one, but the referee didn't arrive and unfortunately no one was prepared to take on the job.
>
> So, what was meant to be a civilized game soon descended into chaos with the bigger kids pushing over the smaller ones and the loudest voices being heard over the less vocal. In fact, as you saw with your own eyes, the people who really wanted to play football were not able to do so. They weren't free to be the best football players they could be.
>
> You realized that what was needed was some external limitation of their freedom (their negative freedom needed to be limited) in order to increase their positive freedom. Following this thought process you understood suddenly that it is best if the government limits negative freedom so that people's positive freedoms can be increased. Is this line of thought plausible?

We might think that this seems reasonable. However, imagine that later you watch a documentary about North Korea. In this documentary the all-powerful leaders are talking about how they strive to develop the freedoms of each individual in society; how they want what is best for all citizens; they want to create opportunities for them to flourish. However, as you know, in the case of this regime it means using brutal labour camps, censoring the newspapers and television networks, using strategies of terror and starvation. They are, in fact, trying to achieve positive freedom by severely restricting the citizens' negative freedom.

But can this be right? Surely not! A state ought not to restrict negative freedom like this. But on the other hand, even if the North Korea example is a bad one, the value of positive freedom development – as brought home to you by the kids 'playing' football without a referee – seems like something in which the State should also be involved. Are these compatible? If so, how do we strike the right balance between a government restricting negative freedoms in order to develop positive freedoms and what is best for the citizens of a state?

Questions

1. What is positive freedom? Give your own examples.
2. What is negative freedom? Give your own examples.
3. List some ways that a state could limit negative and positive freedoms.
4. Do you think that players have more freedom in games where the rules are rigidly enforced?

5 What would a state look like if negative freedom was its only concern?
6 What would a state look like if positive freedom was its only concern?
7 How far do you think a state should restrict negative freedoms in order to encourage positive freedoms?
8 Imagine that the State could mass-hypnotize all its citizens to want nothing but bread and water, which the State could always provide. What, if anything, is wrong with this State?
9 How would you explain (8) in terms of negative and positive freedom?
10 How might this discussion relate to education?

Reading

Blau, Adrian. 2004. 'Against Positive and Negative Freedom', *Political Theory*, 32(4), 547–53.
Kaufman, A. 1962. 'Professor Berlin on "Negative Freedom"', *Mind*, 71(282), 241–3.
MacCallum, G. 1967. 'Negative and Positive Freedom', *The Philosophical Review*, 76(3), 312–34.
Warburton, N. 2001. *Freedom: An Introduction with Readings*, London: Routledge.

---

98 What's so great about democracy?

We live in a state: that is undeniable. So the question that follows is: what sort of government should we have? Most people in the 'West' would think that the best government is one which is *democratic*. Countries have been invaded and sanctions have been put in place in the name of 'democracy'. However, some philosophers have been unconvinced that democracy is in fact the best form of government at all. Here is a simple but powerful argument put forward by Plato:

Thought experiment

If something is a skill, then people can be good or bad at it. If I wanted someone to cook me a quality meal I wouldn't ask the local police officer (unless of course he was a renowned chef). If I want my appendix taken out I don't want someone who is not qualified to carry out the operation. If I am on a ship I *would* want the captain to navigate the ship – as she is trained and skilled in such things.

But arguably *ruling a country* is a skill that requires a very specific and developed set of qualities, characteristics and experiences.

If this is true, then answering 'how do we get the best government?' comes down to the question 'how do we get the most skilled rulers?' Once we put the question like this it isn't at all clear that voting is the best process to achieve this.

> The voters aren't necessarily going to know or perhaps care who the most skilled ruler is. People typically haven't had the time and training to know what the qualities of a skilled ruler are. So democracy isn't the best form of government, for it doesn't guarantee that the right person is given the power to rule.

This line of thought is simple and powerful. Plato, who put forward a sustained attack on democracy along these lines, thought that the most skilled people for ruling are the philosophers. For 'philosophers', though, don't think of academics or philosophy students, think of people in late middle age who have experienced not only a wide variety of sustained academic training, e.g. music, history, maths, etc., but have also experienced much of 'life'. Even if we don't agree with Plato's conclusion about who ought to rule, might we think that there is something disturbing and right about Plato's attack on democracy?

After all, there are some nasty people in our country, some people who aren't clever, or well read. There are people with self-serving agendas, who are bigoted and narrow minded, etc. Why should they get the same say as you or me? Why should we have to abide by rules and laws made by a government which has been put there by such people?

On the other hand, despite these concerns we might challenge this line of thinking:

First, we might think that ruling *isn't* a skill in the same way that being able to cook is. The government isn't one person; it is a collection of people. These people have lots of different skills and are able to tap into the skills of many other people. So the fact that 'the people' aren't experts at identifying the 'key' skills for ruling doesn't really matter.

Second, the most usual response to this type of challenge to democracy is that, although it is not perfect, it is the best we have. If we look at various other systems which have been tried: monarchy, dictatorship, oligarchy, etc., we see a litany of death and destruction, and many unhappy citizens. So, the defender of democracy might suggest that given the options on the table, we should go for democracy.

Questions

1 What is a democratic government?
2 Name two other forms of government.
3 Do you think that 'ruling' is a skill?
4 If you think the answer to (3) is 'yes', then what do you think of Plato's argument? If you think the answer is 'no', what is it?
5 If Plato is right and we should get rid of democracy, what do you think the best form of government is?
6 What do you think of the suggestion that philosophers should rule?
7 If you were to decide who should rule, who would it be and why?

8 Why do you think that democracy has become the default form of government in 'the West'?
9 What might be the biggest problem with putting 'experts' in government?
10 Could it be good to have a democracy, even if a country is worse off under democracy?

Reading

Mason, M. 2009. 'Democracy in Plato's *Republic*', *Southwest Philosophy Review*, 25(1), 93–105.

Santas, G. 2007. 'Plato's Criticisms of Democracy in the *Republic*', *Social Philosophy and Policy*, 24(2), 70–89.

Santas, G. 2012. 'Democracy Then and Now: Plato, Mill, and Rawls on Wealth and Ruling', *Philosophical Inquiry*, 36(1–2), 1–12.

Saunders, B. 2010. 'Democracy, Political Equality, and Majority Rule', *Ethics*, 121(1), 148–77.

---

99 Terrorism

There isn't a day that goes by that the media doesn't talk about 'terrorism'. Terrorists make attacks across the globe, politicians talk about terrorist policy and how their government is best placed to tackle the 'problem of terrorism'. But what is terrorism? And is it always wrong? Political philosophers have recently become more interested in this topic because of recent high-profile terrorist attacks:

Thought experiment

Imagine two people: a Jewish person living in Israel and an Arab living in Palestine. Both feel they have a right to the land that they occupy. Both support action – whatever it takes – against the other to allow the liberation of further land and greater sovereignty.

The Jewish person supports forced evictions, the use of snipers and strategic air strikes; the Arab supports the use of missiles and the bombing of public areas. Both agree that each action is awful – it is not something that they seek as an end itself. If they thought the desired end could be brought about by some other – non-violent – means then they would of course take it. It is just that they can't.

They also both agree that civilians will die; there is no way to bring about their aim without men women and children being killed. They both deny that they are 'terrorists' or are siding with 'terrorists' and each thinks the other is a 'terrorist'.

What do you think? Is either right? Is neither right? How might you try and work out the answer to this question?

