

## Week 1: What is scepticism? Are there any beliefs that are immune to scepticism?

### What is scepticism?

“Scepticism”, “Sceptic”, and “Sceptical” are words commonly used in everyday English; however, they are also technical terms in philosophy, and they have quite a precise meaning.

In **everyday English**, a sceptic is someone who exhibits excessive or extreme doubt about specific claims, or makes a habit of asking for extra justification or evidence. For example, a “climate change sceptic” is someone who (i) doubts scientists’ claims about climate change (e.g. the claim that climate change is happening), and (ii) thinks that more evidence or justification is needed before these claims should be accepted.

In **philosophy**, a sceptic is someone who attacks our claims to **knowledge**. More precisely:

**Scepticism** = the view that it is **impossible** to have knowledge.

Philosophers traditionally distinguish between **global** and **local** scepticism: **local** scepticism is the view that a particular **kind** of knowledge is impossible for humans to obtain, while **global** scepticism says that it is impossible for humans to know **anything at all**. For example, a local sceptic might accept that we know about the contents of our own minds, but deny that we know anything about the external world (the world represented to us by our senses). In fact, **scepticism about the external world** – the view that it is impossible to know anything about the world we inhabit – is the form of local scepticism most commonly discussed by philosophers, and the form of scepticism which most sceptical arguments seek to establish. However, you should also be aware of these other kinds of “local” scepticism:

- Scepticism about **other minds** (the view that we cannot know whether other minds exist);
- Scepticism about **memory** (the view that we cannot know anything based on what we remember);
- Scepticism about the **future** (the view that our experience so far cannot give us knowledge about what will happen next).

It is sometimes argued that **global scepticism is incoherent**, i.e. that what the global sceptic believes (“we can have no knowledge whatsoever”) is somehow irrational or self-defeating. The argument is this: if the global sceptic is claiming to **know** that we can have no knowledge, then he has contradicted himself; but if he doesn’t know that knowledge is impossible, then we have no reason to accept what he says.

There are three responses to this charge. One was famously adopted by Socrates, who claimed to know *only* one thing: namely, that he knew nothing. This “almost-global” scepticism allows the sceptic to go on criticizing everyone else’s claims to knowledge without being accused of inconsistency. An alternative strategy is to say that the argument against global scepticism is mistaken: it is possible to *assert* or *claim* things without thereby claiming to *know* them; we can make claims which we are not completely sure of. Thus the ancient Greek **Carneades** is reputed to have said, c.150BC, “nothing can be known, not even this”. Alternatively, we might take the approach of another ancient Greek sceptical school – **Pyrrhonism** – which recommends simply “withholding assent” from every claim or belief without making any positive claims. (The Pyrrhonists hoped to achieve tranquillity by achieving the “suspension” of all belief in matters of significance.)

Two distinguishing features of philosophical scepticism are:

- Philosophical sceptics generally offer rational **arguments** for their views; you don't get to be a philosophical sceptic simply because you are excessively suspicious or grumpy, as many everyday "sceptics" are.
- Philosophical scepticism is a view about **knowledge**, not about **truth**: philosophical sceptics only want to convince us that our beliefs can't meet the standard needed to count as **knowledge**, not that they are definitely **false**. (This will become important when we look at the sceptics' arguments.)

From now on, when I talk about "scepticism" I'll mean *philosophical* scepticism, not "everyday" scepticism. This is common practice among philosophers, and is a convention you can rely on in your essays as well.

### **Knowledge as a kind of success**

Perhaps the most general claim that can be made about knowledge is that it constitutes a kind of success: to say that someone knows some particular proposition is to say that this person has "got things right" in a certain distinctive way. This means that someone who knows has got to the **truth** of the matter (we can debate this later if you like); but knowledge must be more than merely the possession of some information that just happens to be true. People can believe truths which they don't count as **knowing**, often because – although what they believe is in fact *true* – they don't have good enough reason to believe it. A common example is the optimistic gambler who *always* believes he will win the next bet. Sometimes the gambler does win – and so sometimes his belief that he would win the next bet was *true* – but he doesn't *know* this, because his belief (although true) is irrationally held: it is not based on good **reasons**, but simply on irrational optimism.

Philosophers call a belief **justified** if it is the sort of belief you ought (rationally) to have in the circumstances; justified beliefs are ones which are based on good **reasons** or **evidence** rather than being irrationally held. Someone who has a justified belief is often said to have a **justification** for that belief – which they could in principle explain to someone else.

Another feature which has sometimes been used to distinguish knowledge from simply possession of some information that happens to be true is that knowledge requires you to be totally convinced of what you believe. Thus some philosophers (especially **Descartes**) have suggested that knowledge requires **certainty**: you might believe something which is in fact true, but unless you can be **certain** of it, you don't count as knowing it.

These two putative features of knowledge – that you need to have a **justification** for what you know, and that you need to be **sure** or **certain** of what you know – are very important in understanding how scepticism works.

### **Are there any beliefs that are immune to scepticism?**

Descartes thought he could provide examples of beliefs that he could be entirely certain of, which couldn't be doubted, and which he was entirely justified in believing. Such beliefs could be described as **immune to scepticism** in the sense that we could never be convinced that we don't know them. In particular, you will remember that he claimed that he can know that he exists merely because he knows he is thinking about whether he exists (this is the **cogito** argument). Other beliefs which might be claimed to be immune to scepticism in this way include:

- knowledge of the nature of one's own mental states (e.g. whether you are in pain);
- knowledge of how things appear to you or knowledge of the information delivered by your senses (**sense data**);
- knowledge of *a priori* truths such as simple truths of arithmetic or truths of logic.

It has also been claimed by some philosophers that "common-sense" beliefs are immune to scepticism; beliefs such as "I have two hands" (G. E. **Moore**) and "motor cars do not grow out of the earth" (**Wittgenstein**) are so fundamental to our view of ourselves and our place in the world that they require no further justification and cannot seriously be doubted by any sane person. For that reason it makes no sense to question whether we know such things: to do so would be a form of mental illness rather than a philosophical position! (This view is controversial, and we'll return to it in more detail later on.)

*Homework:* read Descartes, *First Meditation*, and answer these study questions (due next lesson).

1. What does Descartes say he must do “once in his life”?
2. What test does he propose for dealing with his pre-existing beliefs?
3. “The senses sometimes mislead us.” Why then – according to Descartes – should they not be trusted?
4. “Suppose then that I am dreaming...” What does Descartes think remains “certain and indubitable” even if this is true?
5. Descartes ends by considering a third, even more worrying, sceptical scenario. What is it, and what does it force him to doubt?

## Week 2: Arguments for scepticism

### Sceptical Scenarios

Last week we said that sceptics try to persuade us that we cannot **know** what we think we know. Importantly, they don't do this by trying to convince us that our beliefs are **false**, as this can often be a way of *gaining* knowledge: if I have good reason for thinking that my belief that  $p$  is false, this can result in me coming to know something else, namely the negation of my original belief: not- $p$ . Instead, the sceptic's aim is to establish a certain kind of *lack* of certainty – a state where we can neither claim to know that our beliefs are true, nor claim to know that our beliefs are false.

Sceptics generally proceed by suggesting a **sceptical scenario** which, as far as we know, might **actually be the case**. Here are a few of them:

- You are dreaming right now; the real world is very different from what you are experiencing in your dream.
- You are really just a brain in a vat in the lab of a mad scientist, who is using a computer to feed you sense-experience to trick you into thinking you are at school in 2012.
- Your mind has been taken over by a malicious demon who is controlling your experiences (and possibly your thoughts as well).
- Five seconds ago your memories were wiped and replaced with different ones by a mad scientist. Everything you believe about the past is false.
- In five seconds, the laws of physics will change completely.

These hypotheses are used to motivate different kinds of local scepticism: the first three motivate **scepticism about the external world**, since they suggest that we could be entirely mistaken about what is going on in the world outside our heads. The hypothesis about wiped memories motivates **scepticism about memory**, because it suggests that we cannot rely on our memory to tell us what has happened so far, while the last hypothesis (about the laws of physics) suggests that we cannot be sure of any of our predictions about how the world will behave five seconds from now, motivating **scepticism about the future**. (This is sometimes also called **scepticism about induction**, because it suggests that we cannot rely on **inductive arguments** – arguments which predict future events as a result of **generalizing** from what has happened already.)

The most important thing to understand about these sceptical scenarios is that **the sceptic does not need to persuade you that they are really true**. All that has to happen for a sceptical argument to get going is that we need to admit that any of them **could** (as far as we know) be the case. In what follows I'll concentrate on the arguments for scepticism about the external world, but much of what I say could also be applied to scepticism about memory or induction.

### Scepticism and infallibilism

These sceptical scenarios provide one very direct way of establishing that we cannot know what we think we know. At the start of the *Meditations*, Descartes suggests that

“the slightest ground for doubt I find in any [of my beliefs] will suffice for me to reject all of them.”

It seems that Descartes believed that genuine knowledge (“*scientia*” in his original Latin) can be had only when we have eliminated every possibility of error. But if, as far as I know, one of the sceptical

hypotheses might be the case (e.g. for all I know I might be a brain in a vat), then surely there is some reason for doubting the evidence of my senses – some possibility that I might have made a mistake – in which case, I don't know anything about the external world. I might think that I am sitting here in a classroom; but as long as there is the slightest logical possibility that my belief is mistaken – the slightest possibility that I might be a brain in a vat instead – I cannot *know* that I am sitting here in a classroom.

**Important:** this sceptical conclusion (that I do not know that I am sitting in a classroom) also holds even if my belief is **true**: in that case, the sceptic would say, I have a true belief that I am sitting in a classroom, but because my belief is not **infallible** (I could have been mistaken) I cannot describe that belief as something I **know**.

**Also important:** Descartes was not a sceptic! He explicitly says that he raises these sceptical doubts only so that he can find a way to eradicate them – although many philosophers now believe that Descartes' attempt to answer the problem of scepticism was not a success.

This view about knowledge – the view that you can only know something if it is **impossible** for you to have made a mistake about it – is known as **infallibilism**. It can also be described as the view that you can know only what is **logically certain**.

How could we respond to this kind of argument for scepticism? The argument says that I can't **know** anything based on my sense-experience unless I have eliminated every possibility that I might be mistaken, and eliminated all grounds for doubting my experience. Since I can't rule out the possibility that I might be a brain in a vat, it looks like I *can't* eliminate all grounds for doubting my sense-experience, and so can't know anything on the basis of my senses. But should we accept the "eliminate all doubt" principle? Descartes thought that knowledge requires *certainty* – that you could know only what it is impossible or irrational to doubt. But perhaps that is too strong a requirement to impose on knowledge. Think about everyday cases where we accept that people know things:

Bill studies hard for his science test, and gets all the questions right. We'd usually say that Bill "knows" the answers. But is Bill *certain* of the answers he gave? Probably not – after all, he cannot rule out the possibility that he might have mis-remembered something (although in fact he didn't). So Bill knows the answers, but he isn't certain of them. So knowledge doesn't require certainty.

This imaginary example, or "thought-experiment", seems to show that we can know something even if there are some reasons for doubting it, and some possibility that we might have made a mistake; therefore, we might argue, providing reasons for doubting what we believe is not enough to show that we don't know what we think we know. In fact, this is not a solitary example; in everyday life, we use the verb "know" in circumstances where we just have **good reason** for belief, not **certainty**. An "ordinary language philosopher" would argue that it makes no sense to propose a theory of knowledge which ignores the way "know" is normally used in our language.

Moreover, it has been suggested that the appeal of infallibilism derives from a logical mistake: we shouldn't argue from "it is impossible that, if I know that **p**, I am mistaken about **p**" (knowledge entails truth) to "if I know that **p**, it is impossible that I could ever have made a mistake about **p**" (infallibilism).

Perhaps the sceptic has a response: the philosopher Peter **Unger** has argued that knowledge really *does* require certainty; as a consequence he claims that most of the time we are making a mistake when we claim to know something we are not certain of. His reason is that it feels like a contradiction to say something like

“I know it but I’m not certain of it.”

According to Unger, the reason this feels like a contradiction is that, in saying “I know it”, we also imply “I am certain of it” – so what we’re really saying here is “I’m certain of it, but I’m not certain of it”, which really is a contradiction. Similarly, David Lewis suggests that

“If you claim that **S** knows that **p**, and yet you grant that **S** cannot eliminate a certain possibility in which *not-p*, it certainly seems as if you have granted that **S** does not after all know that **p**. To speak of fallible knowledge, of knowledge despite uneliminated possibilities of error, just sounds contradictory.”

Lewis, “Elusive Knowledge”, p.549

### **Scepticism and Justification**

However, sceptics do not have to rely on **infallibilism** to make their case. An alternative sceptical approach is to attack our claim to justification, arguing that – whether or not our beliefs are *true* – they are ultimately **unjustified**. As we said before, a belief which does not have an adequate justification – one which we do not have good enough **reason** for believing – does not count as knowledge. This kind of sceptic often points out an important feature of sceptical scenarios such as being a **Brain in a Vat**: that they are **qualitatively indistinguishable** from real life; there is absolutely no way that I can find out from examining the nature of my experiences whether they are genuine and “**veridical**”, or misleading hallucinations fed to me by the mad scientist. This fact about my experiences – that they would seem exactly the same to me even if I were a brain in a vat – means that any belief formed on the basis of those experiences must be ultimately **unjustified** and therefore cannot be knowledge. If my beliefs about my surroundings are true, the fact that they are true is the result of my good fortune in not being a brain in a vat, rather than the result of my using my abilities in a **reliable** way.

An example might make this clearer: suppose I am asked whether a tree in the park is an ash or an elm. I can’t tell the difference between these trees reliably, but I decide to say that the tree is an elm. I get this right – but *because* I can’t reliably tell the difference my answer is merely a lucky guess rather than genuine knowledge. Similarly, suppose I am asked whether the sun is shining outside. My visual experience seems to suggest that it *is* sunny outside, but I can’t tell the difference between having an experience of genuine sunshine, and having a fake experience of sunshine fed to my envatted brain in a lab on a rainy winter’s day. So – even if I get it right and it really *is* sunny outside – my answer is merely a lucky guess rather than genuine knowledge. The possibility that I might be a brain in a vat suggests that any true belief I form about my environment is true because a piece of **luck**, and beliefs that are true because of luck are not justified and cannot count as knowledge.

Another way of making the point is this: although we can know what the content of our sense-experience is – how things *seem* to us – it is a further **inference** to conclude that these sense-experiences adequately reflect reality, and this inference is **unjustified**. The only way we

could use our sense-experiences to gain knowledge of reality would be if we were justified in making the inference from “this is how things seem” to “this is how things are”, and we could be justified in doing that *only* if we could rule out the possibility that a sceptical scenario obtains. Unfortunately, there does not seem to be any easy way to rule out this possibility. So our beliefs about the external world are unjustified and cannot count as knowledge.

Film: *The Matrix*

*Aside* : the surprising thing about the Matrix is that Neo (Keanu Reeves) completely misses the philosophical lesson he should learn from his experience. Despite discovering that all his experience up to that point has been misleading and unreliable, fed to him by a malicious computer program, he never doubts the veracity (truthfulness) of his experiences *after* he leaves the simulation, and indeed seems to have complete certainty that his new world of fighting robots, submarines and Carrie-Ann Moss in a catsuit is in fact the *real* world. The lesson he *should* take from his experiences is that *no* consistent set of experiences is entirely trustworthy – even the ones “after” he leaves the Matrix. In that respect, the film *Inception* displays slightly more philosophical sophistication.

*Homework*: “Assess the claim that there can be no knowledge of what lies beyond sense experience.” (June 2011)



## Extension material: the Closure Principle

*(Don't worry about this section unless you feel like engaging with some degree-level contemporary philosophy; this material is not essential for the exam!)*

The modern debate about scepticism tries to make the steps in the sceptic's argument clearer. Often the sceptic is represented as arguing like this:

- (i) I don't know that I'm not a brain in a vat;
- (ii) If I don't know that I'm not a brain in a vat, I don't know anything about the outside world;
- (iii) therefore, I don't know anything about the outside world.

Let's look in detail at how steps (i) and (ii) work.

### (i) I don't know that I'm not a brain in a vat.

This seems plausible: after all, what reason could you have for thinking that you **aren't** a brain in a vat? It's not something you could check up on by experimenting on yourself, since the "results" of whatever you try to "do" are simply fed to your brain by a computer controlled by the mad scientist. If that common-sense argument doesn't work for you, try these alternatives:

I don't know that I'm not a brain in a vat because...

- ... I **can't tell the difference** between being a brain in a vat and not being a brain in a vat. Things would seem the same to me whichever scenario was the case – so if my belief that I'm not a brain in a vat happens to be true, it's no more than a **lucky guess**, and we know that lucky guesses don't count as knowledge.
- ... I would still believe that I'm not a brain in a vat even if I **were** a brain in a vat. So my belief that I'm not a brain in a vat does not "**track the truth**" and so does not count as knowledge.