The fact that this thought experiment is about Jewish and Arab people is irrelevant. You can simply re-write the thought experiment with any people you want. But it does raise an interesting question. How might we define a 'terrorist?' 'Terrorist' isn't a label that many people want attached to them, and it has become purely pejorative (this hasn't always been the case). People want to distance themselves from being called 'terrorists' and often call themselves 'freedom fighters' or something similar – anything but 'terrorist'.

But as we have seen in the thought experiment, it does seem that one person's 'freedom fighter' is another's 'terrorist'. Where the Arab is doing all she can for her freedom, the Jewish person sees this as terrorism pure and simple, and vice versa. Perhaps then the best way to approach the task of defining 'terrorist' is not in terms of *who* is carrying out the action – because as the thought experiment shows this might get us nowhere – but rather by focusing on the actions themselves. Once this has been sorted out, then we can just say that whoever is performing those types of actions is a terrorist. This means that individuals, governments, groups, etc. could be terrorists, and aren't ruled out simply because of who they are. Let's try this, then, and see how far we get.

It seems that the first answer to what does 'terrorist' mean must have something to do with bringing about some goal by creating 'terror'. To be a terrorist a terrorist needs to bring about terror and/or fear amongst people. This means that it is possible that a terrorist might never actually be violent. We could imagine *threats* of violence coercing people to act in the way desired by the terrorist.

Moreover, if this is our starting point then we might think that there is no need for violence (or threats of violence) against *people*. If a group claimed they were going to shut down all power to a city, or if they could shut down the Internet, then presumably both these actions *might* not actually bring about any harm to people. But would this mean they were not – in light of that very fact – not terrorist actions? Let's leave this and assume that a terrorist – or terrorist organization – is someone or some group who threatens violence, or is violent, so as to coerce people to bring about their desired end.

But this seems incomplete – or at least controversial – in a number of ways. It is true that coercion through terror is important, but is it true that *any* end is important? Imagine that a group of people started blowing up buses in order to secure a billion dollars from oil firms, so that they could live in luxury. Should this be classed as a terrorist activity? Many would say 'no'. Rather the aim needs to have some weightier reason, perhaps with a 'political' dimension. So if people were blowing up buses in order to secure a billion dollars so that they could then help liberate animals from cosmetic testing, this might be better classed as 'terrorism'.

What though of *who* is killed? If an Army drone commander released a missile into an enemy army barracks killing many soldiers, presumably we wouldn't class this as 'terrorism'. Rather we might say this is an 'act of war'. So perhaps, then, those killed need to be non-combatants or 'innocents'? Although something like this seems right it does open another, bigger issue. Namely, in what sense are people 'innocent'?

If I wanted to stop animal testing and thought it OK to blow up animal testing labs, I might argue that the scientists and cleaners who work there *aren't* 'innocent'. They are part of the problem. Or if I was going to blow up a government building because I wanted an end to the government's divisive foreign policies, I might argue that anyone killed in the bombing – even if they didn't work directly for the government – was not innocent. After all they voted for, or at least don't speak out against, the government's policies, and are then in some way complicit. Hopefully you can see how complex and murky this issue is.

And we haven't even touched on the further issue of the moral status of terrorism. We might think that terrorism is always wrong. But then what of groups who seem to be fighting for what is morally right, for example the French resistance in the Second World War?

Questions

1. Name three terrorist organizations.
2. Do you think that *who* is carrying out the action makes a difference to whether it is an act of terrorism or not?
3. Do you think that a terrorist has to bring about harm to people?
4. Terrorism is thought to be about 'coercion'. What is 'coercion'?
5. Do you think that the aim of a so-called 'terrorist' act could be simply to get money to live in luxury? If you do, then could you categorize some significant robberies as terrorist activities? If not, then what does the aim of a so-called 'terrorist' act have to be?
6. Imagine that two countries are at war. Could one country carry out a terrorist act on the other country?
7. What is it for one country to be 'at war' with another country?
8. Some definitions of terrorism talk about killing 'innocents'. What do you think of this? How might you interpret the idea of 'innocents'?
9. The term 'terrorism' has not always been thought of as pejorative. Why do you think it has come to be thought of in this way?
10. Could an act of terrorism ever be morally acceptable?

Reading

Bauhn, P. 1989. *Ethical Aspects of Political Terrorism: The Sacrificing of the Innocent*, Lund: Lund University Press.

Coady, C. A. J. 1985. 'The Morality of Terrorism', *Philosophy*, 60(231), 47–69.

Jollimore, T. 2007. 'Terrorism, War, and the Killing of the Innocent', *Ethical Theory and Moral Practice*, 10(4), 353–72.

Primoratz, I. 2013. *Terrorism: A Philosophical Investigation*, Cambridge: Polity Press.

100 Multiculturalism

One contemporary issue that quickly divides groups of people is *multiculturalism*. Politicians are often seen as either being 'in favour of' or 'against' multiculturalism; where both labels are used in a pejorative way. It is very difficult to define what 'multiculturalism' actually means, and it is made harder by the fact that we first have to be clear about the meaning of 'culture'.

We can arrive at a rough and ready understanding of multiculturalism by seeing it as a way of *proactively* responding to the fact that the different cultures in modern society should be respected and treated in a positive way. The important word here being 'proactive'. Multiculturalism isn't just acknowledging or tolerating other cultures but is about *actively* making sure that other cultures are respected politically, legally and morally. So, an example of multiculturalism might be providing public information in a number of different languages.

Let's, as far as possible, leave the thorny issue of definitions to one side and instead ask what might be said for or against multiculturalism:

Thought experiment

> Julie has just started secondary school. One day she comes home quite angry. 'We had to sit through two hours of boring Christmas rehearsals today!' You point out that this is meant to be fun and that 'after all everyone has to do it, it isn't a choice.' Julie gets visibly annoyed: 'But that isn't true. The reason I'm so angry is because the Jones children,' who she reminds you are Jehovah's Witnesses, '*don't* have to get involved!'
>
> You explain that this is because their religion excludes them from celebrating Christmas. Julie responds: 'That seems quite convenient! Perhaps I should start a religion that excludes me from going to school! Well I for one have had enough and I'm going to see my teacher tomorrow to complain about it!'
>
> Do you think Julie has a point? Do you think that we should make these types of exceptions for religious groups?

It is true that Jehovah's Witnesses don't celebrate Christmas, and hence can be excluded from certain lessons and celebrations. And there are many examples of where different religions and/or cultures are not only just tolerated but accommodated. For example, schools putting Kosher or Halal food on the canteen menu; Islamic women being allowed to wear different clothing. We might think that preserving distinctive parts of the culture is important for the *individual* to develop a sense of belonging, a sense of history and a shared sense of direction, etc. It might be important for individual well-being.

Furthermore, if we value *freedom*, then we should enable individuals to have the best choice possible and we increase choice by encouraging different cultures. For example, if the government is stopping me worshipping on a Saturday, or stopping

me holding a gay pride march, then my options are limited and my freedom is limited. So multiculturalism increases a value which we all hold, namely freedom.

Finally it might be the case that the dominant culture 'owes it' to the other minority cultures. We can think of examples of multiculturalism such as the enshrining of certain property rights for Native Americans or Aborigines, that are justified by citing how badly they have been treated. So the idea is that some cultures not only need to be 'left alone' but need to be actively 'compensated' through the actions of the dominant culture – e.g. through changes in the law, etc.

However, it does seem that there are some serious problems with multiculturalism. We can use feminism to show one serious issue. Feminists would want all women to have the vote, or be educated, or drive a car, to cite just a few examples, and so if we can conceptualize 'women' as a culture, then feminists can use multiculturalism to argue that societies should be active in supporting and recognizing women.

But, and here is the difficulty, if multiculturalism is used as a key justification for feminism then this means that feminists are unable to oppose the practices of some cultures. In particular, what about cultures which exploit women, which treat women as less than human; which ban women from being schooled, or support female genital mutilation?

Of course this point generalizes. Think of any characteristic you want, say someone's sexual orientation. Imagine a culture – and we probably don't have to think that hard – that treats homosexuality as 'evil' or 'a curse'; that treats homosexuals as an anathema and doesn't allow them voting rights, freedom of expression, etc? It seems that if we support multiculturalism we are legitimizing that behaviour because it is part of a culture.

Questions

1 In the thought experiment, do you think the Jehovah's Witnesses should be allowed to take their children out of lessons connected with Christmas?
2 Think about the same story, but imagine that fundamentalist Christians wish to take their children out of science lessons that deal with evolution and sex. Should this be allowed?
3 Write down four different 'cultural groups'.
4 Look at your answer to (3) and think how you might start to define 'culture'.
5 What is the difference between 'tolerance' and 'multiculturalism'?
6 Why might 'tolerance' be thought to be inadequate?
7 'If individual freedom is important, then multiculturalism is the best position.' What do you think of this claim?
8 One defence of multiculturalism is that history has often shown how one culture has been oppressed and exploited by another. For example, think about how the Native Americans were treated. In that case, the minority culture is 'owed' certain things from what was – in the past – the oppressive culture.

9 Is there any way of reconciling the difficulty that in respecting multiculturalism we are failing to respect certain individuals within the minority groups themselves – e.g., women, atheists, etc.
10 If you rejected multiculturalism, what are the other options?

Reading

Okin, S. 1999. 'Is Multiculturalism Bad for Women?', in J. Cohen, M. Howard and M.C. Nussbaum (eds), *Is Multiculturalism Bad for Women?*, Princeton, NJ: Princeton University Press.

Scheffler, S. 2007, 'Immigration and the Significance of Culture', *Philosophy & Public Affairs*, 35(2), 93–125.

Song, S. 2008. 'The Subject of Multiculturalism: Culture, Religion, Language, Ethnicity, Nationality, and Race?', in B. de Bruin and C. Zurn (eds), *New Waves in Political Philosophy*, New York: Palgrave Macmillan.

Waldron, J. 1992. 'Superseding Historic Injustice', *Ethics*, 103(1), 4–28.

Part XI

PHILOSOPHY OF RELIGION

101 The ontological argument

We tend to think that, on purely conceptual grounds, we are unable to work out what exists and what doesn't. Thinking about the concept of unicorns, chipmunks or chairs won't tell us anything about whether unicorns, chipmunks or chairs do actually exist. Famously, St Anselm thought that God is different, and that if we think about the concept of God we can show that He must necessarily exist:

Thought experiment

> When your atheist friends tell you that 'God doesn't exist' you ask them what they mean by 'God'. They respond that they mean something like: 'The ultimate perfect being that has the features of omniscience and omnipotence, and is all-loving.' That is their idea of God, but they do not believe He exists.
>
> So, you suggest that if something isn't perfect then that thing cannot be God. They agree. The concept of God – which they are quick to remind you is only a concept and doesn't exist – is one of a being than which there is none greater.
>
> Knowing that your friends like football, you ask them if they think it is better to have two existing tickets to the cup final, or two tickets that simply exist in the mind. A bit bemused, they claim it is clearly better for the two tickets to exist and be in their hand and so, you suggest, it must be better to exist than not to exist.
>
> Unsure what you are on about, your friends agree. 'OK, well then, you continue, if God is – by definition – the greatest being we can think of, and it is better for something to exist than not to exist, God Himself must necessarily exist. For if He didn't He wouldn't be God, because He wouldn't be the greatest being.' So, you conclude God by definition must necessarily exist.

This is a rough version of the ontological argument ('ontology' referring to what exists). In fact, more correctly, it is a collection of broadly related arguments. It is a rare form of religious argument, as it is based on 'a priori' reasoning, which means that, whereas other arguments for the existence of God (see chapters 106 and 107) ask us to look out at the world and make observations and generalizations, this argument does not. Rather it asks us to consider and reflect on the very concept of God. A concept which – as the example above illustrates – even the atheist has.

One initial response is to claim the argument fails because it leads to absurd conclusions about what must exist. We might think that if a perfect God must exist then a perfect anything must exist – such as a perfect biscuit, a perfect island, a perfect unicorn. But this is crazy, and consequently the ontological argument must be wrong.

But this misses some fundamental features of the original argument. First, it isn't part of the very concept of a biscuit, unicorn, island, etc. that they are perfect. Second, and more damaging, it isn't clear at all what it means to talk about a *perfect* biscuit, a perfect island, a perfect unicorn. Does the island have to have ten or eleven trees? Golden sand or a pebbly beach? Lots of fish or hardly any fish? Or what about the biscuit – does that have to have a crunch? Or is it soft? Or is it chocolate or a wafer?

Another important clarification is that the argument doesn't rely on the claim that the world is a better place if a particular thing exists rather than doesn't. A world without Hitler would have been better than a world with Hitler. The point is that for something to be the best possible thing *it can* be, it needs to exist. So, for Hitler to be the best Hitler he can be, he needs to exist.

This argument has been influential not just within the philosophy of religion, but also in other areas. For what it asks us to consider, amongst other things, is the nature of existence, the nature of properties, the nature of perfection and the link between concepts and reality.

Questions

1 What does it mean to say that this argument is based on 'a priori' argument?
2 Write out the ontological argument in your own words.
3 What do you think about this form of argument?
4 What does the concept of 'God' mean?
5 When an atheist talks about 'God' not existing, what might she mean?
6 Do you think that it is greater for something to exist than not to exist?
7 What is it for something to be 'perfect'?
8 Does a perfect island exist?
9 Some people think that this argument relies on existence being a property. Does it?
10 Do you think that a rational argument such as the ontological argument can ever 'convert' someone? If not, why not?

Reading

Saint Anselm (Archbishop of Canterbury 1093–1109). 1965. *Proslogion*; with 'A Reply on Behalf of the Fool' by Gaunilo, and 'The Author's Reply to Gaunilo', South Bend: University of Notre Dame Press.

Dummett, M. 1993. 'Existence', in his *The Seas of Language*, Oxford: Oxford University Press, 277–307.

Matthews, G. B., and Baker, L. R. 2010. 'The Ontological Argument Simplified', *Analysis*, 70(2), 210–12.

Tooley, M. 1981. 'Plantinga's Defence of the Ontological Argument', *Mind*, 90(359), 422–7.

102 Pascal's Wager

Blaise Pascal wanted a way to persuade his gambling friends to think seriously about believing in God. To this end he famously put forward what has become known as 'Pascal's Wager':

Thought experiment

Imagine that you are undecided about whether to believe in God. You have two options. You either believe in God or you don't (for Pascal, being undecided is the same as not believing). Pascal thinks that there are two options when you die. Either there is a heaven and a hell. Or there is nothing.