### (ii) If I don't know that I'm not a brain in a vat, I don't know anything about the outside world.

This premise (stage in the reasoning) is harder to defend, and in fact many philosophers think that the best way to deal with scepticism is to reject it. First, let's try a common-sense defence:

"If you can't be sure you're not a brain in a vat, you can't be sure that the world you seem to be perceiving really exists; but if you can't be sure of *that*, you can't be sure of any specific claim to knowledge, e.g. "I know I have hands."

Philosophers who want a more technical explanation talk about something called the "closure principle". There are various ways of formulating a closure principle, but this is the simplest:

**(CP)** If you know that **q** is a necessary precondition of **p**, and you don't know **q**, then you don't know **p** either.

Try this with **p** = "I have hands" and **q** = "I am not a brain in a vat". Obviously **q** is a necessary precondition of **p**, since I can't have hands **unless** I am not a brain in a vat. So by the closure principle **(CP)**, I can't know **p** (I have hands) unless I know **q** (I am not a brain in a vat).

The debate here focuses on whether we should accept **(CP)** in the context of a sceptical argument. Philosophers often argue this point by giving quite mundane examples where **(CP)** seems to work, or doesn't seem to work. For example, here's an example that supports **(CP)**:

“The cake cannot be cooked unless the oven has been turned on. So unless I know that the oven has been turned on, I don't know that the cake is cooked.”

Here's an example that pushes against **(CP)**. Suppose, when walking round the zoo, you reason like this:

“Those animals are zebras only if they're not aliens in disguise. But it's not true that, unless I know that those animals are not aliens in disguise, I don't know that they are zebras.”

Here, it seems that **(CP)** fails: I don't have to rule out the possibility of shape-changing aliens to know that those animals are zebras – provided, of course, that those animals really *are* zebras (knowledge implies truth).

## Week 3: Responses to scepticism

### Descartes on Scepticism about the external world

Descartes' view is simply that the existence of the external world is guaranteed by the existence of a morally good God. Although we cannot know that physical objects ("corporeal things") exist, simply on the basis of the evidence of our senses, we *can* know that they exist once we know that a perfect God exists – something which Descartes claims to have proved. If our sense-experiences were entirely misleading, it would mean that we could do everything we could to check up on our beliefs about the external world and yet still be radically mistaken about what exists – something which, Descartes claims, would make God a "deceiver".

"In the first place, I recognize that it is impossible that He [God] should ever deceive me, since in all fraud and deceit is to be found a certain imperfection... the desire to deceive bears evidence without doubt of weakness or malice, and, accordingly, cannot be found in God.

Secondly, I am aware in myself a certain power of judgement, which undoubtedly I have received from God... and as he would not wish to deceive me, it is certain that he has not given to me a power such that I can ever be in error, if I use it properly."

*Fourth Meditation*

"Since [God]... has given me ... a very strong inclination to believe that [sense experiences] are sent to me or derive from corporeal things, I do not see how he could be excused of deception if in truth these ideas came from or were produced by causes other than corporeal things. And accordingly one must confess that corporeal things exist."

*Sixth Meditation*

### Hume and Mitigated Scepticism

Hume's view is, in a way, a capitulation to scepticism: he accepts that sceptical arguments do succeed in showing that we cannot know that our sense-experiences are caused by continuously existing physical objects (which he calls "external bodies"). However, he attempts to "mitigate" this scepticism (make it less harsh) by arguing that human nature makes it impossible for us to withhold assent from all our beliefs and to stop "judging" that things are a certain way; he also recommends "carelessness and inattention" as a remedy which will enable us to continue believing that the external world exists. However, Hume's view faces a serious problem: he seems to accept that these beliefs about the external world which nature "compels" us to maintain are in fact unjustified and irrational. But how can it be a good thing to retain beliefs that are unjustified and irrational? And why should we be pleased if nature compels us to retain them?

One possible response is this: Hume has shown us that **scepticism is not a practical option** – i.e. he has pointed out that sceptical doubts are impossible to sustain, simply because human nature compels us to get on with our lives to survive, and to do so requires us to maintain beliefs about the external world which may be highly questionable, but which nevertheless are essential if we are to stay alive. (A plausible modern adaptation of Hume's views about "nature" might operate in terms of the theory of **evolution**: we might try to account for our seemingly unshakeable belief in the external world by showing the kind of evolutionary advantages that belief has. Of course, this does not show that such a belief is *justified* in the sense disputed by the sceptic.)

From **Hume**, *A Treatise of Human Nature* (1739):

“Should it here be asked me, whether I sincerely assent to this argument, which I seem to take such pains to inculcate, and whether I be really one of those sceptics, who hold that all is uncertain, and that our judgment is not in any thing possest of any measures of truth and falsehood; I should reply, that this question is entirely superfluous, and that neither I, nor any other person was ever sincerely and constantly of that opinion. **Nature, by an absolute and uncontrollable necessity has determined us to judge as well as to breathe and feel;** nor can we any more forbear viewing certain objects in a stronger and fuller light, upon account of their customary connexion with a present impression, than we can hinder ourselves from thinking as long, as we are awake, or seeing the surrounding bodies, when we turn our eyes towards them in broad sunshine.

... the opinion of the continued existence of body depends on the **COHERENCE**, and **CONSTANCY** of certain impressions ... thus I am naturally led to regard the world, as something real and durable, and as preserving its existence, even when it is no longer present to my perception.

But whatever force we may ascribe to this principle, I am afraid it is too weak to support alone so vast an edifice, as is that of the continued existence of all external bodies.

This sceptical doubt, both with respect to reason and the senses, is **a malady, which can never be radically cured, but must return upon us every moment**, however we may chase it away, and sometimes may seem entirely free from it. It is **impossible** upon any system to defend either our understanding or senses; and we but expose them farther when we endeavour to justify them in that manner. As the sceptical doubt arises naturally from a profound and intense reflection on those subjects, it always increases, the farther we carry our reflections, whether in opposition or conformity to it. **Carelessness and inattention alone can afford us any remedy.** For this reason I rely entirely upon them; and take it for granted, whatever may be the reader’s opinion at this present moment, that an hour hence he will be persuaded there is both an external and internal world.

### **Phenomenalism**

You will remember that the sceptic appeals to the **gap between appearance and reality** – i.e. the sceptic points out that we cannot confidently **infer** the existence of a “real” physical world simply from the “appearances” provided to us in the form of sense-experience. One possible response to this line of thought it simply to deny that there is any such gap – to say that statements apparently about a reality composed of physical objects, are really statements about my actual and possible sense-experience. This view is known as **phenomenalism** and was defended by **A. J. Ayer** among others (although he later abandoned it). To see how it works, consider a specific example: if I make the statement “there is a brown table in this room”, that is not – as it seems to be – a statement about a physical object – but instead must be understood as a claim about the kind of sense-experiences I am having now and could have in the future (e.g. that my experiences are brown-ish and square). So the sceptic is wrong to say that I do not know that there is a table in the room, because *what* I know when I know this is simply some facts about the contents of my own experiences – and most sceptics would agree that we can know facts about the content of our experience, even if they deny that we can know about the ultimate nature of reality.

Although it was immensely popular during the middle of the 20<sup>th</sup> century, phenomenism now has few defenders, owing to at least three major problems. First, it seems impossible to “translate” statements about material objects into statements which mention only sense-experience, as the phenomenalist advises. We cannot say “I am having an experience which looks like a table”, because we are still using a word – “table” – that refers to a physical object rather than a sense-experience. Second, the phenomenalist cannot explain the *consistency* and *regularity* of our experiences. Normally we would say that our experiences follow regular, predictable patterns *because* they are caused by physical objects which continue to exist even when unperceived (for example, if we leave the room and come back in again). But the phenomenalist wants to say that, strictly speaking, there are only actual and possible sense-experiences; how then can we explain why these experiences follow regular patterns?

Finally, there is the problem that phenomenism apparently leads to **solipsism** – the view that only I (and my experiences) exist. The phenomenalist says that all statements about the external world are to be translated into statements about my actual or possible sense-experience. That means that statements about *other people* should also be understood in terms of my actual or possible sense-experience. But then the other people around me are not really *people* like I am, with internal mental processes and experiences of their own; instead they are merely **logical constructs** from experiences I have. This seems like an intolerable conclusion.

### **Transcendental arguments against scepticism**

A **transcendental argument** tries to show that scepticism is mistaken because the very things the sceptic doubts are a precondition of raising meaningful sceptical doubts in the first place. How that works will become clearer through two examples.

First, there is an argument put forward by Gilbert **Ryle**: to be able to apply a term like “mistaken” or “misleading”, or even “false” is to presuppose that something is capable of being “accurate” or “true”, just as talking of “counterfeit” coins presupposes that “real” coins exist:

“In a country where there is a coinage, false coins can be manufactured and passed... An ordinary citizen... might become suspicious of the genuineness of any particular coin that he received. But however general his suspicions might be, there remains one proposition which he cannot entertain... that all coins are counterfeits. For there must be an answer to the question ‘Counterfeits of what?’”  
Gilbert Ryle, *Dilemmas*, pp. 94-95

Similarly, Ryle claims, it makes no sense for a sceptic to suggest that *all* our experiences might be misleading (for example if we were an envatted brain being fed a computer simulation). To describe some experiences as misleading presupposes that there can be experiences which are *not* misleading (“veridical”); therefore it follows that what the sceptic suggests (that *all* of our experiences might be misleading) is in fact an impossibility.

Ryle’s argument has been heavily criticized: first, the analogy with coins suggests merely that there must be *some* veridical experience in the world, not that I can be sure that some of *my* experiences must be accurate. Even if I am convinced that there is some real coinage somewhere in the country, I can come to believe that all the coinage I personally own is counterfeit. The same is true of

experiences – I might be sure that there is *some* veridical experience out there while thinking that all of *my* experience is misleading.

Moreover, it seems more reasonable to say that using terms such as “misleading” and “inaccurate” presupposes only that it is a **conceptual possibility** that there is some veridical experience in the world, not that veridical experience must actually exist. The analogy with coins backs this up: suppose in Ryle’s imagined country forgery was rife, and the forgers were removing genuine coins from circulation at the same time as introducing forgeries. Eventually there would be no real coinage left, and all the actual coins would be fakes. All that would remain would be the conceptual possibility of a real coin being brought into existence, and this would be enough to enable us to understand that all the existing coins are fakes. Similarly, it seems like we can make sense of the claim that *all* the actual experiences in the world are misleading, as long as we can understand the conceptual possibility of someone having a veridical experience.

Another kind of **transcendental argument** is suggested by Hilary **Putnam**. He suggests that the mere fact that the sentence “I am a brain in a vat” is *meaningful* suggests that the sentence must be *false*. So a precondition of the intelligibility of this sceptical scenario is that the sceptical scenario does not actually obtain. The major premise of Putnam’s argument is this:

“One cannot refer to certain kinds of things... if one has no causal interaction at all with them”

Putnam, *Brains in Vats*, p.489

Suppose that I am a brain in a vat. Then all my sense-experience is the result of a complex computer simulation, in which case I have no causal interaction with real things in the physical world, and my words cannot refer to them. My words can refer only to what I have causal interaction with – namely the misleading experiences being fed to me by the computer. So if I make the sounds “I am a brain in a vat”, I am not really saying something containing words that refer to brains and vats:

“[the word] ‘vat’ refers to vats in the image... but certainly not to real vats, since the use of ‘vat’ in vat-English has no causal connection to real vats.” (*op. cit.*, p.488)

So anyone who is capable of using the sentence “I am a brain in a vat” to mean what we think it means has thereby proved that it is false: for if she were a brain in a vat, she would not be able to use these words with their standard English meanings. Of course, if she really is a brain in a vat, speaking vat-English, then her words do not mean the same as they would in normal English, given that they refer only to the images she is fed; indeed, in such situations we might doubt whether her words have meanings at all. Thus, Putnam claims, the claim that I am a brain in a vat is self-refuting.

*Exercise: what, if anything, is wrong with Putnam’s argument?*

### Anti-sceptical appeals to ordinary language and ‘common sense’

One suggestion – which we have already met – is that we can only understand the meaning of a word like “know” by examining how it is used in ordinary language; and since in ordinary language we use “know” in a way that presupposes that we *do* know things about the external world, we are justified in rejecting any sceptical argument that threatens to show that knowledge is impossible: any account according to which knowledge is impossible just isn’t an account of the kind of “knowledge” we talk about in ordinary language, since *that* kind of knowledge is something that human beings can have.

However, sceptics have a response to this: they can claim that they are, after all, talking about the “ordinary” kind of knowledge – all they are doing is drawing out the consequences *implicit* in the everyday concept of knowledge (e.g. that knowledge requires justification) and revealing that – once the concept of knowledge is fully understood – it follows that knowledge is impossible. After all, people are capable of using a concept every day without fully understanding it, so it should be no surprise if a close examination of a complex concept such as knowledge reveals that many people are mis-using it, claiming to have knowledge where none is possible.

A more direct appeal to common sense was made by G. E. **Moore** in the first half of the 20<sup>th</sup> century. Moore claimed that he could prove the existence of the external world by starting from premises which were known through common sense. He held up one hand, saying “Here is a hand”, and then held up another, saying “Here is another.” Since – Moore claimed – he *knew* that what he was saying was true, and since he *knew* that, if there was a hand here then it followed logically that there must exist at least one physical object (i.e. the hand itself), he could conclude that he knew that an external world exists. Indeed, Moore suggests that these “common sense” premises are immune to serious doubt:

“I *knew* that there was one hand in the place indicated... and that there was another in the different place indicated ... How absurd it would be to suggest that I did not know it, but merely believed it, and that perhaps it was not the case! You might as well suggest that I do not know that I am now standing up and talking.”  
Moore, *Proof of an External World*, p.144

Moore’s strategy is sometimes dismissed as inept or foolish: after all, he has apparently “begged the question” against the sceptic by presupposing exactly what the sceptic questions, namely whether there are any physical objects in the first place. However, the reasoning is more complex than that. Moore is using a logical strategy known as **reversing the argument**: while the sceptic argues like this

- (1) I don’t know that there is an external world; therefore
- (2) I don’t know that I have hands

Moore argues the reverse:

- (1) I know that I have hands; therefore
- (2) I know that there is an external world.

Both arguments are equally valid, so we have to *decide* which one to accept. Moore's point is that his version of the argument starts with a premise we are really very sure of, but the sceptics argument starts with a premise which is extremely controversial. But of course, given a choice between two arguments, we ought to accept the one with the most reasonable premises – namely, Moore's argument.

Further viewing: <http://tinyurl.com/MoorePEW><http://tinyurl.com/MoorePEW>

### **Internalism and Externalism about knowledge**

**Internalists** are philosophers who hold that the difference between knowledge and mere true belief depends only on what is going on “inside your head” – for example, whether you possess good reasons for what you believe, or can explain your justification when challenged. **Externalists** deny that internalism is true, claiming that the difference between knowledge and true belief depends on whether the subject is “hooked up to the outside world” in the right way. For example, an externalist might say that whether our beliefs about the world count as knowledge depends only on whether they are formed using a **reliable method**, not on whether we can give an intellectual justification of those beliefs. So – according to the externalist – the fact that we cannot rationally *prove* that we are not brains in vats, and cannot prove that our senses are not deceiving us, does not prevent us from gaining knowledge about the world: as long as our senses really *are* reliable we can use them to gain knowledge, regardless of whether we can *prove* that they are reliable, and regardless of whether we even have any *reason* for believing that they are reliable. Thus the externalist says that our true beliefs about the world can count as knowledge even if we cannot give a full rational justification of those beliefs.

Notice that the externalist argues that our ability to have knowledge of the world does not depend on whether we can *prove* we are not brains in vats; however the externalist does concede that we can have knowledge about the world *only if* we are in fact not brains in vats. This is not surprising: for one thing, if we really were brains in vats, then our beliefs about the external world would be false and so could not count as knowledge anyway; for another, we should remember that the sceptic promised to show that knowledge of the external world is *impossible* by showing that we cannot know anything about the external world *even if* we aren't brains in vats and our beliefs are true, on the basis that our beliefs about the world cannot be fully justified even in the situation where those beliefs are true. We'll return to externalism again in coming weeks.

### **Epistemic Contextualism**

**Contextualists** claim that the meaning of the word “know” is different in different contexts of use; as a result, what would count as an appropriate and adequately-supported claim to knowledge in one context (e.g. everyday life) might not be appropriate (or even true) in another context – for example, in a philosophy seminar. In the early 1990s Michael **Williams** developed this into a way of responding to scepticism: he claims that, although the sceptic is right to say that we cannot “know” facts about the external world in *some* contexts – i.e. when engaged in philosophical reflection – we should not agree that knowledge is impossible in other, everyday contexts, when sceptical doubts are not at issue.

“The sceptic takes himself to have discovered, under the conditions of philosophical



reflection, that knowledge of the world is impossible. But in fact, the most he has discovered is that *knowledge of the world is impossible under the conditions of philosophical reflection.*"

Williams, *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*, p.

130

This view has two major advantages: first, it recognizes the force of sceptical arguments, agreeing with the sceptic that we can no longer claim to know what we think we know when we are considering sceptical arguments in our philosophy class. However, it does not force us to agree with the sceptic's conclusion that knowledge of the external world is *impossible*, because the contextualist can say this: it does not follow, from the fact that knowledge is impossible in *one* context (philosophical discussion) that knowledge is impossible in *every* context. So the contextualist view has a second major advantage: it shows that knowledge is possible (as long as we are not thinking too philosophically at the time!)