Given these options what is it *rational* for us to believe? If we are going to gamble with our lives we'd better think hard! We should work out the desirability of the outcome of each possible decision.

If we believe in God and there *is* a God, and a heaven and hell, then we go to heaven and get eternal bliss. Pascal thinks that if this were the outcome then we would achieve an *infinite positive value*.

If we believe in God and there is *no God* and no heaven and hell (just nothingness) then we will have lived a good life and Pascal thinks we'll be happy. So let's assign this outcome an arbitrary numerical value of 500. So if we *do* believe in God there are two possible outcomes: either infinite value or a 500 value.

Now imagine that we *don't* believe in God. If we don't believe in God and there *is a God* then we'll go to hell which is an *infinite negative value*.

If we don't believe in God and there is no God then there will be some value, but Pascal argues not as much as 500 because we won't have lived as good or as happy a life as someone who had lived their life believing in God. Let's call this value 300. So if we don't believe in God the two possible outcomes are infinite negative value or a 300 amount of value.

Looking at the options and outcomes, Pascal concludes that belief in God is going to give you the best outcome. Therefore, it is more rational for us to believe in God than not believe in God.

First, we might think that even if Pascal is right, we can't simply *make* ourselves believe in God. We can't simply choose what to believe. If you were asked to believe that the book you are reading is made of chocolate, and moreover if you did you would get a million pounds, you couldn't do it. Beliefs are not the sort of thing we can simply adopt at will. Although this point seems right, this is no challenge to Pascal. For Pascal isn't using his wager to *force* people to believe in God. He is simply giving people one more reason to think that belief in God is reasonable. Having considered his argument we might, for example, start attending church or reading the Bible to try and cultivate the belief in God.

A second response runs something like this. If we are atheists and we think God doesn't exist, then we can't assign *any* value to the options which talk about God existing. After all, we do not believe that God exists. However, this response

doesn't work either. For unless the atheist can show that it is *logically* impossible for God to exist, we must be allowed to say that it is a possibility (however remote) that God exists. And if it is possible then we can assign a value to that outcome occurring.

A third response is the following. What Pascal's Wager has shown us is that it is rational to believe in a God *or* Gods. That is to say, there is nothing in Pascal's Wager to show it is more reasonable to believe in a Christian God than, say, the many Gods of Sikhism or the God of Islam, etc. Perhaps, though, Pascal would not be worried by this? If he has shown that it is more reasonable to be a theist of whatever stripe, then arguably he has done what he has set out to achieve.

A final and more powerful consideration is about how Pascal set out the wager in the first place. For his wager to work we need to be allowed to assign an *infinite positive value* to the scenario of belief in God, and God existing. Recall that this was meant to be because we will be in heaven with God for eternity.

However, we might think that the notion of an infinite value makes no sense. We can imagine a great experience, and a greater, relatively speaking, experience. But an infinitely valuable experience? Can we make sense of this?

Questions

1. Considering Pascal's Wager do you think it is more reasonable to believe in God than not to believe in God?
2. If someone thought the answer to (1) was 'no', then what mistake has Pascal made?
3. More sophisticated versions of Pascal's Wager introduce probabilities of the outcomes obtaining. So for example, although the outcome of believing in God and God existing is infinite, the likelihood of such an outcome is, say, 0.000000001. Does that affect the argument at all? (Tip: anything multiplied by infinity is infinity.)
4. Do you think we can make ourselves believe something?
5. If you could show that it is logically impossible for God to exist, then this argument would fail. How might someone show that it is logically impossible that God exists?
6. Do you think that the argument would work if we substituted 'faith' for 'belief'?
7. Do you think Pascal would have been happy with the fact that his argument concludes that it is reasonable to believe in any God or Gods?
8. Do you think a life lived as if God existed – even if He does not – is better than a life lived as if God didn't exist – even if He does?
9. Do you think it makes sense to talk about infinite value?
10. Consider your answer to (7) and how your answer might be changed by the observation that in some religions, belief in many Gods or the 'wrong' God means that we don't get eternal bliss?

Reading

Carter, A. 2000. 'On Pascal's Wager, Or Why All Bets Are Off', *The Philosophical Quarterly*, 50(198), 22–7.

Hájek, A. 2003. 'Waging War on Pascal's Wager', *The Philosophical Review*, 112(1), 27–56.

Pascal, B. 1968. *Pensées*, translated with an introduction by A. J. Krailsheimer, Harmondsworth: Penguin.

Saka, P. 2001. 'Pascal's Wager and the Many Gods Objection', *Religious Studies*, 37(3), 321–41.

---

103 Divine foreknowledge and free will

Many religions and religious people talk as if we have free will (see Chapter 37). For example, they talk about us *freely choosing* to obey God, or *intending* to sin, or *following* the religious commandments, etc. However, philosophers of religion have been concerned that this isn't possible:

Thought experiment

> You are trying to comfort your friend who has just undergone a painful break-up with his partner. You point him to a passage in the Bible that says that our actions are known by God, that He has a plan for us, and that this was true even when we were in our mother's womb. 'Someone is looking after you,' you say.
>
> Far from comforting him this makes him even more upset. 'Does this mean that God caused me to suffer in this way! That is really evil!' You explain to him that even though there is a plan for everyone's life, you don't believe that God *makes* a person act in a certain way. You explain that as God is omniscient, He does know what will happen, but that does not mean that humans do not have free will.
>
> But your friend is unconvinced. 'But if He really did know that I was going to split up with my partner, then I couldn't have done otherwise, could I?'
>
> 'Why not?'
>
> 'Well, if I did otherwise, He *wouldn't* have known it!'

Now you see his point. If God *knows* what we are going to do, even if he doesn't intervene in our lives, it seems that we haven't got the power to do other than what we in fact do. But if we don't have the power to do anything else than what we in fact do, then we aren't free. Because it seems that free will is precisely about having the power to do otherwise. So do you think that we have free will and believe that there is an all-*knowing* God?

There are a number of things to make clear from the start. It is easy to think that this problem is about fore*ordination*. It isn't. There is an issue with that, but we are not concerned with that in this case. The question is about *knowledge*. Could it be the case that God *knew* that I was going to go, say, to the shop – which surely He did – and yet I was still free to choose not to go to the shop?

Notice also that this isn't about *predictability*. It isn't a solution to this problem to say that, because God knows the *type* of person you are, He will have an idea how you will behave. Because having an idea of what we will do isn't the same as *knowing* what we'll do. I can know that in the past John has become angry when he lost at squash, but knowing that about John's character doesn't mean that I *know* what he is going to do tomorrow if he loses the squash game.

A useful way of looking at the problem is to remember that the problem of divine foreknowledge is silent on *how it is* that God could know these things. The point is simply that it seems like the claim (a) we are free, and (b) God is omniscient are simply incompatible.

Questions

1 What is preordination?
2 What is free will?
3 The problem is outlined in such a way that it sees free will as having 'the power to do otherwise'. Is this what we mean by free will?
4 Why might a theist not want to give up human free will?
5 Explain in your own words the seeming incompatibility of God's foreknowledge and human free will.
6 One response to this problem is to claim that God isn't omniscient. What do you think of that response?
7 Sometimes God is thought to exist 'outside time'. If that is true, does this help with this problem?
8 If God is outside time does that mean He can't be a personal or interacting God?
9 What is it to know that someone will act in a certain way? Focus on what 'knowledge' means.
10 What, in the end, is wrong with giving up on human free will?

Reading

Hasker, W. 1985. 'Foreknowledge and Necessity', *Faith and Philosophy*, 2(2), 121–56.

Pike, N. 1965. 'Divine Omniscience and Voluntary Action', *The Philosophical Review*, 74(1), 27–46.

Rhoda, A. R., Boyd, G. A. and Belt, T. G. 2006. 'Open Theism, Omniscience, and the Nature of the Future', *Faith and Philosophy*, 23(4), 432.

Zagzebski, L. 1985. 'Divine Foreknowledge and Human Free Will', *Religious Studies*, 21(3), 279–98.

104 Worship

Brown and Nagasawa (2005) have recently developed a fascinating argument about the traditional theistic claim that people have an *obligation* to worship God. They argue that *God cannot oblige us to worship Him*.

Typically, Christians hold that if, for example, God commands us to give money to charity then we have an *obligation* to give money to charity (see Chapter 74 on Divine Command Theory). This is because they are committed to the general thesis that if God commands x, then we are obliged to do x. The obligation arises not because of *what* is being commanded but by the very fact that God is commanding it; that is, Christians typically think that the reason that humans are obliged to do x is precisely for *the reason* that God commands x.

Furthermore, it is a central tenet of Christianity – and in fact many other world religions – that humans have an *obligation* to worship God. It is not just good of us if we happen to worship God, but rather if we don't worship God we have failed in what we *ought to* do. The challenge arises because the claim about God's commands creating obligation, and the claim about being obliged to worship God, seem to be in tension:

Thought experiment

> Jokingly your friend tells you that because you are '*so* rich' you – personally – *ought* to pay off world debt. This clearly would be *impossible* . . . . " . . . you aren't *that* rich!" You can't be *obliged* to do something that you just *can't do*. Generally speaking then, if you ought to do something, then you must be able actually to do that thing.
>
> So you get thinking about the fact that you *ought* to worship God. Is it true that you *can* worship God? Is it possible? It certainly seems so. You know many Christians who claim to be living their lives as worship to God. However, now you see a problem . . .
>
> Jokingly you respond to your friend that given that you are so rich, he must worship you. You command it! Your friend looks blankly at you and shrugs his shoulders. 'If you are commanding me to worship you, I *can't* do that. You can't command me to worship you any more than you can command me to love you or command me to want to spend time with you.'
>
> This seems right, you can't command someone to worship someone. Well, strictly speaking, you can, it is just that they can't fulfil that command. It seems that people have to worship *freely*. But then you realize that God commanding us to worship Him means that, simply based on the fact that it is a command, we *can't* worship Him.
>
> But we have already agreed that if God commands something then, in virtue of that very fact, we *ought* to do it. So something has gone wrong. On one hand we ought to worship God because he commands it, but on the other hand if He commands it we can't worship Him.

This seems like a genuine problem in the notion of worship. But to get fully clear on the issue we need to make some clarifications:

- The claim *isn't* that it is *impossible* to worship God, because this is just false. We just need to visit a church, mosque or synagogue to see this. Of course, we could state that even though all these millions of people think they are worshipping God, they are in fact not doing so. But even if all these people were deluded it would not be relevant to the argument at hand, because this is about the notion of worship as it fits within a theistic framework. Not what happens 'on the ground'.
- The claim isn't that people can't justify why we might worship God. If, for example, you ask Christians *why* they worship God, they will cite many reasons. They might talk about being made in God's image, they might talk about the gift of grace; about being saved; about responding to God in love; they might talk about the 'fear of God', etc. So there are many plausible and legitimate reasons that people might, and do, give for worshipping God.
- Rather, the problem is the two claims that we ended the thought experiment with. It seems that *either* we have to give up the general claim that if God commands something we ought to do it, *or* we have to give up the claim that God commands us to worship Him. The problem is that both seem very plausible.

After all, many religious people think that God is the ultimate authority for why we should behave in certain ways. God is the ultimate arbitrator, the great 'I AM', the ultimate sanction; He is all-powerful, all-knowing, all-loving, etc. We are nothing in relation to Him, so when He commands us to, say, give money to the poor, it really does seem like that command is enough to ground our obligation to do so.

Perhaps then the way to proceed is to deny that God commands us to worship Him. This thought might be hard to reconcile with the way that the Holy Scriptures talk about God. Islam, for example, is about 'submission to God and the worship of Allah as the only true God'; Christianity has a commandment that humans should worship God alone. So it seems like there is a centrality in religions that we have to worship God. But maybe there is a gap here? Perhaps we could find a way to claim that although we ought to worship God, there is no evidence that God *commands us* to worship Him?

There is a final point here worth considering. We might revisit an assumption in the thought experiment. Perhaps, contrary to what is written, it *is* possible for someone to command someone else to worship and based on that fact we worship them?

Questions

1 Can you put this challenge in your own words?
2 Do you think that Christians (or any other theists) think we *ought* to worship God?
3 Do you think that Christians (or any other theists) have to be committed to the claim that if God commands x, then we ought to do x?
4 Thinking about (3) do you think that Christians have to be committed to the claim that the reason we ought to do x, is that God commanded x.

5 Is it true that if we ought to do x, then we must be able to x?
6 What is it to worship someone or something?
7 Do you think that we can be coerced to worship something?
8 If you think that the answer to (7) is no, why? Is your reasoning going to generalize to other mental attitudes such as love, want, desire?
9 What do you think counts as coercion? For example, what if you were told as a child that if you don't worship God then you might not go to heaven. Is that coercion?
10 Considering your answers to (1)–(9) how do you think a Christian should respond to this challenge?

Reading

Bayne, T. and Nagasawa, Y. 2007. 'The Grounds of Worship Again: A Reply to Crowe', *Religious Studies*, 43(4), 475–80.
Brown, C. and Nagasawa, Y. 2005. 'I Can't Make You Worship Me', *Ratio*, 18(2), 138–44.
Crowe, B. D. 2007. 'Reasons for Worship: A Response to Bayne and Nagasawa', *Religious Studies*, 43(4), 465.
Smuts, Aaron. 2012. 'The Power to Make Others Worship', *Religious Studies*, 48(2), 221–37.

---

105 The problem of suffering

One issue that has, above all, turned people away from belief in God is the problem of needless suffering. It just seems repugnant to people to believe in an all-loving and powerful God when there is so much suffering in the world. This issue has produced a vast amount of literature. Philosophers of religion have separated the issue into two broad problems. On the one hand we have the suffering caused by people on people: murder for example. On the other hand there is the suffering caused by 'nature', e.g. cancer and tsunamis. In this chapter we'll be concerned with the first. We can get to the heart of the problem of needless suffering caused by humans by a simple idea:

Thought experiment

> Imagine that a father sees his son being punched and teased in the street. It seems very obvious to him that his son is scared, upset and in pain. It seems perfectly sensible for us to think that the father will stop the bullying. But if he *doesn't*, what might we say?

If, contrary to our expectations, he didn't stop this we would be surprised, and we would think that this inaction was something that demanded an *explanation*. For example, perhaps:

- The father *didn't know* the bullying was taking place (maybe he thought it was part of a joke his son and his friends were playing). Or maybe:
- The father was scared and simply *wasn't powerful* enough to confront the bullies and felt 'out of his depth'. After all, he is going to be a better father alive than in the hospital, or dead. Or, finally, maybe:
- He *doesn't love* his son and rather thinks it would be good if his son gets a good kicking. It will 'make a man of him', etc.