However, a problem for this view is that Williams also needs to claim that the kind of "knowledge" that is at issue in the context of philosophical discussions of scepticism is no *better* than the kind of knowledge which we can have in everyday contexts; it is important for his view that we are not "missing out" on anything by being unable to claim to know – in the strict, philosophical sense – that we are not brains in vats. That seems highly questionable. Moreover, we might worry that the view concedes too much to the sceptic: what we are worried about is whether we can "really" know that the external world exists, and this view concedes that we can only know that the external world exists in "everyday" contexts where scepticism isn't under consideration. The conclusion that we can know that the external world exists only when we are not considering the problem of scepticism seems a bit too much like ignoring the problem in the hope that it will go away.

### **Wittgenstein's *On Certainty***

During the last two years of his life, Wittgenstein made a series of notes developing a response to the problem of scepticism; some years after his death in 1951 these were collected and published as the book *On Certainty*. His views are complex, but worth considering in detail. His basic idea is that **the starting point for sceptical arguments is unintelligible** – i.e. that what the sceptic proposes is somehow incoherent, and so cannot be accepted.

In the second half of his philosophical career, Wittgenstein often appealed to the idea that our words get their meaning through the ways in which we use them. Thus the meaning of a word like "pain" comes from the variety of ways in which that word could be used appropriately within our society. Wittgenstein used the term **language-game** to refer to a shared set of language-using practices of this kind. (The use of "game" here is primarily to emphasize that the use of words follows shared *rules*.) Within each society there might be many different "language-games" relating to different kinds of activity, and it is crucially important to remember that a word or sentence which has a meaning within one language-game might have a different meaning – or even *lack* a meaning – within a different language-game. The reason a word could lack a meaning within a language-game is that meaning is determined by use, and if there is no agreement about how the word should be used in the context of that particular language-game, then there is no such thing as "the meaning" of the word.

Wittgenstein uses this overall view of language to develop a reply to the sceptic: he says that the kind of beliefs which the sceptic is trying to convince us that we do not know – fundamental propositions such as “there are physical objects”, “I have two hands” or “the world was not created 150 years ago” – are such that it makes no sense to claim to know them – but also makes no sense to claim not to know them, and no sense to doubt them. His reason is that our language-games have rules for the appropriate use of “know” and “doubt” only in certain situations, namely when one is in a position to offer rational support for that knowledge or doubt. Since – according to Wittgenstein – there is nothing that could count as rational support either for doubting such propositions *or* for claiming to know them, it follows that such claims about knowledge have no rules for their use, and so are literally meaningless, or at least deeply inappropriate:

“ ‘I know that I am a human being.’ In order to see how unclear the sense of this proposition is, consider its negation. At most it might be taken to mean ‘I know I have the organs of a human’... But what about such a proposition as ‘I know I have a brain’? Can I doubt it? Grounds for *doubt* are lacking!”

Wittgenstein, *On Certainty* §4

“Now, can one enumerate what one knows...? Straight off like that, I believe not. - For otherwise the expression ‘I know’ gets misused.” (§6)

“We just do not see how very specialized the use of ‘I know’ is.” (§11)

“One says “I know” when one is ready to give compelling grounds. “I know” relates to a possibility of demonstrating the truth. [...] But if what he believes is of such a kind that the grounds that he can give are no surer than the assertion, then he cannot say that he knows what he believes.” (§243)

However, Wittgenstein’s view is controversial. The fact that we find it hard to find situations in everyday life where it is appropriate to assert things like “I know I have hands” does not immediately prove that such claims are meaningless and cannot be accepted. An alternative theory of meaning is that sentences get their meaning not from the context of their use, but rather from the meanings of their constituent words and the way they are arranged. This **compositional theory of meaning** suggests that any sentence is meaningful if it is formed from meaningful words put together in a meaningful way - regardless of whether we can imagine using that sentence in everyday life. So the mere fact that we cannot imagine a situation when we would have a *use* for “I know I have hands” in everyday life does not demonstrate that this sentence does not express a proposition that can be debated meaningfully.

*Extension Material:* Duncan Pritchard, *Wittgenstein on Scepticism*. (Available online at <http://www.philosophy.stir.ac.uk/staff/duncan-pritchard/documents/WittOnScepticism.pdf>)

*Homework:* “The fact that our beliefs are subject to error shows that knowledge is impossible. Discuss” (June 2009)

## Week 4: Belief and Justification

### Propositional Attitudes

You will remember from last year that philosophers are especially interested in **propositional knowledge**: this is knowledge *that* something is the case, and is usually distinguished carefully from “know-how”, which can be a kind of practical ability which you need not be able to explain in words, and knowledge by acquaintance, which is a kind of familiarity with a person or place (“Do you know Paris well?”). When sceptics claim that knowledge is impossible, or philosophers enquire about the correct definition of knowledge, it is propositional knowledge that they are primarily interested in.

Propositional knowledge gets its name because it is usually conceived as a relationship between the knowing **subject** and a **proposition**. The proposition is the **content** of your knowledge; it is what is expressed by the words following “that...” in sentences of the form “S knows that ...”. The standard view is that propositions are **expressed** by sentences, but are not the same as sentences, as two different sentences can express the same proposition – for example if they are saying the same thing in different languages.

### The dual-component view of belief

Knowledge is just one kind of propositional attitude: others include believing, hoping, doubting, and considering. The view that belief divides into two separate elements – the proposition that is the content of the belief, on the one hand, and the attitude we take to that belief, on the other – is known as the **dual-component view of belief**. It claims that believing that something is the case is distinguished from other kinds of propositional attitudes by the precise nature of the attitude we take towards that proposition. This view faces two problems: first, whether every belief has to consist in a relationship to one specifiable proposition; and second, how precisely we should specify the nature of our attitude to the proposition concerned.

First, we might wonder whether every belief has to have a content which can be specified by a specific “that-” clause. Couldn’t there be a case where I believed something without quite being able to work out how it should be put into words? Moreover, we often quite happily attribute beliefs to animals – e.g. dogs and cats – which lack the intellectual ability to understand a structured proposition. For example, I might say that my cats know that there is food in a certain room in the house, even though they almost certainly lack the intellectual ability to grasp a structured proposition such as “there is food in this room”. This problem leads some people to adopt **instrumentalism about belief**, as we’ll see below.

Second, how do we explain what kind of attitude to a proposition we take in believing it? Believing that something is the case is a continuous state, so it cannot be explained in terms of a momentary event in the mind, for example a moment at which I consciously approve the belief then file it away for later use. David **Hume** claimed that belief is

“a more vivid, lively, forcible, firm, steady conception of an object, than what the imagination alone is ever able to attain.”

However, it does not seem that this can be right: our beliefs cannot always be “vivid and lively”, for the simple reason that we do not always have our beliefs directly before our consciousness. I have

believed that Paris is the capital of France for many years, but this is something I rarely consider consciously in such a way that the belief might deserve the descriptions “vivid” and “lively”, and in fact even when I do consider this belief it does not seem to come with the kind of experience attached that Hume suggests it should: in fact, although I am aware *that* I believe that Paris is the capital of France, there is no distinctive kind of conscious experience that I associate with that belief – nothing it *feels like* to be someone who is holding that belief.

*Exercise:* you may disagree with that last point. If it’s wrong, *why* is it wrong?

A more promising account of the “attitude” involved in belief comes via the connection between belief and *truth*: to believe that something is the case is to believe that something is true; it seems paradoxical to claim something like

“It’s true that it’s raining but I don’t believe that it’s raining.”<sup>1</sup>

Nevertheless, the explanation of belief in terms of truth can be challenged: for one thing, it is circular, presupposing a prior understanding of the term it sets out to explain. If I tell you that “believing that *p*” is the same as “believing that *p* is true”, I have explained something about belief, but I haven’t managed to tell you what belief *is*, since “believing that” recurs in my proposed explanation, and someone couldn’t understand that explanation unless she already knew what belief was. Moreover, the proposal faces the problem that there can be *degrees of belief*, but not degrees of truth (at least on most standard logics). One can express the level of our belief in a proposition in terms of percentages corresponding to the likelihood of that proposition being true, with 100% being perfect certainty, whereas truth and falsity are absolutes. This suggests that we should not explain “believing a proposition” simply in terms of assigning a proposition the value “true”; if we did we would not leave any room for explaining how we believed some propositions with less certainty than others.

### **Realism and Instrumentalism about belief**

The **dual-component** view of belief suggests that belief involves a certain kind of **mental representation** – i.e. that the *content* of our belief (the proposition) is somehow represented in our mind. **Realism about belief** is the view that beliefs are mental representations of propositions – of ways the world could be. One vivid way of putting the thought is to say that beliefs are like “sentences in the head” – that the content of our belief is represented somehow in the structure of our minds. This view is associated with Jerry **Fodor**, who has vigorously defended the idea that there is a **Language of Thought** – distinct from any particular human language we might speak – which enables our minds to represent the world as being a certain way. One of Fodor’s reasons for his view is that it explains how belief is “productive” and “systematic” – i.e. the capacity to form one belief enables us to form other beliefs by recombining the same elements:

“The property ... that I have in mind is one that inheres in the ability to understand and produce sentences. That ability is - as I shall say, systematic, by which I mean that the ability to produce/understand some of the sentences is intrinsically connected to the ability to

---

<sup>1</sup> This is a version of “Moore’s paradox”, so-called after its originator, G.E. Moore. The interesting feature of such sentences is that they can easily be true, and do not involve an obvious logical contradiction, but nevertheless they seem self-contradictory in some deeper sense that is very hard to explain.

produce/understand many of the others. You can see the force of this if you compare learning a language the way we really do learn them with learning a language by memorizing an enormous phrase book... The point that I'm now pushing is that you can learn any part of a phrase book without learning the rest. Hence, on the phrase book model, it would be perfectly possible to learn that uttering the form of words 'Granny's cat is on Uncle Arthur's mat' is the way to say that Granny's cat is on Uncle Arthur's mat and yet have no idea how to say that it's raining (or, for that matter, how to say that Uncle Arthur's cat is on Granny's mat).

... [Our] cognitive capacities must be at least as systematic as linguistic capacities, since the function of language is to express thought. To understand a sentence is to grasp the thought that its utterance standardly conveys; so it wouldn't be possible that everyone who understands the sentence 'John loves Mary' also understands the sentence 'Mary loves John' if it weren't that everyone who can think the thought that John loves Mary can also think the thought that Mary loves John. You can't have it that language expresses thought and that language is systematic unless you also have it that thought is as systematic as language is."

Fodor, *Why we still need a Language of Thought*

The alternative to **realism** about belief is **instrumentalism**. This is the view that beliefs are not real elements of our mind, but rather just convenient fictions made up to describe our disposition to behave in certain ways. This view is attractive primarily because it enables us to ascribe beliefs to non-human animals on the basis of their behaviour, even though we do not believe that these animals are capable of considering structured propositions or having a complex "Language of Thought".

One way of developing this view is through the **logical behaviourism** associated with the Oxford philosopher Gilbert **Ryle**, who attempted to explain away what seemed to be "internal" mental states such as belief and knowledge in terms of their outwardly visible manifestations:

"To believe that the ice is dangerously thin is to be unhesitant in telling oneself and others that it is thin, in acquiescing in other people's assertions to that effect, in objecting to statements to the contrary... But it is also to be prone to skate warily, to shudder, to dwell in imagination on possible disasters and to warn other skaters." Ryle, *The Concept of Mind*, p.129

The logical behaviourist suggests that all talk apparently about beliefs and desires can and should be reinterpreted by translating it into statements about actual behaviour and dispositions to behaviour. One advantage of this view is that it shows how the meaning of words such as "belief" is linked to the evidence we use when attributing beliefs to others; it also explains how we can have beliefs which we are not consciously aware of until we need to act on them. However, it has become very unpopular with philosophers, because of two major problems: first, if we say that beliefs are really nothing more than dispositions to action, we can no longer use beliefs to *explain* our dispositions to act in certain ways: if our beliefs just *are* our dispositions, then one cannot explain the other. Secondly, it gradually became apparent that statements about individual beliefs cannot in fact be translated into statements about specific dispositions to behaviour: if I believe that it is raining outside, what I do as a result depends entirely on what my *other* beliefs and desires

are. For example, this belief might dispose me to stay inside (if I want to stay dry), or to go outside (if I believe that I have left the sunroof of my car open), or to build an Ark (if I believe this is the start of a Biblical flood). Given that any specific belief can result in *any* action – or even total inaction – depending on the content of our other beliefs, it does not seem plausible to say that beliefs should be “reduced to”, or even “explained in terms of” dispositions to action, as the logical behaviourist suggests.

An alternative **instrumentalist** account is offered by Daniel **Dennett**’s theory of the **intentional stance**, which you met last year. Dennett’s idea is that we are justified in attributing beliefs to *any* complex creature (whether human, animal, or machine), so long as we find it useful to use the vocabulary of belief and desire to predict how that creature will respond to stimuli. For example, when playing a computer chess program I might consider what the computer “expects” or “believes” I will do next. Once I start using the vocabulary of “beliefs” and so on when dealing with something I am described as having “taken the intentional stance” towards that object. Dennett claims that the “beliefs” we talk about when we use the intentional stance do not correspond to anything “real” inside whatever it is we are talking about – even in the case of humans. Moreover, he denies that we can explain any single belief in terms of specific dispositions to behaviour; instead, we predict behaviour by assigning a set of beliefs which collectively enable me to work out what the machine will do next.

One problem with Dennett’s approach is that, while these “beliefs” enable us to predict the behaviour of humans and others, they do not *explain* that behaviour, since the actual cause of the behaviour has to be something *real*, i.e. the underlying mechanisms in the brain or circuitry of the thing we are interested in. Another problem is that the realist about belief will argue that there is a crucial difference between humans and computer programs or other things we might take the intentional stance towards: using the vocabulary of “belief” is just a *metaphor* when it is applied to computers and other machines, and it is a metaphor that we understand only because we understand the *literal* application of words such as “belief” to creatures which genuinely have minds. Finally there is the problem of understanding how beliefs *can* successfully predict the behaviour of human beings, if they do not correspond to anything “real” inside our minds. Surely the fact that beliefs are a useful predictive tool is good evidence for the view that beliefs really exist.

### **Can beliefs be voluntary?**

Can you *decide* to believe something? A belief could be described as **voluntary** if it is the result of a choice or (in the old-fashioned vocabulary) a “volition of the will” – an act of willing. We generally think that there can be such a thing as voluntary *action*, in so far as our actions can be the result of our own choices, but it is not so clear that we can *choose* to believe something, at least in this sense: I cannot believe that the balance of evidence is in favour of proposition *p* and yet decide to believe not-*p* for other reasons (for example, perhaps I really want *p* to be true). I might *pretend* to believe not-*p* to impress others, although I know that the balance of evidence is in favour of *p*, but even then I would be misleading people about my real beliefs rather than genuinely believing something that conflicted with my own grasp of the evidence for and against. It seems that beliefs

are not voluntary, then; our beliefs are *determined* by whether they seem to us to have compelling evidence in favour of them.

However, there are situations where the picture doesn't seem so simple. Even if someone cannot choose to hold beliefs in defiance of what seems to them to be the balance of the evidence, they can be held responsible for mistaken beliefs which are the end result of other choices they made, or what you might call wilfully ignoring evidence. For example, someone who supports an extremist political party might deliberately avoid mainstream media reports about politics deliberately to avoid confronting evidence that might force them to reassess their views; or they might decide to ignore any evidence that they *do* encounter if it is in conflict with their views. We may be compelled to form our beliefs in line with what seems to us to be the strongest evidence in our assessment; but this does not prevent us from voluntarily being selective or biased in forming our assessment.

Another reason for thinking that beliefs can be voluntary is offered by William Van Ormand **Quine**. His view is that there is no such thing as a piece of evidence that *forces* us to accept or reject a particular belief: our beliefs form an interconnected web which connects directly to experience "only along the edges". Quine emphasizes that any belief that seems to conflict with our experience can be retained providing we are willing to revise other beliefs so that the totality of our system of beliefs fits the totality of our experience. Thus we have to *choose* which beliefs to retain and which to reject whenever our evidence changes: there is never such a thing as our evidence *forcing* us to accept or reject a particular belief. That explains why extremists, creationists and conspiracy theorists are able to maintain their crazy central beliefs in the face of what seems to us to be overwhelming evidence to the contrary: they make adjustments in their background beliefs to ensure that, whatever evidence they are confronted with, they can still hang on to their convictions. A particularly common example of this is when people who present opposing evidence are dismissed as "part of the conspiracy" or "in the pay of the government". The belief that people are trying to trick you is a belief that enables you to dismiss their evidence without having to think too much about it. Here's how Quine makes his point:

"The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery occasions readjustments in the interior of the field. Truth values have to be redistributed over some of our statements. Re-evaluation of some statements entails re-evaluation of others, because of their logical interconnections -- the logical laws being in turn simply certain further statements of the system, certain further elements of the field. Having re-evaluated one statement we must re-evaluate some others, whether they be statements logically connected with the first or whether they be the statements of logical connections themselves. But the total field is so undetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to re-evaluate in the light of any single contrary experience. No particular experiences are linked with any particular statements in the interior of the field, except indirectly through considerations of

equilibrium affecting the field as a whole.