Hopefully the analogy to Theism is clear. The father is meant to represent God. Once we make this link we can see how powerful the problem of suffering is. For if the father is God then it seems that the three explanations above aren't available.

- If the father in our story is God, then He is *all-knowing*, so it couldn't be the case that he didn't know the bullying was going on. It couldn't be the case that God says 'I could have helped stop the suffering, I wanted to help, I just didn't know it was going on.'
- If the father in our story is God, then He is *all-powerful*, so there would be no fear and worry about getting involved in the situation. Hospital, death, mugging, etc. don't bother God.
- Finally, imagine the father is *all-loving* such that he loves his son completely and utterly – more than anyone could love anyone else. The response that he just wants to see his son get beaten up isn't available.

Now we can see how serious the problem of suffering is. For it seems that the existence of such a God – a God who is all-knowing, all-powerful and all-loving, seems incompatible with human suffering. Many people have denied the existence of God for this very reason. Others, wanting to believe in the existence of God, have tried to answer the problem in many ways. Here are a few:

- One response is to *deny that needless suffering exists*. Perhaps it is, for example, 'all in the mind'. Of course, if there is no suffering then there is no incompatibility with God. If there is no incompatibility, then we don't have to deny that God exists. But on the face of it this seems like a very odd response. Just open your eyes. And even if we think suffering is somehow an illusion – it is still suffering!
- Another response is to plead *ignorance*. If it is true that God exists and is all-knowing, all-loving and all-powerful, then that just shows that *if* there is such suffering in the world then it *is* compatible with God. It just *seems* incompatible because we are finite and largely ignorant beings. We just don't know the reasons for what is going on. God knows how He and suffering can coexist; we don't.
- Another response is to say that if God was going to create the *best* possible world, then that world needed to be one in which humans have free will. This is because the relationships we form with each other and God are possible and are enriched if we enter into them freely (see Chapter 104 on worship). So God

has to create a world in which humans are free. But if humans are free, then humans are free to cause great harm and suffering for each other. Suffering, then, is a by-product of free will, and humans have to be free because God has to create the best possible world for us. This is why the existence of God is compatible with suffering.

Questions

1 Explain what the problem of suffering is.
2 What do you think of the analogy between God and the father in the example?
3 Why do you think that this argument is often thought to be the most powerful argument against God's existence?
4 What do you think of the claim that there is needless suffering in the world? How might someone deny this?
5 What do you think of the claim that only God knows why needless suffering occurs and hence this problem arises because we are finite and ignorant beings?
6 What do you think of the claim that suffering maybe a good thing?
7 Do you think that a world where people are free is better than a world where people are not free? Why?
8 Could God create a world in which humans are free but where the humans cause no needless suffering?
9 What do you think about changing God's attributes as a response to this problem?
10 If you were a theist what might you say in response to the problem of suffering caused by 'nature', such as tsunamis and cancer?

Reading

McBrayer, J. and Howard-Snyder, D. (eds) 2013. *The Blackwell Companion to the Problem of Evil*, Hoboken, NJ: Wiley-Blackwell.
Stump, E. 1985. 'The Problem of Evil', *Faith and Philosophy*, 2(4), 392–423.
Teehan, J. 2013. 'The Cognitive Bases of the Problem of Evil', *The Monist*, 96(3), 325–48.
Van Inwagen, P. 2006. *The Problem of Evil: The Gifford Lectures Delivered in the University of St Andrews in 2003*, Oxford: Oxford University Press.

---

106 The cosmological argument

I just dropped my cup. That caused it to break. I suspect that reading this book is causing you to have thoughts. There is something obvious about all of this. Things happen and are caused to happen. But what has any of this got to do with the existence of God? As the thought experiment indicates, quite a bit:

Thought experiment

> Stop what you're doing right now. If you followed the instruction, then it caused you to stop. That stopping was caused by some reading. And that reading was also caused – perhaps by your decision to pick up the book. The picking up was also caused. We can carry on in this way for quite some time. In fact, or so it seems, every event that we've come across has been caused. And, if we think it through carefully, we see that it's very plausible that *every* event has a cause. So now backtrack mentally through all of those causes. There must, it would seem, be a *first* cause. And if there is a first cause, it must be a thing that created the Universe – that is, it must be a thing that created space and time. That creative force must then be outside space and time. Such a being must be God.

In crude outline, this is a version of the cosmological argument. As stated, it's not quite an argument, more an invitation to think through some possibilities. But notice that the end point of our musings is that we find ourselves committed to the existence of a God. Of course, there have been a variety of responses to this now very old argument:

- For instance, one might allow that it's perfectly possible for there to be infinitely long causal chains, such that there *is* no first cause.
- Or, if we preferred, and given evidence of a Big Bang, we might deny that a first cause needs to be a personal and creative force. That is, might not the first cause just be a cause in the same way that the blowing of the wind can cause the door to slam?
- Finally, might we not say that there is an obvious error in the line of reasoning given in the thought experiment? After all, do we really need a first *cause* at all? Do we really need to say that absolutely *every* event is caused? We might think that we could say something much weaker – to the effect that although every event we've ever observed had a cause, that need not be taken to entail that all events have causes.

As we might expect with such an old and venerable argument there are replies to these ideas, and further replies to those! To illustrate, let's just take the last point, that perhaps we might make do with one uncaused event – the first event in the created world (whatever that might be – the Big Bang, perhaps). Two thoughts:

- First, the idea of something being *genuinely* uncaused can sometimes seem alien to us. After all, if the preceding is right, then *all* of our experiences are of caused events (the causes of which themselves have causes).
- Second, and ignoring the last point, we might still have an argument here for God's existence. Since *all* of the events that we've seen have been caused, that gives us some good reason to think that a creation event (like the Big Bang) was itself caused. The likely cause of such an event is God.

One point is of central importance, however. When evaluating arguments in the philosophy of religion we must be very careful to avoid simply looking for arguments that confirm our own pre-established beliefs. In doing good philosophy, we should evaluate the arguments on their own terms.

Questions

1 Write out the cosmological argument in premise–conclusion form.
2 Do you think that the idea of an uncaused event makes sense?
3 How does your answer influence your response to the cosmological argument?
4 Do you think it plausible that we have infinitely long causal chains, such that there *is* no event of creation?
5 If there is a first cause, what reasons can you think of for identifying that first cause with God? Would it make sense, for instance, to think that there is a first cause, but that it is not God?
6 Why might it be tempting to say that God is the cause of the Universe, but that God is uncaused?
7 Suppose we say that not every event is caused. How does this bear on the cosmological argument? Can we still develop an argument in favour of the existence of God?
8 Suppose God did not exist. Does it make any practical difference to us?
9 Would the cosmological argument prove that one religion is true, rather than another?
10 Do you think that an atheist and theist should approach this argument differently? Explain your answer.

Reading

Craig, W. L. and Sinclair, J. D. 2009. 'The Kalām Cosmological Argument', in William L. Craig and James P. Moreland (eds), *The Blackwell Companion to Natural Theology*, London: Blackwell, 101–201.

Gale, Richard M. and Pruss, A.R. 2003. 'A New Cosmological Argument', in Gale and Pruss (eds), *The Existence of God*, Burlington, VT: Ashgate, 365–80.

Oppy, G. 2002. 'Arguing about the Kalām Cosmological Argument', *Philo*, 5(1), 34–61.

Russell, B. and Copleston, F. 1964. 'Debate on the Existence of God', in John Hick (ed.), *The Existence of God*, New York: Macmillan, 167–90.

---

107 The design argument

Another alleged proof for the existence of God is given by the design argument – sometimes called the teleological argument. The starting point for this argument

is the idea that the world around us exhibits purpose and, seemingly, that this provides us with some evidence of a designer. To get to the nub of matters, let's consider a very famous thought experiment that we borrow from Paley (with some of our own changes to bring the case up to date):

Thought experiment

> Suppose that you are walking across a stretch of barren countryside, where you can't see anyone or any signs of life. As you walk, you find (to your surprise) a smartphone. Astonished, you turn it on. You examine the phone. It exhibits design. It exhibits purpose. Having turned it on, played with it, and taken it apart, you infer that the phone must have been designed by a creative force. It is simply too incredible to suppose that such a complex object could possibly have come to be merely by chance. Something like a smartphone simply *has* to have been created by an intelligent designer.

What relevance does the thought experiment have for arguments purporting to show the existence of God? Well, consider the Universe around you. Consider the incredible complexity involved in the interactions between sub-atomic particles. Think about the way in which these are bound together by laws and are arranged into structures like people, planets and suns. Think about the ways in which these further objects interact with one another.

In the case of the smartphone it seems overwhelmingly obvious to us that it must have been created. Yet even the least complex smartphone is orders of magnitude less complex that the Universe around us. If we are tempted, in the case of the phone, to infer the existence of a designer, then so too in the case of the Universe we should conclude that we have evidence of a designer. There are many interesting lines of reply:

- First, we might question whether this argument speaks to the existence of a God. Following Hume, we might note that although there is great regularity and order, this doesn't necessary license the conclusion that God exists. All that we need is a very powerful designing agent. That seems some way less than God, we might think.
- Second, we might worry about a regress. Suppose that we agreed that the Universe exhibited sufficient design and order for us to infer the existence of a creator. Presumably, that being would exhibit a good deal of complexity; it would be highly sophisticated, certainly. In that case, we should infer that the designer had a designer . . . and so on?
- Last: here (though many words have been written on this topic), we might worry about whether the uniqueness of the Universe makes a difference to our inferential practices. After all, things like smartphones (and watches – for watches were the subject of Paley's original case) are the kind of thing that are repeated many times over. It's much less clear what we would think if we were crossing a piece of moorland and encountered a very sophisticated object that we'd never encountered before and never encountered again.

Questions

1 Would the designer specified in the design argument have to be anything like the God of Judaeo-Christian religion? Explain your answer.
2 Do you think that smartphones exhibit purpose and regularity?
3 Do you think that the alleged uniqueness of the Universe makes a difference in the case of the design argument?
4 Do you think that, properly considered, the Universe really is unique?
5 Can you think of very complex and sophisticated objects that aren't designed? What difference (if any) do these cases make to the plausibility of the design argument?
6 What does the infinite regress argument (sketched above) show? How does it impact on the cosmological argument?
7 Thinking about biological complexity: how does evolution bear on this issue?
8 Do some parts of the Universe exhibit a lack of order and purpose? What (if anything) does this show about the design argument?
9 What would it mean to say that something happened 'by chance'? How does this differ from the talk of design present here?
10 What do you think is strongest about the design argument?

Reading

Davis, J. L. 1987. 'The Design Argument, Cosmic "Fine-Tuning", and the Anthropic Principle', *The International Journal of Philosophy of Religion*, 22(3), 139–50.

Hacking, I. 1987. 'Inverse Gambler's Fallacy: The Argument from Design', *Mind*, 96(383), 331–40.

Manson, N. (ed.) 2003. *God and Design: The Teleological Argument and Modern Science*, New York: Routledge.

Swinburne, R. 1968. 'The Argument from Design', *Philosophy*, 43(165), 119–212.

---

108 God and time

Discussions of God are frequently very abstract and complex in the philosophical literature. Discussions of time, too, are often very similar. So when it comes to discussing God's relationship to time, things can often get pretty difficult.

A number of points intersect in this debate. Let us first note that God is to be regarded as a maximally perfect being. We must then try to work out what God's relationship is with time:

- Suppose, to start with, that God exists in time, and that He undergoes change – just like us.
- Suppose, second, that God exists in time, but does so *unchangingly*.
- Third, suppose that God exists *outside* time.

The second option doesn't seem so plausible to some. After all, God is described as undergoing change: taking form, speaking to followers, and so on. So might God exist outside time? Perhaps, but it's not quite clear what that would mean. So, finally, might God exist in time? Here is a concern with the idea:

Thought experiment

> Sheila thinks that she has the perfect house. Everything is exactly as she wants it to be. The garden is just so. The decoration is just the way that she likes it. Everything is right! But then, one day, she visits a friend's house. *They* have a feature wall in the living room. Although feature walls had fallen out of fashion for a while, they're on the cusp of returning to fashion. Sheila rushes home and changes her living room so that she too has a feature wall. Once again, Sheila is convinced that she has the perfect house. But now Sheila talks to her friends, whom she has copied. They say to her that her house *was not perfect*. After all, it required change so that it matched *their* perfect house. Their house – unchanging – is perfect, they say, where Sheila's *was* not, but is now.

The thought experiment here is intended as an analogue for the case involving God. Just as Sheila's friends judged her house as imperfect, because it had to undergo a change, so we might view God as imperfect if *He* must undergo a change. After all, if God is *perfect*, then why would He undergo a change? Surely any change to God's nature would make Him either more or less perfect than before. And since God is taken to be maximally perfect, He cannot undergo any change.

This might make it tempting to revisit the other options – perhaps we can come to terms with the slightly mysterious claim that God is to be viewed as outside time. But yet more problems will arise. Suppose that God really is outside time; how, then, can God interact with the world in response to prayer or to bring about miracles? Interacting with physical objects *takes time*. If God is engaged in activities that *take time*, it is very hard to see that He can properly be said to be outside time and time*less*. Something around here will have to go; the question is what?

Questions

1. What might be meant by 'outside time'?
2. Are Sheila's friends right that her house *was* less than perfect?
3. Do you think that a being that undergoes change must thereby become more or less perfect? Use a non-religious example to illustrate.
4. Could God be unchanging, though in time, and yet still perform the tasks commonly attributed to Him?
5. Could God exist 'outside time' and still perform the tasks commonly attributed to Him?
6. Could something go from being *in* time to *outside* time?

7 Suppose that we denied that God is maximally perfect. Would that solve the problem?
8 Could someone who believes in God deny that God is maximally perfect? Would that be a consistent position?
9 Provided we think that God exists, does it matter whether God is in time or timeless?
10 Do you think that we have the resources here for an argument against God's existence? Explain your answer.

Reading

Craig, W. L. 2001. *God, Time and Eternity*, Dordrecht: Kluwer.
Gale, R. 1986. 'Omniscience-Immutability Arguments', *American Philosophical Quarterly*, 23(4), 319–35.
Kretzmann, N. 1966. 'Omniscience and Immutability', *The Journal of Philosophy*, 63(14), 409–21.
Leftow, B. 1991. *Time and Eternity*, Ithaca, NY: Cornell University Press.

---

109 Miracles

A miracle is normally understood to be (something like) a transgression or violation of a law of nature by a theistic being. Our question here is whether or not we could ever have good evidence (that is, reason to believe) that a miracle has occurred. Thus, even supposing that we'd seen one, should we believe it?