If this view is right, it is misleading to speak of the empirical content of an individual statement -- especially if it be a statement at all remote from the experiential periphery of the field. Furthermore it becomes folly to seek a boundary between synthetic statements, which hold contingently on experience, and analytic statements which hold come what may. Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. Even a statement very close to the periphery can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws." Quine, *Two Dogmas of Empiricism*

### **Justification: internalism and externalism**

In the first week we suggested that having justified beliefs could be thought of in terms of "having a justification" for what we believe, where this justification is a set of **reasons** or **evidence** which you could explain to someone else if prompted. This view of justification is known as **internalism about justification**: it is the view that justification (and hence knowledge) depends on factors **internal** to you, which you could be aware of: these "internal" factors may be **evidence** (this view is known as "evidentialism"), or may be **reasons**. Internalism is initially plausible because it follows "common sense": common sense tells us that we shouldn't claim to **know** things for which we have no evidence or reason.

The alternative view is known as **externalism about justification**: this is the denial of internalism, i.e. it claims that justification does not (only) depend on factors internal to the believer. Externalists say that justification depends on the relationship between the believer and the object of her belief: whether the believer is "hooked up to the world" in the right way. For example, the externalist might claim that my beliefs are justified if they have been formed by a reliable method – a method which generally results in true beliefs. It doesn't matter whether I can provide *reasons* for my beliefs; those beliefs are justified just so long as they are reliably formed, regardless of what I can say to "justify" them.

Some **externalists** go even further and say that **knowledge** is nothing to do with *justification*, but is simply something you get through your senses being hooked up to the world in a reliable way. We'll return to that more extreme form of externalism in a few weeks.

### **Problems for internalism about justification.**

A first problem for internalist theories is that they may have the consequence that only those with fully developed rational faculties can have justified beliefs: animals and small children, for example, don't have the capacity to be aware of their reasons for belief, so according to the internalist they can't have justified beliefs and so can't know anything. (NB: this depends on whether internalism requires the ability to articulate/explain your reasons; perhaps even a dog is "aware" of its reasons for believing that the cat is up the tree, in the sense that the dog's evidence for this belief (its visual experience) is something that it has "access" to.)

Secondly, internalist theories are vulnerable to **scepticism**. If I can't rule out the possibility that I might be a brain in a vat, then perhaps I no longer have any **reason** to trust my senses; in which



case, according to the internalist, I have no justification for forming beliefs on the basis of my senses and can't know anything about the outside world.

Finally, internalist accounts of justification face the threat of a **regress of reasons**. Suppose that my belief that  $p$  is justified by another belief that is my reason for believing that  $p$  – for example, my belief that  $q$ . What is my belief that  $q$  justified by? If it is justified by another belief, then it seems like we have the beginning of an **infinite regress**, in which every belief is justified by some other belief, in a chain that never ends. That would make the task of justifying my original belief seem like something that was constantly deferred but never actually completed, and many philosophers would argue that a belief that is part of an infinite chain of justification is never *really* justified after all.

Instead, some philosophers have adopted the approach that the chain of justification must end with some beliefs which are “self-supporting” and so provide the foundations for everything else we believe. This **foundationalism**, according to which some beliefs are self-justifying and need not be supported by other beliefs, was developed by **Descartes**, who suggested that belief in his own existence, and the logical principles needed to prove the existence of God, are deliverances of the “natural light” of reason, and do not need to be supported by other beliefs. However, there is room for doubt about whether these principles really are self-justifying, and there is room for doubt about whether a small set of foundational beliefs is enough to justify all of our other beliefs about the external world.

For those reasons, some philosophers have endorsed **coherentism**, which is the view that a belief is justified by how well it fits or “coheres” with the other beliefs that we hold. This explains how someone who lived 1,000 years ago can be described as having a *justified* belief that the Earth was created in six days: given what else this person believed (e.g. about the authority and literal truth of scripture), such a belief is exactly the one which fits best and so is the one which is justified for that person. However, coherentism faces the problem that we do not want to allow just *any* beliefs to count as justified on the basis that they form a coherent set. As we noticed already, many extremists and conspiracy theorists maintain sets of beliefs that are internally consistent with each other; but their crazy beliefs do not become justified merely because they form an internally consistent set.

### **Problems for externalism about justification**

The serious problems faced by an internalist account of justification (in terms of *reasons* or *evidence* to which the believer has conscious access) has persuaded many philosophers to adopt theories of justification according to which justification depends on factors which the subject may be unaware of. A popular version of this **externalist** theory is **reliabilism**, according to which your beliefs are justified if they are formed using a reliable method – regardless of whether you know that the method is in fact reliable. However, externalism about justification faces problems at least as severe as those threatening the internalist.

First, externalists say that you can be justified in believing (and hence know) things without being aware of the reasons or evidence for that belief. A counter-example is given by Norman the clairvoyant. He uses his powers to form beliefs about the future which are in fact reliably true (although he has never checked up on them) – so Norman has true beliefs, and is “hooked up to the

world” in a reliable way. The externalist would want to say that Norman has justified beliefs about the future – but how can we describe these beliefs as “justified” if – as far as Norman knows – he has *absolutely no reason* to trust his predictions? This example was used by Laurence **BonJour** to suggest that the externalist is wrong, and reasons or evidence are in fact required for justification.

Second, externalists are committed to the view that we can’t have justified beliefs when our methods of forming beliefs have gone wrong somehow and we *aren’t* “hooked up to the world” in a reliable way. But consider a situation where you really *are* a brain in a vat being fed misleading images of a fake “external world”. The natural thing to say about this situation is that, although you have a great deal of false beliefs about your environment (e.g. that you are in a classroom now), those beliefs are nevertheless *justified* – they are the appropriate beliefs for you to have given the evidence of the sensory experience fed to you by the machine and the fact that you have no reason for believing that this evidence is misleading. This is another example which suggests that the correct account of justification will talk about the believer’s evidence or reasons, as the internalist claimed.

A final problem for externalism is that externalists accept that we may be justified in believing (and hence know) things without being able to check *whether* we are in fact justified. So we can know things in circumstances where it is impossible for us to **know that we know**. For example, suppose that I am in fact *not* a brain in a vat. Then in fact my visual experience is a reliable way of forming beliefs about the world, so my beliefs about the world are justified. But the fact that justifies my beliefs (the fact that my visual experience is reliable) is not a fact I can ever test or check up on, since there is absolutely no way I can check whether I am a brain in a vat.

That way of thinking has led some philosophers to argue that externalists **miss the point** of a philosophical account of justification and knowledge, since externalists have nothing to say about **why** we should trust our beliefs. Thus Barry **Stroud** says

“Suppose that... humans form their beliefs about the physical world through the operation of belief-forming mechanisms which are on the whole reliable... this fact alone will do us no good as theorists who want to understand human knowledge in this philosophical way... we want to know or have good reasons for thinking that what we believe... is true”

Stroud’s complaint is this: the search for a philosophical theory of justification is in part an attempt to find out *why we ought* to believe what we believe; but the externalist has nothing to say about whether we *ought* to trust our beliefs, because according to them the factor that makes our beliefs trustworthy (their reliable connection to the world) is something we can never know to be present or absent.

*Homework:* “My beliefs are justified only if I can state my reasons for them”. Critically evaluate this claim.

## Week 5: Knowledge: the tripartite definition

One popular view about scepticism is that is not a philosophical position that we might seriously consider accepting (for there are very few sincere global sceptics in the world); rather, it is thought that avoiding scepticism is a **challenge** facing our attempts to reach a theory of knowledge – or more precisely, that scepticism sets up a **test**: if our theory of knowledge has the consequence that all knowledge is impossible (global scepticism), that just shows that our theory of knowledge is **wrong** and needs to be changed. Duncan **Pritchard** endorses this view of scepticism:

“we are not to think of the ‘sceptic’ as a person – as someone who is trying to convince us of something – but rather as our intellectual conscience which is posing a specific kind of problem for our epistemological position in order to tease out what our view really involves and whether it is a plausible stance to take” (Pritchard, *What is this thing called Knowledge?*, p.137)

With that in mind, the focus of the debate shifts, from “is knowledge possible?”, to “what is knowledge, anyway?”. Such a shift intuitively makes sense: how can we discuss whether we *can* know anything unless we are pretty sure we know what knowledge *is*? Unfortunately – as we’ll see in the next few weeks – there is no universally accepted philosophical definition of knowledge.

### The Tripartite Definition

Tri-partite: “having three parts”. The **tripartite definition of knowledge** defines knowledge in terms of **three** conditions:

Subject S knows that p if and only if:

- S has a belief that p;
- p is true;
- and p is justified

This is sometimes shortened to the claim that knowledge is “justified true belief”. Notice that the claim is that each of the three conditions is individually **necessary** (“needed”) for knowledge – i.e. *you can’t have knowledge unless each condition is met*. Notice also that the conditions together are claimed to be jointly **sufficient** (“enough”) for knowledge – i.e. the definition says that every situation in which subject S meets all three conditions is a situation in which S has knowledge.

The tripartite definition is almost as old as philosophy itself: a version of it is first considered by Plato, and it was widely accepted as correct until the second half of the 20<sup>th</sup> century.

“When, therefore, any one forms the true belief about anything without rational explanation, you may say that his mind is truly exercised, but has no knowledge; for he who cannot give and receive a reason for a thing, has no knowledge of that thing; but when he adds rational explanation, then, he is perfected in knowledge”. – *Plato, Theaetetus 202c*

NB: Plato himself did not fully endorse the definition: he found himself unable to give a satisfactory account of what is meant by a “rational explanation” (λογος) .

This week we’ll consider the three conditions individually, asking whether each of them really is **necessary** for knowledge in every situation; next week we’ll consider arguments that suggest that the conditions are not jointly **sufficient**– i.e. that there are cases of justified true belief that are *not* cases of knowledge. Notice that definitions are usually tested by giving **counterexamples**. For example, the biographer Diogenes Laertius tells us that when Plato attempted to define man as a

“featherless biped”, Diogenes the Cynic went into the back yard, plucked a chicken, and came back waving it, saying “this is Plato’s man”.

*Exercise:* Did Diogenes show that Plato’s definition did not provide **necessary** conditions for being a man, or that the definition was not **sufficient**? Why?

### 1. Does Knowledge require Belief?

The tripartite definition represents knowledge as a **kind** of belief – it is a belief that is also justified and true. But often you will hear people say things like “I *don’t* believe it; I know it!”. The view that knowledge and belief are entirely different things was endorsed by Plato in the *Republic* (although he later seems to have rejected that view in favour of something like the tripartite definition). According to Plato, belief, or “*doxa*”, is fallible, and concerns the changeable things in this visible world, whereas knowledge (“*episteme*”) is infallible, and concerns only the eternal and unchanging world of Forms such as Justice and Beauty. So you cannot both believe and know the same thing: the things that you know must be different from the things that you believe, and knowledge cannot be a kind of belief.

However, we do not have to accept Plato’s view. If we reject his suggestion that knowledge has to be **infallible**, then we can know and have beliefs about the very same everyday objects; the fact that our beliefs might be mistaken does not show that they cannot count as knowledge. Moreover, there is a simple answer to the example of the person who says “I don’t believe it; I know it!” – that they have missed out the word “*merely*” from what they say, and what they intend to say is that they don’t *merely* believe something – they also know it.

In fact, it often sounds like a contradiction for someone to claim to know something without believing it: saying “I know that it is raining but I don’t believe that it is raining” sounds every bit as contradictory as “I know that it is raining but it is not raining”.

Nevertheless, recent discussion has suggested that there might be *some* good counter-examples to the claim that knowledge requires belief. Suppose that a conscientious student has studied hard for an exam, writes the correct answers on the exam paper, but does not feel confident enough to believe that those answers are correct. You might think that such a student *knows* what the answers are (after all, she has studied hard), but does not *believe* that the answers are the ones she has given; so this is an example of knowledge without belief.

A response to such an example could take one of two forms: first, to accept that the student knows the answers, but to say that she also believes the answers: after all, it might seem plausible to say that her actions in deliberately choosing *those* answers rather than any others demonstrated that she had some degree of belief in them! Alternatively, we could say that the student neither believes nor knows the answers: if she is so unsure of the answers that she has given that she does not really believe that they are the right ones, how can she be said to *know* that these are the right answers?

## 2. Does Knowledge require Truth?

Most philosophers agree that knowledge entails truth – you can't know what is not true. Even if you have a well-justified belief that *p*, this belief can't count as knowledge unless it is in fact true that *p*. Evidence for this comes from our everyday use of "know": in most everyday conversations "I know that *p*" always implies "*p* is true". Plus, no-one would ever say "I know that *p* but *p* isn't true" – this seems like a straightforward contradiction. Thirdly, saying that knowledge requires truth enables us to begin to explain why knowledge is something *useful* or *desirable*. As Aristotle said, "All men by nature desire to know". Saying that knowledge must be true shows why we might want knowledge.

One very common response to this is to point out that people often *think* they know that *p*, or *claim* to know that *p*, in circumstances where *p* is in fact not true– so knowledge doesn't require truth after all! But this simply confuses *thinking you know* with *actually knowing*. Plenty of people think they know things without *actually* knowing them; when you find out what you were claiming to know is not true, you might say something like "I thought I knew that *p*, but in fact I *didn't* know that *p*: I had made a mistake – *p* is not true!" So the claim that knowledge requires truth survives this criticism.

Alternatively, you might give the example of people in distant cultures who have "different knowledges" – e.g. by saying that "Ordinary people in the ancient world 'knew' the earth was flat". This seems to be an example of "knowledge" which does not require truth. This kind of "knowledge" even made it into a scene in the film "Titanic":

"He [the captain] figures anything big enough to sink the ship they're going to see in time to turn. But the ship's too big to, with too small a rudder... everything he knows is wrong."

The standard answer to examples like these is to say that it is simply imprecise use of language. People sometimes use "know" to mean something like "believe strongly" or "be really convinced that" – but this is not the normal meaning of "know", and it is not the kind of "knowledge" that philosophers are interested in.

A neat way of putting the point that knowledge requires truth is to say that "knows" is a **factive verb**, i.e. that what you know is always a **fact**, and therefore always something **true**.

*Exercise:* which of the following are statements containing a "factive verb"?

- (1) Matilda discovered that *p*.
- (2) Lily forgot that *p*.
- (3) Holly surmised that *p*.
- (4) Ashling admitted that *p*.
- (5) Emma supposed that *p*.
- (6) Rebecca was happy that *p*.
- (7) Abi regretted that *p*.
- (8) Banuta saw that *p*.

### 3. Does Knowledge Require Justification?

We've already discussed "justification" to some extent. Traditionally, justification has been thought of in terms of **reasons** or **evidence** the knower has access to, which make it **rational** for her to believe what she believes. My beliefs count as "justified" if I (or someone else) can explain these reasons or evidence to demonstrate that – in the circumstances – I ought rationally to believe what I believe. For example, I might claim to have a justification for believing that I am in school because I can specify certain experiences I am currently having which count as my **evidence** for believing that I am in school. And I might claim to be justified in believing that a complex mathematical theorem is true because I understand a proof which gives me a really good reason for believing the theorem.

The main motivation for believing that **justification is necessary for knowledge** is to prevent "lucky guesses" from counting as knowledge. Sometimes our beliefs turn out to be true, not because we have formed them in a sensible or rational way, but because we just got lucky. Such beliefs, which are true by accident, do not count as knowledge. For example: an optimistic gambler who is always convinced her horse will win the race finally gets lucky – her horse wins the race. So the gambler had a true belief – *that her horse will win the race* – but she did not **know** that her horse would win the race, because it was just luck that her belief turned out to be true. Obviously knowledge requires more than just having a true belief, regardless of how that belief was formed. So it might seem sensible to say that only those beliefs that are based on good reasons or solid evidence should count as knowledge. Then we can say that the optimistic gambler did **not** know that her horse would win, because her true belief that the horse would win was not **justified** – it was irrational, not based on reasons or evidence.

Another reason for thinking that knowledge requires justification was suggested by Plato in the *Meno*: he said that an important difference between knowledge and true belief is that knowledge is permanent and stable, whereas mere belief is liable to change at any moment. Here he compares beliefs to the mythical statues of Daedalus, which were said to be automata which would walk off their plinths unless chained down:

"Now this is an illustration of the nature of true beliefs: while they abide with us they are beautiful and fruitful, but they run away out of the human soul, and do not remain long, and therefore they are not of much value until they are fastened by the tie of reason; and this fastening of them, friend Meno, is recollection, as you and I have agreed to call it. But when they are bound, in the first place, they have the nature of knowledge; and, in the second place, they are abiding. And this is why knowledge is more honourable and excellent than true opinion, because fastened by a chain." *Plato, Meno 97*

So we might say – following Plato – that justification is what explains why knowledge is more permanent than belief: understanding the reasons behind our beliefs enables us to see why they are true, and so to avoid changing our mind in the face of every new bit of evidence. One example might be finding out the answers to a maths exercise: if you just look over the shoulder of the person sitting next to you, you have a changeable **belief** which is easily altered if you see someone else with a different answer, but if you work out the answer for yourself you have a piece of justified **knowledge** which you will not let go of as easily.

However, there are also arguments to suggest that justification is **not** always needed for knowledge. We have met some of these already. It cannot be the case that every belief is justified by some *other* belief; that would lead to either an **infinite regress** of justification, or a chain of justification that is **circular**, with A justifying B justifying C justifying D justifying A again. So some beliefs must **not** be justified by other beliefs to avoid regress or circularity.

*Exercise:* What, if anything, is wrong with an infinite, or circular, chain of justification?

Moreover, some philosophers have argued that certain propositions can be known without justification because they are **epistemically basic** – they are propositions that you “just see” have to be true, without having to explain how/why you believe they are true. Common examples include:

- Knowledge that I exist
- Propositions about how things seem to me – e.g. “I seem to be in a philosophy classroom”
- Propositions about the content of my mind – e.g. “I am in pain”
- Simple mathematical propositions – e.g.  $2+2=4$
- Knowledge of God’s existence (known through faith and so no justification required)

Some **externalist** philosophers have gone even further, saying that knowledge is nothing to do with having a “justification” at all. For example, you could know that the Battle of Hastings took place in 1066 without being able to explain your reasons or evidence (Perhaps you have forgotten how you came to know this fact). Externalists often argue that knowledge is not a question of being able to justify or support your beliefs, but rather depends on whether you are “hooked up to the world” in the right way – for example, whether you formed the belief by a reliable process. In support of this suggestion they might point out that non-human animals and infants can “know” things even though they are incapable of understanding or expressing any kind of justification for what they believe.

*For discussion:* can animals know things?

*From David Lewis, "Elusive Knowledge" (1996: p.551):*

"I don't agree that the mark of knowledge is justification.

First, because justification is not sufficient: your true opinion that you will lose the lottery isn't knowledge, whatever the odds. Suppose you know that it is a fair lottery with one winning ticket and many losing tickets, and you know how many losing tickets there are. The greater the number of losing tickets, the better is your justification for believing you will lose. Yet there is no number great enough to transform your fallible opinion into knowledge - after all, you just might win. No justification is good enough - or none short of a watertight deductive argument, and all but the sceptics will agree that this is too much to demand?

Second, because justification is not always necessary. What (non-circular) argument supports our reliance on perception, on memory, and on testimony? And yet we do gain knowledge by these means. And sometimes, far from having supporting arguments, we don't even know how we know. We once had evidence, drew conclusions, and thereby gained knowledge; now we have forgotten our reasons, yet still we retain our knowledge. Or we know the name that goes with the face... by relying on subtle visual cues, without knowing what those cues may be."

*Homework: "Critically assess the view that justified true belief is necessary for knowledge."*



## Week 6: Gettier cases

Last week we considered whether each of the three conditions in the **tripartite definition** (belief, truth, and justification) is **individually necessary**. Now we need to ask whether these conditions are **jointly sufficient** – i.e. is the combination of belief, truth, and justification always enough for knowledge? To show that the conditions are *not* sufficient, we simply need to find a **counter-example** – a case in which someone has a justified true belief which nevertheless is not knowledge. Notice that this counter-example does not have to be a real-life situation: because the claim under discussion is that justified true belief is knowledge *by definition*, it follows that every **conceivable** case of justified true belief ought to be knowledge, whether or not the situation we are conceiving actually occurs.

We have already met one way of rejecting the claim that justified true belief is sufficient for knowledge: if you think that knowledge requires **certainty**, then there will be plenty of cases where someone has a belief that is both justified and true, but does not really know because she is not really certain or sure of what she believes. I believe that Paris is the capital of France, because I have read this in many different books. And in fact Paris is the capital of France, so I have a belief which is true as well as being justified. But I am not totally certain of this belief – there remains the tiny possibility that I could have mis-remembered or made some other mistake. So if knowledge requires certainty, then I don't really know that Paris is the capital of France, even if I have a justified true belief that Paris is the capital of France. However, the view that knowledge requires certainty is controversial: normally we *would* say that I “know” that Paris is the capital of France in this situation, so we need some further philosophical argument to convince us that knowledge is impossible without certainty.

Another possible kind of counter-example to the tripartite definition comes from what is sometimes called the **KK principle**: that you can't know something unless you **know that you know** it. I might have a justified true belief that Paris is the capital of France, but be unsure whether my information comes from a reliable source. So I don't know whether my belief counts as knowledge, so don't know that I know, so don't know that Paris is the capital of France. Again, however, the principle that you don't know something unless you know that you know it is controversial – apart from anything else, it means that children and animals cannot know anything, as they do not have the concept of knowledge in the first place and so can't reflect on whether they know things.

Nevertheless, there is a kind of counter-example to the tripartite definition which many philosophers regard as decisive, proving that justified true belief is *not* sufficient for knowledge. These counter-examples are known as **Gettier cases**, named after Edmund Gettier, who introduced them in his 1963 article, “Is Justified True Belief Knowledge?”. Gettier originally gave two such cases; read the excerpts from his paper and answer the questions that follow each one.

### Gettier case I

Suppose that Smith and Jones have applied for a certain job. And suppose that Smith has strong evidence for the following conjunctive proposition:

(d) Jones is the man who will get the job, and Jones has ten coins in his pocket.

Smith's evidence for (d) might be that the president of the company assured him that Jones would in the end be selected, and that he, Smith, had counted the coins in Jones's pocket ten minutes ago. Proposition (d) entails:

(e) The man who will get the job has ten coins in his pocket.

Let us suppose that Smith sees the entailment from (d) to (e), and accepts (e) on the grounds of (d), for which he has strong evidence. In this case, Smith is clearly justified in believing that (e) is true.

*Questions:*

What is Smith's justified true belief? \_\_\_\_\_

Why is it not something that he *knows*? \_\_\_\_\_

### Gettier case II

Let us suppose that Smith has strong evidence for the following proposition:

(f) Jones owns a Ford.

Smith's evidence might be that Jones has at all times in the past within Smith's memory owned a car, and always a Ford, and that Jones has just offered Smith a ride while driving a Ford. Let us imagine, now, that Smith has another friend, Brown, of whose whereabouts he is totally ignorant. Smith selects three place names quite at random and constructs the following three propositions:

(g) Either Jones owns a Ford, or Brown is in Boston.

(h) Either Jones owns a Ford, or Brown is in Barcelona.

(i) Either Jones owns a Ford, or Brown is in BrestLitovsk.

Each of these propositions is entailed by (f). Imagine that Smith realizes the entailment of each of these propositions he has constructed by (f), and proceeds to accept (g), (h), and (i) on the basis of (f). Smith has correctly inferred (g), (h), and (i) from a proposition for which he has strong evidence. Smith is therefore completely justified in believing each of these three propositions. Smith, of course, has no idea where Brown is.

But imagine now that two further conditions hold. First Jones does not own a Ford, but is at present driving a rented car. And secondly, by the sheerest coincidence, and entirely unknown to Smith, the place mentioned in proposition (h) happens really to be the place where Brown is. If these two conditions hold, then Smith does not know that (h) is true, even though (i) (h) is true, (ii) Smith does believe that (h) is true, and (iii) Smith is justified in believing that (h) is true.

*Questions:*

What is Smith's justified true belief? \_\_\_\_\_

Why is it not something that he *knows*? \_\_\_\_\_

Gettier cases have this feature in common: although the subject has a justified true belief, it is just **luck** that their belief is true – and so their belief isn't knowledge, because a belief can't count as knowledge if it is true just through chance or luck. That feature comes out strongly in another, simpler Gettier case (not one that he included in his paper):

### **Gettier case III**

I look at the clock and notice that it is showing ten o'clock. I know that the clock has been working correctly during the previous week, so I form a justified belief that it is ten o'clock. In fact, it is ten o'clock. So my belief is true. However, what I don't know is that the clock is broken: it stopped working exactly 12 hours ago. It is sheer luck that I happened to look at the stopped clock at the one moment of the day when it is in fact showing the correct time. So it is sheer luck that my belief is true. So I don't *know* that it is ten o'clock, even though I do have a justified true belief that it is ten o'clock.

### **Making your own Gettier cases**

Gettier cases are easy to make up, using a two-stage process. First, imagine a situation in which someone has a justified *false* belief. For example, imagine someone using a stopped clock to tell the time. Second, alter the story slightly so that the belief that they form is – by some fluke or stroke of luck – actually true. For example, suppose that the stopped clock the person is looking at is – just by chance – telling the right time. Then you should have a good example of a justified true belief that isn't knowledge.

### **The Gettier game**

Form teams. Each team has 15 minutes to come up with as many different Gettier cases as possible (nominate someone to make a note of them). The winner is the team with the most genuine cases of justified true belief that isn't knowledge at the end of that time. However, the opposing teams can try to reduce each others' scores when going through the Gettier cases at the end of the time, by raising **objections** to individual cases. Objections could include:

- That isn't really a **justified** belief!
- In that situation the subject actually **does** have knowledge!
- What the subject believes isn't really **true**!

*Homework:* Briefly explain (in less than one side of A4) why Gettier cases show that knowledge is not the same as justified true belief.

## Week 7: Responses to Gettier

One possible (but perhaps unconvincing) response to Gettier cases would be to say that these really *are* cases of knowledge after all – it is just that sometimes you can have knowledge as the result of chance or luck. But most philosophers don't adopt that approach; the point of making a distinction between knowledge and true belief in the first place is that someone who *knows* has arrived at her belief using a method that somehow connects her reliably with the truth of the matter: her belief is true *because of* the method she used to arrive at it. This means that beliefs which are true merely by chance cannot count as knowledge. In this section we'll consider some of the most popular **responses to Gettier** in recent philosophy. What is meant by a "response to Gettier" is this: an attempt to alter or amend the tripartite definition so that it is no longer vulnerable to Gettier-style counter-examples.

### Infallibilism again

If it were possible to argue that the subject in the Gettier cases doesn't really have a justified true belief, then Gettier cases wouldn't be counter-examples to the tripartite definition. One way to do this is to argue that the subject in these cases isn't really **justified**; and the way to argue for this is to go back to a suggestion about knowledge we met already: **infallibilism**. Infallibilists claim that we can have knowledge only when we have investigated so thoroughly that it is impossible for us to have made a mistake; thus they sometimes introduce the concept of **infallible justification** – that we are *really* justified only when we have formed our belief in such a way that it is impossible for us to be mistaken. In the original Gettier cases, Smith had beliefs that seemed reasonable – but they weren't **infallible**, and in the circumstances he could easily have been mistaken. So an infallibilist might say that Smith's beliefs weren't really justified; in fact we have justified beliefs only in situations where it is impossible for us to have made a mistake.

This suggestion has not been very popular with philosophers, for three reasons.

- First, it seems at odds with our normal concept of justification: normally we would say that the subject in a Gettier case *is* justified. After all, the subject in a Gettier case has a belief that it is rational to hold in the circumstances – a belief that she holds for good reason. Why then should we deny that she is justified in holding the belief?
- Second, saying that justification is infallible justification means that you can be justified in holding only beliefs which are **true**: if justification requires the impossibility of having made a mistake, then justification requires truth, because anyone who has a false belief has certainly made a mistake! Defenders of the view that justification is infallible justification are committed to saying that there is no such thing as a **justified false belief**, since a subject who has made a mistake is (for that reason) not justified in believing what she believes. This seems implausible: surely there are plenty of cases in which we are entirely justified in believing what we do, although what we believe is false.
- Finally, there is the problem of a return to **scepticism**. In fact we have few (if any) beliefs which it is impossible for us to be mistaken about; imposing the requirement that knowledge requires infallibility makes it very hard if not impossible to know anything at all. But most people would say that we do know at least some things. So infallibilism is not a convincing account of our normal concept of knowledge.

## Indefeasibility

**Indefeasibility** is the suggestion that we can get round the Gettier cases by adding an extra fourth condition to the tripartite definition, to prevent beliefs from counting as knowledge if the believer has made some kind of mistake. An early version, “**no false lemmas**” or “no false reasons”, suggested that we should add the requirement that justified true belief is knowledge only if the justification is not based on a false premise or mistake in reasoning.

A problem with this is that not all justified beliefs are justified by *reasoning*. For example, I can see a sheep in a field and directly form a justified belief that there is a sheep in the field without going through any process of reasoning or inference to arrive at my belief. For that reason the “no false lemmas” requirement was replaced by the requirement that there be no “defeaters”:

Defeater =<sub>df</sub> a truth **D** such that, if **S** were justified in believing **D**, then **S** would not be justified in believing her original belief **p**

A problem for Indefeasibility is that the presence of relevant false belief doesn't always prevent me from knowing things: I can know that my friend is giving me a lift tonight even if there are other truths (e.g. that his car broke down this morning) that would “defeat” my justification. (We have to assume that his car broke down this morning and then was fixed this afternoon – so I get my lift after all.)

More generally, the “no defeaters” requirement makes knowledge too hard to achieve and opens the way to the return of **scepticism**. Most of my justified beliefs about the world are **defeasible**, in the sense that there is some other truth which would remove my justification if I came to know it. Saying that knowledge is **indefeasible justified true belief** makes knowledge impossible in many everyday situations.

## Reliabilism

**Reliabilism** is the suggestion that knowledge is true belief that has been formed by a **reliable process**. Reliabilists sometimes go so far as to say that justification is irrelevant to knowledge – what matters is simply the reliability of the process used to form the belief. This approach was defended by Fred **Dretske**:

“Those who think knowledge requires something other than, or at least more than, reliably produced true belief, something (usually) in the way of justification for the belief that one's reliably produced beliefs are being reliably produced, have, it seems to me, an obligation to say what benefits this justification is supposed to confer.... Who needs it, and why? If an animal inherits a perfectly reliable belief-generating mechanism, and it also inherits a disposition, everything being equal, to act on the basis of the beliefs so generated, what additional benefits are conferred by a justification that the beliefs are being produced in some reliable way? If there are no additional benefits, what good is this justification? Why should we insist that no one can have knowledge without it?” (Dretske, “The Need to Know”, p. 95)

Part of Dretske's motivation is the thought that our definition of knowledge should not exclude animals and children from having knowledge; even creatures which are incapable of understanding a justification of their beliefs should be able to “know” things.

However, reliabilism faces some serious problems. First, it is not clear that the reliabilist succeeds in avoiding Gettier cases. Looking at a clock is generally a **reliable** way of coming to tell the time, so in the “stopped clock” Gettier case the reliabilist may be forced to say that **S** knows what time it is after all. Similarly, try this example from Alvin Goldman:

“Henry is driving in the countryside with his son. For the boy's edification Henry identifies various objects on the landscape as they come into view. “That's a cow,” says Henry, “That's a tractor,” “That's a silo,” “That's a barn,” etc. Henry has no doubt about the identity of these objects; in particular, he has no doubt that the last-mentioned object is a barn, which indeed it is. Each of the identified objects has features characteristic of its type. Moreover, each object is fully in view, Henry has excellent eyesight, and he has enough time to look at them reasonably carefully, since there is little traffic to distract him.

...

Suppose we are told that, unknown to Henry, the district he has just entered is full of papier-mache facsimiles of barns. These facsimiles look from the road exactly like barns, but are really just facades, without back walls or interiors, quite incapable of being used as barns. They are so cleverly constructed that travelers invariably mistake them for barns. Having just entered the district, Henry has not encountered any facsimiles; the object he sees is a genuine barn. But if the object on that site were a facsimile, Henry would mistake it for a barn. Given this new information, we would be strongly inclined to withdraw the claim that Henry knows the object is a barn.”

Goldman, “Discrimination and Perceptual Knowledge”, p.772-3

In Goldman's example of driving through **fake barn county**, the traveller Henry forms a true belief (“that's a barn”) using a **reliable method**, namely visual perception. But he doesn't **know** that it is a barn – it could easily have been a fake barn, and Henry wouldn't have been able to tell the difference.

The defender of reliabilism might respond that, although perception is normally a reliable way of finding out whether there are barns around, it isn't a reliable method in fake barn county, because perception doesn't guarantee the right answer when used to look for barns in fake barn county. But this runs the risk of a return to **infallibilism**: if a “reliable method” is one that **guarantees** we get the right answer in the circumstances we are actually in, that means that we can know something only when the method we have used makes it impossible for us to have made a mistake.

Another problem for reliabilism is explained by Duncan **Pritchard**:

“Imagine, for example, that you find out what the temperature of the room is by looking at the thermometer on the wall... Suppose, however, that unbeknownst to you someone is playing a trick on you. The thermometer is in fact broken and is fluctuating randomly. Crucially, however, this isn't making the thermometer an unreliable indicator of what the temperature in the room is for the simple reason that someone is hidden in the room and adjusting the temperature of the room to match whatever reading is on the thermometer whenever she sees you look at the thermometer.”