This might seem surprising. If I see someone drop a packet of crisps on the street, that gives me reason to think that she did it. Nonetheless, there's something a bit different about the case involving miracles, that stems from the nature of something's being miraculous:

Thought experiment

> One day, Ellis is driving along. He sees Michael carelessly step out into the road without having checked for oncoming traffic. Ellis is of the view that he is travelling much too fast to stop in time. He cannot turn the car off the collision course without either mowing down many innocent pedestrians or careering into a number of cyclists who are outside him. Suddenly, Ellis sees a being that satisfies the description of what God would look like step between Michael and the car. The car stops, instantly, with no harm being done to the car driver. Michael, oblivious and unharmed, carries on his way. Ellis, somewhat stunned, leans out of the driver's-side window and asks 'who are you?' To which the God-like being responds, 'I am God'.
>
> Our question, then, is whether our driver should believe that the have seen a miracle – a transgression of a law of nature.

Here is a reason that he should not. A miracle is, by definition, a violation of the laws of nature. To believe in such a thing, we'd need to have very good evidence. Now, think about what is more likely: that the driver is mistaken about what he saw and wrongly estimates the speed of the vehicle as it approaches Michael; or that God *really has* appeared, and (to boot) has performed a task that is physically impossible.

To get a flavour of the concern: imagine that someone who is very slow (and so very unlikely to win a running race) was reported to you as having won Olympic gold over the 100 metres. You'd be very surprised. You'd be incredulous. You'd want to see some pretty compelling evidence if you were to come to believe this. Indeed, you'd probably want to see more than one piece of evidence.

Yet this case (of an unlikely 100-metre victor) is still many times more probable than the occurrence of a *physically impossible* event. After all, very strange things do happen. What does *not* happen, though, is that we see laws of nature being violated. So, just as we would naturally ask for a higher standard of evidence in cases involving unlikely scenarios, so we must raise the standards of evidence even higher in physically impossible scenarios.

The concern, then, is that the standard of evidence for giving us grounds to believe in the occurrence of miracles is simply so high that we could not ever have grounds for believing that we'd seen one.

Questions

1. If you were Ellis, what would you believe what happened? Who was that mysterious being?
2. If you were Ellis, what do you think that you should believe happened? If there is a difference between your last two answers, explain why.
3. Suppose Michael saw everything unfold and corroborated Ellis's story. Would that make a difference?
4. What if 10,000 people saw everything unfold and corroborated the story. Would *that* make a difference?
5. Find real examples of alleged miracles. Write down five.
6. Do you think that the definition of what it is to be a 'miracle' in a sense 'stacks the deck' against the religious believer?
7. Do you think that the definition of a miracle is right? If not, what would you replace it with?
8. How should we interpret the 'law of nature' in the definition of what it is to be a miracle? Does this make any difference to the case?
9. How much difference (if any) do you think it would make to religious belief as a whole if we were obliged to deny that miracles occurred? Could we, for instance, still maintain a belief in God?
10. Do you think that reports of miracles from many different religions make those individual religions any less plausible? Explain your answer.

Reading

Hájek, A. 2008. 'Are Miracles Chimerical?', in Jonathan Kvanvig (ed.), *Oxford Studies in Philosophy of Religion*, Oxford: Oxford University Press, vol. I, 82–104.

Holder, R. 1998. 'Hume on Miracles: Bayesian Interpretation, Multiple Testimony, and the Existence of God', *British Journal for the Philosophy of Science*, 49(1), 49–65.

Millican, P. 2013. 'Twenty Questions about Hume's "Of Miracles"', in Anthony O'Hear (ed.), *Philosophy and Religion* (Royal Institute of Philosophy Supplement 68), Cambridge: Cambridge University Press, 151–92.

Owen, D. 1987. 'Hume versus Price on Miracles and Prior Probabilities: Testimony and the Bayesian Calculation', *Philosophical Quarterly*, 37(147), 187–202.

---

110 Petitionary prayer

People can pray to God for a number of reasons. But sometimes people *ask* for things. Thus, the parent of a sick child might pray for her swift recovery. An ardent sports fan might pray for their team to win. A sailor might pray for good weather before going out in a storm. Indeed, petitionary prayer seems to be an integral part of at least some religious practices. We might, for instance, ask God for forgiveness. But is the request genuine?

It would be a natural starting point to think that if we are sincerely petitioning God, then we have some reasonable expectation that God listens to and acts upon our request. But is that viable?

Thought experiment

> Anne is talking to her friend. He tells her that God is perfect. Anne agrees with this. She also asks her friend whether or not, if she prays, God will listen and act accordingly. The friend responds that, of course, God is patient, kind and loving and that it is within God's power to behave in such a way as to act on her request. But, Anne responds, wouldn't that involve changing God's mind? After all, He wasn't going to do what I asked for until I asked for it. Just so, says the friend. But in that case, Anne worries, how could God be thought of as perfect? I, a mere mortal, persuaded Him to change His mind.

There are a range of different responses available to the proponent of petitionary prayer. To give just a flavour, here are three:

- First, we might say that, in engaging in the act of petitionary prayer, we are not really petitioning God to change His mind. Rather, in asking for something from God, we are coming into a relationship with Him and changing ourselves in the process. Although we go through the motions of asking questions of God, this is not the same as the sort of requesting that we engage with when talking to our friends and family.

- Second, rather than seeing our asking as a means of changing God's mind, perhaps our asking for some state to be realized, and turning to Him in our time of need, is nothing more than an unfolding of God's plan. On both of these two options, God remains immutable and perfect, and we must reinterpret what is being asked for.
- Third, we might say, instead, that God is not in fact perfect. Perhaps we should say that although God is merciful, loving and kind (as well as very powerful and very knowledgeable), that does not mean that He is perfect. Whilst for some that might mark a significant departure in terms of their conception of God, it would certainly solve the current problem. For if God is not perfect, then it is open to him to be moved by the power of petitionary prayer.

Questions

1 Do you think that it would be rational to ask for something if we knew that there was genuinely no hope of getting it?
2 How does that bear on the issue of petitionary prayer?
3 How significant to faith do you think are acts of petitionary prayer? Are any particular instances of it key to central theological narratives?
4 Is it plausible to treat prayer as a means of coming into a relationship with God, without actually being engaged in genuinely asking for something?
5 Do you think we should see the act of petitionary prayer as an unfolding of God's plan?
6 If we understood petitionary prayer in this way, what (if anything) would this tell us about God's character and nature?
7 Should we respond to this puzzle by denying that God is perfect?
8 If we deny that God is perfect, does this threaten arguments for His existence?
9 Do you think that we *should* petition God?
10 Why might it matter whether we really can petition God?

Reading

Basinger, D. 2004. 'God Does Not Necessarily Respond to Prayer', in Michael L. Peterson (ed.), *Contemporary Debates in Philosophy of Religion*, Malden, MA: Blackwell, 255–64.

Howard-Snyder, Daniel and Howard-Snyder, Frances. 2011. 'The Puzzle of Petitionary Prayer', *European Journal for Philosophy of Religion*, 3, 43–68.

Murray, Michael and Meyers, Kurt. 1994. 'Ask and It Will Be Given to You', *Religious Studies*, 30(3), 311–30.

Stump, E. 1979. 'Petitionary Prayer', *American Philosophical Quarterly*, 16(2), 81–91.

Index