Pritchard, *What is this thing called Knowledge?* p.56

In this example, you have a true, reliably-formed belief about the temperature in the room which is nevertheless *not* knowledge, because you can't know the temperature of the room by looking at a broken thermometer!

## Truth-tracking

**Truth-tracking** is the theory (defended by Robert **Nozick**), that true beliefs count as knowledge if they would “track the truth” – i.e. if my belief that **p** is **sensitive to** or **depends on** whether **p** is true or not. More precisely:

**S** knows that **p** if and only if

**p** is true

**S** wouldn't have believed that **p** if **p** wasn't true

**S** wouldn't have failed to believe that **p** if things were different but **p** was still true

The second requirement deals with the Gettier cases: in a Gettier case it's just luck that my belief is true, and I **would** go on believing it even if the belief was in fact false. (cf. the “stopped clock”: whatever time it is, I'm going to believe it's 10am on the basis of the clock – it's just luck that the stopped clock shows the right time.)

Nevertheless, Truth-Tracking faces several problems:

- Does my belief that I'm not a brain in a vat track the truth? Suppose I'm not a brain in a vat; then I have a true belief. But if I **were** a brain in a vat, I would still believe what I believe now. So my true belief that I'm not a brain in a vat does not “track the truth” and so I don't know that I'm not a brain in a vat. But then surely I don't know anything else either!

(Nozick agrees that I don't know I'm not a brain in a vat, but denies that you can infer from (i) “I don't know that I'm not a brain in a vat” to (ii) the more specific conclusions “I don't know that I'm here in this room / have hands / have a body” etc. Does this sound sensible? If not, why not?)

- Does this analysis get us any closer to understanding *what knowledge is*? Suppose it is **true** that, if we know something, our belief “tracks the truth”: nevertheless, we won't understand knowledge until we know *what it is* about “truth-tracking” beliefs that enables them to track the truth – *why* does knowledge track the truth?

## Knowledge as appropriately-caused belief

The **Causal Theory of Knowledge**, proposed by Alvin Goldman, says that **S** knows that **p** if and only if **S**'s belief was caused by the **fact** that **p**. So we can explain why the case of the broken thermometer doesn't count as knowledge: my belief about the temperature of the room wasn't caused by the fact that the room was a certain temperature; instead the random fluctuations of the broken thermometer were both causing my beliefs about the temperature and causing the concealed observer to change the room temperature to match.

Nevertheless, this theory faces some problems of its own:

- Can it deal with the visit to fake barn county? Henry's true belief that there is a barn in the field is caused by the fact that there is a barn in the field – but Henry doesn't *know* that there is a barn in the field because he could just as easily have been looking at many of the fake barns in the vicinity, and wouldn't have been able to tell the difference.
- What if my true belief that **p** was caused by the fact that **p**, via an unreliable process? Bob's anger causes Pete to tell me that Bob is angry – but Pete is a very unreliable judge of



character who I shouldn't really believe. Although my belief that Bob is angry was (indirectly) caused by the fact that Bob is angry, it doesn't seem like I can claim to **know** it.

- This suggestion seems to make it impossible to have knowledge of the future. My current belief that **p** will happen would have to be caused by a fact in the future – but then causation would have to run *in reverse*, from future to past.
- This suggestion involves us in a metaphysics of **facts**, conceived as entities “out there in the world” which cause our beliefs. But this is hard to understand and controversial (although Wittgenstein did famously say that “the world is the totality of facts, not of things”).

### **Abandoning the project of analysing knowledge**

Before leaving this topic we should briefly note two recent approaches which suggest that it was a mistake to try to discover a single definition which provides necessary and sufficient conditions for knowledge. One is **contextualism**, the view that the standards for a correct use of the verb “know” vary between different contexts of use. Where the stakes are low – the truth or falsity of a belief is relatively unimportant – the standards for knowledge are correspondingly lower and someone may count as “knowing” simply in virtue of having some correct information. This is the sense of “know” used in questions like “How many children in this class know that Paris is the capital of France?”. However, where the stakes are high, and it is very important that the belief is true, then we use the verb “know” much more stringently, demanding more in the way of justification or certainty than we would have done if the object of our knowledge had been less important.

You will remember that this **contextualist** approach also enables us to deal with some of the challenge posed by scepticism: we can say that, in everyday contexts, when not much is at stake, we do “know” that we have a physical body, and that the external world exists; however, in the context of a philosophical discussion, where the standards for knowledge are more rigorous, we might concede that we *don't* “know” that the external world exists, because we don't “know” we are not a brain in a vat. There is nothing inconsistent about this: the contextualist is really saying that there is no one single kind of knowledge; the meaning of the verb “know” is different in different contexts, so we can claim to “know” in one context, and “not know” in another without contradicting ourselves.

Finally, you should be aware that some philosophers (most notable Timothy **Williamson** at Oxford) now believe that knowledge **has no analysis** – i.e. that knowledge is conceptually basic, and cannot be defined in terms of more fundamental concepts such as belief, truth, and justification. This idea might seem appealing – after all, there is no universally-agreed solution to the Gettier problem. Moreover, it is not the case that every concept must be explicable in terms of other, more basic concepts on pain of regress. Against this suggestion, other philosophers argue that knowledge is obviously a *kind* of true belief; all the Gettier problem showed is that it is far from obvious how we should distinguish the true beliefs which *are* knowledge from those which are not. So two-thirds of the analysis of knowledge is complete – it is something that is both true and believed; all that is lacking is to say what else must be added to true belief to give knowledge, if it is not simply justification. Rather depressingly, this is exactly the dialectical situation that Plato found himself in at the end of his dialogue *Theaetetus*, early in the 4<sup>th</sup> century BC.

*Homework: "Gettier cases prove that no analysis of knowledge is possible. Discuss"*

## Week 8: Nominalism, Realism, and Conceptualism

The debate between Nominalists and Realists has a long history: it concerns a theory that originated with Plato, and has remained controversial throughout the medieval period and into the present day. (In fact, it has sometimes proved *too* controversial: there are reports of riots between rival nominalist and realist factions in medieval Paris, and in 1092 the good people of Soissons persuaded the nominalist Roscellinus to change his philosophical position on this issue by threatening to stone him to death.) So what do Nominalists and Realists believe?

### Realism

**Realists** believe that the mind-independent world contains not only particular objects (“**particulars**” for short), but also **universals**. You might think that you have never heard of universals before, but the realist will say that you talk about them all the time: according to the realist, universals are what you talk about when you talk about the **properties, qualities, or attributes** that objects have, and the **relations** they stand in. According to the realist, universals are also the **referents of general terms** – i.e. universals are what we refer to when we use **adjectives** such as “white” and “square”, and when we use **abstract nouns** such as “justice” and “courage”.

We’ll look at the arguments for realism in more detail next week, but for now we should note that there is one very straightforward argument for the existence of universals: namely, that much of what we say in both everyday and scientific discourse **presupposes** or **commits us** to their existence, whether we like it or not. To deny that universals exist seemingly requires us to give up many natural ways of expressing truths about the world, for example:

- Napoleon had all the **qualities** of a great general.
- Quarks and electrons do not have any **properties** in common.
- Your courage is the **attribute** I most admire in you.
- Spiders have many of the same anatomical **features** as insects.

You might respond by saying that these ways of talking must be misleading and should be abandoned – but then the onus is on you to suggest a way of expressing the same truths without using the vocabulary that commits you to the existence of universals.

Realists also say that these universals explain how **predication** works (predication = forming a sentence by applying a **predicate** such as “...is wise” to a noun or noun-phrase, such as “Socrates”). According to the realist, the reason we say something true when we say that Socrates is wise, is that the predicate “...is wise” refers to the universal *wisdom*, and the sentence (truly) asserts that Socrates “has”, “partakes in”, “shares in”, or “instantiates” this universal.

The first defender of a theory of universals was **Plato**, although (of course) he did not know the world “universal”; instead he tended to use Greek words which are standardly translated into English as “Form” (capital “F”) or “Idea”. Although it is clear that Plato’s Forms are essentially the same as the modern realist’s universals, the scholarly convention is to use the word “Form” when discussing Plato’s views, and reserve “universal” for talking about the modern debate.

Plato believed that the Forms are responsible for things in this “visible” world being the way they are: for example a white square is white and square because it shares in the Forms of whiteness and squareness. Similarly, he claimed, the Form of the Good is responsible for all the goodness in

the world; however, he denied that the Forms exist *within* the visible world – instead, he suggested, they inhabit a separate “realm” which is eternal and unchanging, and which can only be accessed by philosophers by means of a kind of mystical insight.

Modern philosophers who call themselves “platonists” (small “p”) defend the view that universals do not have a physical location, and exist eternally; however, they generally do not accept the idea of a separate “realm” or “Platonic heaven” in which universals exist. Instead, they say that the claim the universals do not have a location means precisely that they have *no* location, not that they have a location in some kind of separate realm, as Plato suggests.

In opposition to modern platonists are so-called “aristotelians” (small “a”), who claim that universals are part of the physical world, and have a physical location: so the universal *whiteness* is located in the physical world wherever its instances are. This view avoids the strangeness of believing in universals which exist despite not having a location to exist *at*. Although Plato may not have known that his student Aristotle would eventually defend this view, he was certainly aware of it, and criticized it in a passage of his dialogue *Parmenides*, which dramatizes an exchange between a young Socrates and an ageing pre-Socratic philosopher Parmenides:

“Socrates, he said, I admire the bent of your mind towards philosophy; tell me now, was this your own distinction between Forms in themselves and the things which partake of them? and do you think that there is a Form of likeness apart from the likeness which we possess, and of the one and many, and of the other things which Zeno mentioned?”

I think that there are such ideas, said Socrates.

....

But I should like to know whether you mean that there are certain Forms of which all other things partake, and from which they derive their names; that similars, for example, become similar, because they partake of similarity; and great things become great, because they partake of greatness; and that just and beautiful things become just and beautiful, because they partake of justice and beauty?

Yes, certainly, said Socrates that is my meaning.

Then each individual partakes either of the whole of the Form or else of a part of the Form? Can there be any other mode of participation?

There cannot be, he said.

Then do you think that the whole Form is one, and yet, being one, is in each one of the many?

Why not, Parmenides? said Socrates.

Because one and the same thing will exist as a whole at the same time in many separate individuals, and will therefore be in a state of separation from itself.

Nay, but the Form may be like the day which is one and the same in many places at once, and yet continuous with itself; in this way each idea may be one; and the same in all at the

same time.

I like your way, Socrates, of making one in many places at once. You mean to say, that if I were to spread out a sail and cover a number of men, there would be one whole including many-is not that your meaning?

I think so.

And would you say that the whole sail includes each man, or a part of it only, and different parts different men?

The latter.

Then, Socrates, the Forms themselves will be divisible, and things which participate in them will have a part of them only and not the whole Form existing in each of them?

That seems to follow.

Then would you like to say, Socrates, that the one Form is really divisible and yet remains one?

Certainly not, he said.

Suppose that you divide absolute greatness, and that of the many great things, each one is great in virtue of a portion of greatness less than absolute greatness-is that conceivable?

No.

Or will each equal thing, if possessing some small portion of equality less than absolute equality, be equal to some other thing by virtue of that portion only?

Impossible.

Or suppose one of us to have a portion of smallness; this is but a part of the small, and therefore the absolutely small is greater; if the absolutely small be greater, that to which the part of the small is added will be smaller and not greater than before.

How absurd!

Then in what way, Socrates, will all things participate in the Forms, if they are unable to participate in them either as parts or wholes?

Indeed, he said, you have asked a question which is not easily answered.

## **Nominalism**

**Nominalists** deny that universals exist; they claim that reality contains only **particulars**. A pressing question for the nominalist, then, is how to explain the function of the parts of language that the realist claims refer to universals, such as **general terms** and **predicates**. The realist claims that a predicate "... is white" applies to an object in virtue of a universal, *whiteness*, which that object "instantiates"; what explanation can the nominalist give? Here are two options:

- 1) All explanations have to stop somewhere; it is just a “brute fact” about the world that predicates apply to some things and not to others – this is not something that has to be explained in terms of mysterious entities such as universals that the predicates refer to.
- 2) Predicates apply in virtue of the **resemblances** between particulars; the predicate “... is white” applies to this object because it resembles all the other white things. (This view is often called **resemblance nominalism**.)

### Conceptualism

**Conceptualism** is an attempt to find a compromise between nominalism and realism. Conceptualists accept that predicates and general terms refer to *something*, but claim that what is referred to is not a mind-independent universal that would have existed even if there had been no creatures around to think about it, but rather a **concept** that depends for its existence on the existence of minds. Proponents of this view sometimes say that these concepts are **human constructs** rather than independently real. Whereas realists claim that universals “carve nature at the joints” – i.e. classify reality according to distinctions that already exist, independently of human thought – conceptualists claim that the division of reality into categories of similar objects is the result of our distinctively human ways of thinking or **conceptual scheme**; the world does not come to us with its own mind-independent structure, but rather we impose a structure on it by interpreting our experience according to a system of concepts which need not correspond to anything “out there” in reality.

This view might seem appealing, but it faces some serious objections. One is that it does not fully explain why our concepts apply to some objects and not to others. Surely the fact that these two objects are white cannot be explained only in terms of our decision to apply the concept of whiteness to them? Instead it might seem more plausible to say (as the nominalist does) that the objects are white *in virtue of* an objectively existing resemblance between them? Moreover, it seems more plausible for some concepts than others. We might accept that concepts such as *justice* are simply the result of our distinctively human ways of thinking; but it is not so easy to accept the view that “scientific” properties such as *mass*, *charge* or *being an electron* are “imposed” on reality by the activity of human thought, rather than counting as pre-existing features of reality that we *discover* instead of creating.

Finally, there is the question of the **ontological status** of these “concepts”. They are **mind-dependent**, in so far as they would not have existed if there had been no minds to think about them; but it does not seem like they exist *in* the minds of individual people, since they have to be shared: you and I share a concept of justice, for if we each had different concepts of justice then we could not discuss or debate it, as we would be talking past one another. So where are our concepts? If they exist in a separate “realm”, or even exist without having a physical location, then it seems that this view is every bit as strange and counter-intuitive as the platonic realism it was designed to replace.

*Homework:* read the following chapter from Russell, *The Problems of Philosophy* (1912), and answer the questions that follow.

## **CHAPTER IX – THE WORLD OF UNIVERSALS**

At the end of the preceding chapter we saw that such entities as relations appear to have a being which is in some way different from that of physical objects, and also different from that of minds and from that of sense-data. In the present chapter we have to consider what is the nature of this kind of being, and also what objects there are that have this kind of being. We will begin with the latter question.

The problem with which we are now concerned is a very old one, since it was brought into philosophy by Plato. Plato's 'theory of ideas' is an attempt to solve this very problem, and in my opinion it is one of the most successful attempts hitherto made. The theory to be advocated in what follows is largely Plato's, with merely such modifications as time has shown to be necessary.

The way the problem arose for Plato was more or less as follows. Let us consider, say, such a notion as justice. If we ask ourselves what justice is, it is natural to proceed by considering this, that, and the other just act, with a view to discovering what they have in common. They must all, in some sense, partake of a common nature, which will be found in whatever is just and in nothing else. This common nature, in virtue of which they are all just, will be justice itself, the pure essence the admixture of which with facts of ordinary life produces the multiplicity of just acts. Similarly with any other word which may be applicable to common facts, such as 'whiteness' for example. The word will be applicable to a number of particular things because they all participate in a common nature or essence. This pure essence is what Plato calls an 'idea' or 'form'. (It must not be supposed that 'ideas', in his sense, exist in minds, though they may be apprehended by minds.) The 'idea' justice is not identical with anything that is just: it is something other than particular things, which particular things partake of. Not being particular, it cannot itself exist in the world of sense. Moreover it is not fleeting or changeable like the things of sense: it is eternally itself, immutable and indestructible.

Thus Plato is led to a supra-sensible world, more real than the common world of sense, the unchangeable world of ideas, which alone gives to the world of sense whatever pale reflection of reality may belong to it. The truly real world, for Plato, is the world of ideas; for whatever we may attempt to say about things in the world of sense, we can only succeed in saying that they participate in such and such ideas, which, therefore, constitute all their character. Hence it is easy to pass on into a mysticism. We may hope, in a mystic illumination, to see the ideas as we see objects of sense; and we may imagine that the ideas exist in heaven. These mystical developments are very natural, but the basis of the theory is in logic, and it is as based in logic that we have to consider it.

The word 'idea' has acquired, in the course of time, many associations which are quite misleading when applied to Plato's 'ideas'. We shall therefore use the word 'universal' instead of the word 'idea', to describe what Plato meant. The essence of the sort of entity that Plato meant is that it is opposed to the particular things that are given in sensation. We speak of whatever is given in sensation, or is of the same nature as things given in sensation, as a particular; by opposition to this, a universal will be anything which may be shared by many particulars, and has those characteristics which, as we saw, distinguish justice and whiteness from just acts and white things.

When we examine common words, we find that, broadly speaking, proper names stand for particulars, while other substantives, adjectives, prepositions, and verbs stand for universals. Pronouns stand for particulars, but are ambiguous: it is only by the context or the circumstances that we know what particulars they stand for. The word 'now' stands for a particular, namely the present moment; but like pronouns, it stands for an ambiguous particular, because the present is always changing.

It will be seen that no sentence can be made up without at least one word which denotes a universal. The nearest approach would be some such statement as 'I like this'. But even here the word 'like' denotes a universal, for I may like other things, and other people may like things. Thus all truths involve universals, and all knowledge of truths involves acquaintance with universals.

Seeing that nearly all the words to be found in the dictionary stand for universals, it is strange that hardly anybody except students of philosophy ever realizes that there are such entities as universals. We do not naturally dwell upon those words in a sentence which do not stand for particulars; and if we are forced to dwell upon a word which stands for a universal, we naturally think of it as standing for some one of the particulars that come under the universal. When, for example, we hear the sentence, 'Charles I's head was cut off', we may naturally enough think of Charles I, of Charles I's head, and of the operation of cutting of his head, which are all particulars; but we do not naturally dwell upon what is meant by the word 'head' or the word 'cut', which is a universal. We feel such words to be incomplete and insubstantial; they seem to demand a context before anything can be done with them. Hence we succeed in avoiding all notice of universals as such, until the study of philosophy forces them upon our attention.

Even among philosophers, we may say, broadly, that only those universals which are named by adjectives or substantives have been much or often recognized, while those named by verbs and prepositions have been usually overlooked. This omission has had a very great effect upon philosophy; it is hardly too much to say that most metaphysics, since Spinoza, has been largely determined by it. The way this has occurred is, in outline, as follows: Speaking generally, adjectives and common nouns express qualities or properties of single things, whereas prepositions and verbs tend to express relations between two or more things. Thus the neglect of prepositions and verbs led to the belief that every proposition can be regarded as attributing a property to a single thing, rather than as expressing a relation between two or more things. Hence it was supposed that, ultimately, there can be no such entities as relations between things. Hence either there can be only one thing in the universe, or, if there are many things, they cannot possibly interact in any way, since any interaction would be a relation, and relations are impossible.

The first of these views, advocated by Spinoza and held in our own day by Bradley and many other philosophers, is called monism; the second, advocated Leibniz but not very common nowadays, is called monadism, because each of the isolated things is called a monad. Both these opposing philosophies, interesting as they are, result, in my opinion, from an undue attention to one sort of universals, namely the sort represented by adjectives and substantives rather than by verbs and prepositions.

As a matter of fact, if any one were anxious to deny altogether that there are such things as universals, we should find that we cannot strictly prove that there are such entities as qualities, i.e. the universals represented by adjectives and substantives, whereas we can prove that there must



be relations, i.e. the sort of universals generally represented by verbs and prepositions. Let us take in illustration the universal whiteness. If we believe that there is such a universal, we shall say that things are white because they have the quality of whiteness. This view, however, was strenuously denied by Berkeley and Hume, who have been followed in this by later empiricists. The form which their denial took was to deny that there are such things as 'abstract ideas'. When we want to think of whiteness, they said, we form an image of some particular white thing, and reason concerning this particular, taking care not to deduce anything concerning it which we cannot see to be equally true of any other white thing. As an account of our actual mental processes, this is no doubt largely true. In geometry, for example, when we wish to prove something about all triangles, we draw a particular triangle and reason about it, taking care not to use any characteristic which it does not share with other triangles. The beginner, in order to avoid error, often finds it useful to draw several triangles, as unlike each other as possible, in order to make sure that his reasoning is equally applicable to all of them. But a difficulty emerges as soon as we ask ourselves how we know that a thing is white or a triangle. If we wish to avoid the universals whiteness and triangularity, we shall choose some particular patch of white or some particular triangle, and say that anything is white or a triangle if it has the right sort of resemblance to our chosen particular. But then the resemblance required will have to be a universal. Since there are many white things, the resemblance must hold between many pairs of particular white things; and this is the characteristic of a universal. It will be useless to say that there is a different resemblance for each pair, for then we shall have to say that these resemblances resemble each other, and thus at last we shall be forced to admit resemblance as a universal. The relation of resemblance, therefore, must be a true universal. And having been forced to admit this universal, we find that it is no longer worthwhile to invent difficult and implausible theories to avoid the admission of such universals as whiteness and triangularity.

...

Having now seen that there must be such entities as universals, the next point to be proved is that their being is not merely mental. By this is meant that whatever being belongs to them is independent of their being thought of or in any way apprehended by minds. We have already touched on this subject at the end of the preceding chapter, but we must now consider more fully what sort of being it is that belongs universals.

Consider such a proposition as 'Edinburgh is north of London'. Here we have a relation between two places, and it seems plain that the relation subsists independently of our knowledge of it. When we come to know that Edinburgh is north of London, we come to know something which has to do only with Edinburgh and London: we do not cause the truth of the proposition by coming to know it, on the contrary we merely apprehend a fact which was there before we knew it. The part of the earth's surface where Edinburgh stands would be north of the part where London stands, even if there were no human being to know about north and south, and even if there were no minds at all in the universe. This is, of course, denied by many philosophers, either for Berkeley's reasons or for Kant's. But we have already considered these reasons, and decided that they are inadequate. We may therefore now assume it to be true that nothing mental is presupposed in the fact that Edinburgh is north of London. But this fact involves the relation 'north of', which is a universal; and it would be impossible for the whole fact to involve nothing mental if the relation

'north of', which is a constituent part of the fact, did involve anything mental. Hence we must admit that the relation, like the terms it relates, is not dependent upon thought, but belongs to the independent world which thought apprehends but does not create.

This conclusion, however, is met by the difficulty that the relation 'north of' does not seem to exist in the same sense in which Edinburgh and London exist. If we ask 'Where and when does this relation exist?' the answer must be 'Nowhere and nowhen'. There is no place or time where we can find the relation 'north of'. It does not exist in Edinburgh any more than in London, for it relates the two and is neutral as between them. Nor can we say that it exists at any particular time. Now everything that can be apprehended by the senses or by introspection exists at some particular time. Hence the relation 'north of' is radically different from such things. It is neither in space nor in time, neither material nor mental; yet it is something.

It is largely the very peculiar kind of being that belongs to universals which has led many people to suppose that they are really mental. We can think of a universal, and our thinking then exists in a perfectly ordinary sense, like any other mental act. Suppose, for example, that we are thinking of whiteness. Then in one sense it may be said that whiteness is 'in our mind'. We have here the same ambiguity as we noted in discussing Berkeley in Chapter IV. In the strict sense, it is not whiteness that is in our mind, but the act of thinking of whiteness. The connected ambiguity in the word 'idea', which we noted at the same time, also causes confusion here. In one sense of this word, namely the sense in which it denotes the object of an act of thought, whiteness is an 'idea'. Hence, if the ambiguity is not guarded against, we may come to think that whiteness is an 'idea' in the other sense, i.e. an act of thought; and thus we come to think that whiteness is mental. But in so thinking, we rob it of its essential quality of universality. One man's act of thought is necessarily a different thing from another man's; one man's act of thought at one time is necessarily a different thing from the same man's act of thought at another time. Hence, if whiteness were the thought as opposed to its object, no two different men could think of it, and no one man could think of it twice. That which many different thoughts of whiteness have in common is their object, and this object is different from all of them. Thus universals are not thoughts, though when known they are the objects of thoughts.

We shall find it convenient only to speak of things existing when they are in time, that is to say, when we can point to some time at which they exist (not excluding the possibility of their existing at all times). Thus thoughts and feelings, minds and physical objects exist. But universals do not exist in this sense; we shall say that they subsist or have being, where 'being' is opposed to 'existence' as being timeless. The world of universals, therefore, may also be described as the world of being. The world of being is unchangeable, rigid, exact, delightful to the mathematician, the logician, the builder of metaphysical systems, and all who love perfection more than life. The world of existence is fleeting, vague, without sharp boundaries, without any clear plan or arrangement, but it contains all thoughts and feelings, all the data of sense, and all physical objects, everything that can do either good or harm, everything that makes any difference to the value of life and the world. According to our temperaments, we shall prefer the contemplation of the one or of the other. The one we do not prefer will probably seem to us a pale shadow of the one we prefer, and hardly worthy to be regarded as in any sense real. But the truth is that both have the same claim on our impartial attention, both are real, and both are important to the metaphysician. Indeed no

sooner have we distinguished the two worlds than it becomes necessary to consider their relations. But first of all we must examine our knowledge of universals. This consideration will occupy us in the following chapter, where we shall find that it solves the problem of a priori knowledge, from which we were first led to consider universals.

- (1) According to Russell, what conclusion did Plato reach when trying to explain the similarity of the “multitude of just acts”?
- (2) Which kinds of words “stand for” universals, according to Russell?
- (3) Russell claims that even those who want to reject universals such as whiteness and triangularity are “forced to admit” at least one universal. Which one is it?
- (4) Why?
- (5) State – in one sentence – Russell’s reason for thinking that the relation “is North of” is part of the mind-independent world rather than existing in people’s minds.
- (6) What is the difference between “existing”, on the one hand, and “subsisting” or “having being” on the other?



## Week 9: Why believe in Universals?

The previous week's material offered a *prima facie* case for the existence of universals; however, the arguments that realists use to establish their conclusions are often subtle and complex. We'll consider four of the most important. These concern the nature of **sameness**, how **language** works, how we gain **concepts**, and how **a priori knowledge** is possible.

### The 'One over Many'

We know from Aristotle that the followers of Plato invoked an argument they called the "One over Many"; unfortunately he does not record the details of their argument. However, in recent debate the "One over Many" argument is usually formulated by asking one or other of these questions:

- How is it possible for two different things to be *the same* – i.e. to share the same nature?
- How is it possible for *one* general term (e.g. "yellow" or "cat") to apply to *many* different particulars?

The realist suggests that these questions can be answered by invoking the existence of universals. The first question is a question about **qualitative sameness** – about sameness "in kind"; it asks how it is possible for two **numerically different** things to be **qualitatively the same**. The realist responds that the two things are the "same" in the sense that they share a common property, and this property is a universal. So qualitative sameness (being "the same in some way") is explained in terms of numerical sameness ("sharing in one and the same universal"). As long as we understand what numerical sameness is, we can understand what qualitative sameness is, and this seems like a theoretical advance – one thing which we don't seem to understand is explained in terms of something that we do understand.

The realist gives a similar answer to the second question: all the things that fall under a general term do so because they all share in or "instantiate" the same universal; the reason the word "yellow" applies to many *different* things is that all these different things instantiate the universal *being yellow* – *yellowness*, if you will. This also explains why general terms apply to some particulars but not to others; the presence or absence of the relevant universal makes the difference between the general term's applying or not applying. It also explains why the application of general terms is not purely **subjective**: whether or not something is an electron, for example, depends on how the world is (whether the universal of *electronhood* is present), and not merely on whether we feel like applying the term "electron" from our own subjective perspective.

How should the nominalist respond to this argument? Perhaps the most popular answer is to say that **resemblance** is a fundamental feature of our understanding of reality, which does not have any further explanation in terms of universals. (Philosophers sometimes make this point by saying that resemblance is a **primitive** term in their theory.) To say that two particulars are "qualitatively the same" is merely to say that they resemble one another, and it is resemblance that explains the ability of one general term to apply to many particulars: what these particulars have in common is that they all resemble each other; what makes it the case that the term "yellow" apply to this particular object is that it is part of a **set** or **class** of objects, all of which resemble each other.

However, this nominalist explanation in terms of resemblance faces a notorious difficulty. According to the **resemblance nominalist**, a collection of things should be said to “have something in common” or “share a common property” if they form a set of objects, all of which resemble each other. But what about this collection of objects?

- Object A is red, square and heavy
- Object B is red, round, and light
- Object C is blue, round, and heavy

Each of these objects resembles each of the other two objects, but there is no single property which they all share. If the mere fact that a set of objects all resemble each other does not guarantee that they all share a common property, then this kind of mutual resemblance cannot explain what it is to share a common property; it seems that “having something in common” cannot be explained in terms of resemblance, but instead must be explained in terms of shared universals, as the realist suggests. This is known as the **imperfect community problem** and was first discovered by the American philosopher Nelson **Goodman**.

*Exercise:* can you think of a way to respond to the imperfect community problem on behalf of the nominalist?

### **Predication**

As we said last week, realists also argue that their theory provides a good explanation of how **predicates** work in human languages. More or less every sentence that is used to make a statement (as opposed to a command, question or exclamation) can be divided into one or more names or **subject-expressions**, and a **predicate** – the bit of the sentence that is left when the subject-expressions are removed. So the sentence “Socrates is wise” divides into a subject-expression “Socrates” and a predicate “... is wise”. Realists claim that, just as we explain the linguistic function of the name “Socrates” by saying that it refers to something in the world – namely, the particular human Socrates – so too should we explain the function of the predicate by saying that it refers to something in the world – namely, the universal *wisdom*. Otherwise we would be forced to say that over half the sentence makes a contribution to what is said by the sentence without being related to anything at all in the world. This point is eloquently made by David **Armstrong**. According to him, the nominalist gives predicates

“what has been said to be the privilege of the harlot: power without responsibility. The predicate is informative, it makes a vital contribution to telling us what is the case, the world is different if it is different, yet ontologically it is supposed not to commit us. Nice work: if you can get it.”

Armstrong, *Against Ostrich Nominalism*, p.443

Armstrong asks how a predicate can make a difference to what is said by a sentence if that predicate does not correspond to something in the world; how can it be that the predicate does not commit us “ontologically” - i.e. commit us to the existence of something (a universal) to which it refers?

Against this argument, nominalists sometimes claim that it is a mistake to say that words can make a contribution to the meaning of a sentence only if they pick out, or “refer to”, something in the

world. For example, the words “and”, “or”, and “not” make a contribution to the meaning of the sentences that contain them, although no-one seriously argues that there are universals of “and-ness” “or-ness” and “not-ness”. However, the realist can respond: the reason these words do not need to refer to anything in the world is that they are **truth-functional operators** whose contribution can be fully explained in terms of the contribution they make to the sentences they are applied to. For example, “not p” is true if and only if “p” is false, while “p or q” is true when at least one out of p and q is true. So these words do not need to have their meaning explained in terms of reference to something in the world. By contrast, the meaning of predicates is not so easily understood, and the realist will claim that the only way to understand the contribution a predicate makes is by associating it with a universal to which it refers.

### Concept acquisition

You should remember from last year that empiricists such as **Hume** and **Locke** worked hard to establish a theory of how we acquire general concepts on the basis of individual particular experiences. Locke claimed that perception “lets in” our “ideas” (concepts), while Hume claimed that “ideas” are copies of our “impressions” or sense-experience. Neither of them believed that we derive our concepts from independently existing universals. Thus Locke says:

“The next thing to be considered is,- How general words come to be made. For, since all things that exist are only particulars, how come we by general terms; or where find we those general natures they are supposed to stand for? Words become general by being made the signs of general ideas: and ideas become general, by separating from them the circumstances of time and place, and any other ideas that may determine them to this or that particular existence. By this way of abstraction they are made capable of representing more individuals than one; each of which having in it a conformity to that abstract idea, is (as we call it) of that sort.

...

To return to general words: it is plain, by what has been said, that general and universal belong not to the real existence of things; but are the inventions and creatures of the understanding, made by it for its own use, and concern only signs, whether words or ideas. Words are general, as has been said, when used for signs of general ideas, and so are applicable indifferently to many particular things; and ideas are general when they are set up as the representatives of many particular things: but universality belongs not to things themselves, which are all of them particular in their existence, even those words and ideas which in their signification are general. When therefore we quit particulars, the generals that rest are only creatures of our own making; their general nature being nothing but the capacity they are put into, by the understanding, of signifying or representing many particulars. For the signification they have is nothing but a relation that, by the mind of man, is added to them.”

Locke, *An Essay concerning Human Understanding*, III.3

The problem for the nominalist here is to explain *what* it is that we perceive in objects that enables us to form general concepts. A nominalist won't want to say that we form our concepts by noticing the *features* or *properties* of objects, since this suggests that objects have independently existing properties, which sounds a concession that universals exist after all – what else is a “property” of an

object if not a universal? Instead, the claim will have to be that we form our concepts by “abstracting” them from similarities between our experiences; but here the realist will say that we are able to classify objects as similar to one another only because we can recognize and distinguish the sameness of their properties: our ability to form concepts depends on our ability to classify objects, but our ability to classify objects depends on our ability to perceive the properties that these objects have. So it seems that we could not acquire our concepts unless we formed them in response to the presence of universals (properties) belonging to objects in the world.

Another, related, realist argument points out our ability to apply concepts to **objects I have not yet encountered**. Suppose I formed my concept of yellowness through exposure to a variety of yellow things – say a car, a door, a rubber duck. I can then go on confidently to apply that concept to any number of other yellow things, and distinguish between yellow and non-yellow things, even with things which resemble the objects I used to form my concept in no other way apart from being yellow. How can I be so sure that I am applying the concept correctly? The realist will say that I know I am using the concept correctly because I know which universal it is associated with, and am able to recognize that universal when I encounter it instantiated in objects I have not seen before. Otherwise, how could I be sure that I am classifying these new objects correctly?

This argument is sometimes criticised by appealing to the idea of **family resemblance concepts**, originating with **Wittgenstein**. He pointed out that some common concepts (e.g. the concept of a **game**) are such that there is no one feature in virtue of which they apply in every case, as there is no one feature that all games have in common:

“Compare chess with noughts and crosses. Or is there always winning and losing, or competition between players? Think of patience. In ball games there is winning and losing, but when a child throws his ball at the wall and catches it again, this feature has disappeared. Look at the parts played by skill and luck; and at the difference between skill in chess and skill in tennis... we see a complicated network of similarities overlapping and criss-crossing”

Wittgenstein, *Philosophical Investigations* s.66

So it seems that *some* concepts (“family resemblance” concepts, such as “game”) do not apply in virtue of any one feature shared by all their instances, and cannot be associated with one single universal “out there” in the world. However, this does not show that universals do not exist; the realist will say that Wittgenstein has shown merely that some concepts apply in virtue of a *family* of universals, not that universals play no role in explaining the application of our concepts to reality. After all, to make Wittgenstein’s point it is necessary to talk about the different “properties” of games, saying that some games share this property, and others share that property; thus it might be argued that the theory of “family resemblances” presupposes the existence of properties (universals) rather than refuting it.

### ***a priori* knowledge**

A final argument for belief in the existence of universals is that it enables us to explain how *a priori* knowledge is possible. If knowledge of things in the physical world is possible only through sense-experience, it can seem puzzling how we could ever know anything *a priori*, independently of sense-experience. But suppose that – as the platonist argues – universals are not located in the physical world, and are not apprehended through the senses, but rather lack a physical location and



are apprehended by “pure reason” or through the operation of the intellect. Then *a priori* knowledge is possible, as long as *a priori* knowledge is knowledge about universals, rather than knowledge of the physical world. This view was defended by Bertrand **Russell**:

“Returning now to the problem of a priori knowledge, which we left unsolved when we began the consideration of universals, we find ourselves in a position to deal with it in a much more satisfactory manner than was possible before. Let us revert to the proposition 'two and two are four'. It is fairly obvious, in view of what has been said, that this proposition states a relation between the universal 'two' and the universal 'four'. This suggests a proposition which we shall now endeavour to establish: namely, All *a priori* knowledge deals exclusively with the relations of universals. This proposition is of great importance, and goes a long way towards solving our previous difficulties concerning a priori knowledge.

...

Thus the statement 'two and two are four' deals exclusively with universals, and therefore may be known by anybody who is acquainted with the universals concerned and can perceive the relation between them which the statement asserts. It must be taken as a fact, discovered by reflecting upon our knowledge, that we have the power of sometimes perceiving such relations between universals, and therefore of sometimes knowing general *a priori* propositions such as those of arithmetic and logic. ...

It will serve to make the point clearer if we contrast our genuine a priori judgement with an empirical generalization, such as 'all men are mortals'. Here as before, we can understand what the proposition means as soon as we understand the universals involved, namely man and mortal. It is obviously unnecessary to have an individual acquaintance with the whole human race in order to understand what our proposition means. Thus the difference between an *a priori* general proposition and an empirical generalization does not come in the meaning of the proposition; it comes in the nature of the evidence for it. In the empirical case, the evidence consists in the particular instances. We believe that all men are mortal because we know that there are innumerable instances of men dying, and no instances of their living beyond a certain age. We do not believe it because we see a connexion between the universal man and the universal mortal. It is true that if physiology can prove, assuming the general laws that govern living bodies, that no living organism can last for ever, that gives a connexion between man and mortality which would enable us to assert our proposition without appealing to the special evidence of men dying. But that only means that our generalization has been subsumed under a wider generalization, for which the evidence is still of the same kind, though more extensive. The progress of science is constantly producing such subsumptions, and therefore giving a constantly wider inductive basis for scientific generalizations. But although this gives a greater degree of certainty, it does not give a different kind: the ultimate ground remains inductive, i.e. derived from instances, and not an *a priori* connexion of universals such as we have in logic and arithmetic.”

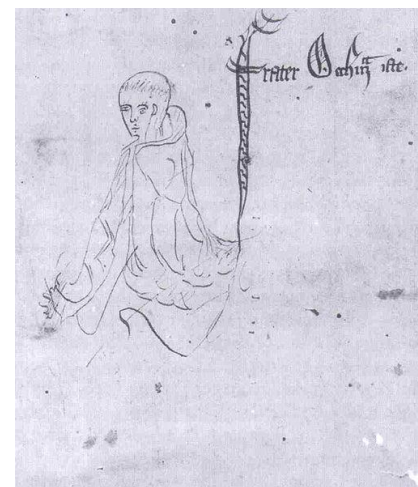
Russell, *The Problems of Philosophy*, chapter 10

*Homework:* "The existence of universals cannot be established by reason or experience and must therefore be rejected. Discuss."

## Week 10: Problems for the theory of Universals

### Ockham's Razor

William of **Ockham** (1288-1347) supposedly appealed to the principle subsequently known as “Ockham's Razor” (or “Occam's Razor”) to support his rejection of universals. Although it never appears in his extant writings in this form, it is usually quoted as the principle **entia non sunt multiplicanda praeter necessitatem**, or “entities should not be multiplied beyond necessity”. This is also sometimes described as the **principle of parsimony**. It states that we should not introduce more things to our theory than we need to: all other things being equal, we should prefer the theory that requires us to believe in the smallest number of entities. Ockham's Razor is sometimes (mis)described as the principle that we should prefer the *simpler* of two competing theories, but this misses the point that the original use of the principle was to limit the number of *entities* or kinds of entity introduced by a theory.



In its original form – as a principle about ontological parsimony (preferring the theory that is committed to the existence of fewer things) – the principle is intuitively plausible. For example, atheists (and some Christians) accept that if everything in the world can be explained by a theory that does not include the existence of God, then it follows that we should not believe that God exists; this is part of the reason why both sides of the debate take such an interest in the explanatory power of the theory of evolution. It might seem, then, that the nominalist can appeal to Ockham's Razor to justify the rejection of the existence of universals. After all, the nominalist's theory accepts the existence of *one* kind of entity (particulars), while the realist accepts the existence of *two* kinds (particulars *and* universals).

However, the nominalist's appeal to Ockham's Razor faces a serious problem: the principle says only that entities should be excluded from our theory when they are *not* needed by the theory; if removing those entities will limit the explanatory power of the theory then they are in a sense “needed” rather than “superfluous”. No-one thinks that we improve a theory by reducing the number of entities it is committed to, regardless of the damage that would do to the theory's ability to explain things – that would be like saying that you could “improve” physics by removing its commitment to sub-atomic particles even though that would leave physics unable to explain what it sets out to explain. We have already seen how the realist about universals claims that belief in the existence of universals explains phenomena such as **qualitative sameness**, **predication**, and **a priori knowledge**; if excluding universals from our theory results in an inability to explain any of these things, then the nominalist cannot claim that universals are a superfluous addition to the theory, and so cannot argue that they should be excluded on the basis of Ockham's Razor.

## The location problem

We have already seen that realists disagree about where universals are located, with “platonists” claiming that they do not have a physical location, but exist unlocated or in some other “realm of universals”, while “aristotelians” claim that universals are located (and wholly present) where their instances are. Both theories are attacked by nominalists. First, if universals are located at their instances, they have to be wholly present in two places at the same time. Obviously a normal physical object can occupy more than one place, by have different *parts* in different locations; but universals are not supposed to have parts – a universal is *one* unified thing which accounts for the similarity between many different particulars. Moreover, the nominalist argues, the “platonist” idea that universals do not have a physical location is even more ridiculous: how can something *exist* without existing *somewhere*?

However, the realist can respond to these complaints: what the nominalist is doing here, it seems, is taking rules that govern the behaviour of *particulars*, such as “no two particulars can be in the same place at the same time” and “every particular has a physical location”, and trying to make those rules apply to universals as well. The realist can say that it is no surprise that the rules that apply to particulars do *not* apply to universals: according to the realist, universals are a fundamentally different kind of thing from particulars, and we should not expect both kinds of entity to obey the same “rules” about how and where they can be located.

## Empiricism

A common complaint against the theory of universals is that the existence of universals is **empirically unverifiable**, i.e. that we cannot say what kind of experience would count as evidence that universals exist, and so cannot know that they are part of reality. (As we’ll see next week, a particularly strong version of this criticism says that, because claims like “universals exist” are unverifiable, they are in fact meaningless – literally nonsense.)

Realists can adopt a variety of strategies to deal with this criticism. One is to say that our knowledge of universals is *a priori*, and does not need to be justified by appeal to experience. This is the approach endorsed by Bertrand **Russell** in the selection quoted earlier; however realists then face the problem of explaining the mysterious mental “faculty” that enables them to grasp the nature of universals without consulting any kind of sense-experience.

A second strategy has become more popular in recent debate. This is to suggest that belief in universals is a form of **inference to the best explanation**. The idea is that, although we may not directly *experience* universals, nevertheless belief in universals enables us to *explain* many features of our experience (e.g. qualitative sameness) which would otherwise be puzzling. The realist will point out that belief in an entity can be “empirically justified” not merely because we can see it, but also because including that entity in our theory enables us to explain aspects of our experience. There is an apparent similarity with electrons and sub-atomic particles here: although no-one has ever “seen” these entities, they allow us to explain features of what we *do* experience (e.g. the readings from the sensors at the Large Hadron Collider) and so belief in their existence is empirically justified on the basis of “inference to the best explanation”, just as the realist claims belief in universals is justified.

## The “Third Man”

One of the earliest objections to the existence of universals was discussed by Plato in his dialogue *Parmenides*. See if you can understand what the problem is:

“Well, said Parmenides, and what do you say of another question?

What question?

I imagine that the way in which you are led to assume one Form of each kind is as follows: you see a number of large objects, and when you look at them there seems to you to be one and the same Form (or nature) in them all; hence you conceive of largeness as one.

Very true, said Socrates.

And if you go on and allow your mind in like manner to embrace in one view the Form of largeness and of large things which are not the Form, and compare them, will not another largeness arise, which will appear to be the source of all these?

It would seem so.

Then another Form of largeness now comes into view over and above absolute largeness, and the individuals which partake of it; and then another, over and above all these, by virtue of which they will all be large, and so each Form instead of being one will be infinitely multiplied.

But may not the Forms, asked Socrates, be thoughts only, and have no proper existence except in our minds, Parmenides? For in that case each Form may still be one, and not experience this infinite multiplication.

And can there be individual thoughts which are thoughts of nothing?

Impossible, he said.

The thought must be of something?

Yes.

Of something which is or which is not?

Of something which is.

Must it not be of a single something, which the thought recognizes as attaching to all, being a single form or nature?

Yes.

And will not the something which is apprehended as one and the same in all, be an Form?

From that, again, there is no escape.

Then, said Parmenides, if you say that everything else participates in the Forms, must you not say either that everything is made up of thoughts, and that all things think; or that they are thoughts but have no thought?

The latter view, Parmenides, is no more rational than the previous one. In my opinion, the Forms are, as it were, patterns fixed in nature, and other things are like them, and resemblances of them-what is meant by the participation of other things in the Forms, is really assimilation to them.

But if, said he, the individual is like the Form, must not the Form also be like the individual, in so far as the individual is a resemblance of the Form? That which is like, cannot be conceived of as other than the like of like.

Impossible.

And when two things are alike, must they not partake of the same Form?

They must.

And will not that of which the two partake, and which makes them alike, be the Form itself?

Certainly.

Then the Form cannot be like the individual, or the individual like the Form; for if they are alike, some further Form of likeness will always be coming to light, and if that be like anything else, another; and new Forms will be always arising, if the Form resembles that which partakes of it?"

Now fill in the gaps to provide an explanation of Parmenides' argument against the Forms:

1. Whenever some particulars are large, that is because they all share a \_\_\_\_\_
2. The Forms are like patterns fixed in nature and particulars \_\_\_\_\_ them.
3. Therefore, the Form of largeness must be \_\_\_\_\_ .
4. So there must be a new Form over largeness which is set over both the \_\_\_\_\_ and the original \_\_\_\_\_ .
5. This reasoning can be repeated endlessly, so instead of having one Form we have an \_\_\_\_\_ of Forms of largeness.

This argument depends on the realist accepting the view that particulars *resemble* the Forms that they instantiate; it is only because Socrates in the dialogue accepts that the Form of largeness must itself be large that we need to introduce a second Form of largeness to account for the largeness which the original Form shares with the plurality of large particulars. Since most modern realists deny that instantiation (the link between particular and universal) should be explained in terms of resemblance, the "Third Man" argument is not often directly appealed to by modern nominalists. However, it does pose a genuine problem for the realist: if instantiation cannot be explained in terms of resemblance, how can it be explained? That question forms the basis of our final problem for realism.

## Explaining Instantiation

You will remember that realists believe that universals function not just as the properties that particular objects have, but also as the *relations* that objects stand in. So relational predicates such as "... loves ...", which have to be completed by two or more particulars, stand for *relational universals*. But what should the realist say about the relational predicate "... instantiates ...", for example in the sentence "Particular *a* instantiates universal *F*"? It seems that she ought – for consistency's sake – accept that this predicate stands for a relation universal of *instantiation*. But then we have a situation in which we should say that "particular *a* and universal *F* instantiate the universal *instantiation*". And this sentence *also* uses a relational predicate "... instantiates ...", which it seems should stand for a further relational universal, *instantiation\**. But if we try to describe the situation now, in which particular *a* and universals *I* and *F* instantiate *instantiation\**, we have yet another relational predicate, which requires yet another relational universal, *instantiation\*\**, and we come to see that, however long we go on we will never come to the end of the list of universals involved when *a* instantiates *F*.

This argument is known as **Bradley's Regress**, after the Victorian philosopher F. H. **Bradley** who first proposed it. Bradley believed he had shown that belief in the existence of relational universals is untenable:

"we are hurried off into the eddy of a hopeless process, since we are forced to go on finding new relations without end. The links [relations] are united by a link, and this bond of union is a link which also has two ends; and these require each a fresh link to connect them with the old. The problem is to find how the relation can stand to its qualities, and this problem is insoluble. If you take the connection as a solid thing, you have got to show, and you cannot show, how the other solids are joined to it. And if you take it as a kind of medium or unsubstantial atmosphere, it is a connection no longer." Bradley,  
*Appearance and Reality* p.33

There is considerable debate about how precisely this regress argument should be understood, but one interpretation is that the realist cannot explain what is responsible for the combination of particular and universal: what is it that "glues" a particular to the universal it instantiates? It cannot be that the gluing is done by the relational universal *instantiation*, since this in turn will need "gluing" to the particular and universal which it relates; and adding more relational universals simply provides more things that need to be glued together and never any actual "metaphysical glue". This interpretation was proposed by C.D. Broad early in the 20<sup>th</sup> century:

"It is plain that Bradley thinks of A and B as being like two objects fastened together with a bit of string, and he thinks of [relation] R as being like the string. He then remembers that the objects must be glued or sealed to both ends of the bit of string if the latter is to fasten them together. And then, I suppose, another kind of glue is needed to fasten the first drop of glue to the object A on the one side and to the bit of string on the other; and another kind of glue is needed to fasten the second drop of glue to the object B on the one side and to the string on the other. And so on without end." C. D. Broad, *An Examination of McTaggart's Philosophy*, p.85

One response to this problem is to reject the assumption on which it is based: namely, that we

need to explain what is “responsible” for the combination of particular and universal in instantiation. Perhaps the realist can say that no further explanation of instantiation is necessary – it is just a “brute fact” about the world that particulars instantiate some universals and not others. However, this leaves the realist with the problem that she has to recognize one relational predicate – “... instantiates ...” – which does not stand for a relational universal. And if *one* predicate can function meaningfully in a sentence without standing for a universal, why can’t *every* predicate function meaningfully without standing for a universal? By conceding that one predicate can function without referring to a universal it seems like the realist is in danger of conceding that we do not, after all, need universals to explain how predicates work, and this threatens to play into the hands of the nominalists, and that is what the nominalist was claiming all along.

*Homework:* “The theory of universals raises more problems than it solves. Discuss.”

*Extension material:* Peacock, “Bradley’s Regress, Truthmaking, and Constitution” (draft copy online at

[http://ucl.academia.edu/HowardPeacock/Papers/106139/Bradleys\\_Regress\\_Truthmaking\\_and\\_Constitution](http://ucl.academia.edu/HowardPeacock/Papers/106139/Bradleys_Regress_Truthmaking_and_Constitution) )



## Week 11: The status of metaphysics: verification and falsification

### What is metaphysics?

Metaphysics is the branch of philosophy that deals with the **ultimate nature of reality** – the branch of philosophy which attempts to find out **how things really are** rather than merely how they seem to us. Metaphysicians investigate the nature of phenomena such as causation, time, and existence; in the history of philosophy there has been a tendency to propose metaphysical theories which seem in drastic opposition to common sense. Notable among these are:

- Plato's view that there is a separate eternal unchanging realm of perfect **Forms**
- The **Idealism** of Bishop **Berkeley**, who argued that the only things that exist are perceptions or "ideas", and God himself.
- The **monism** of F. H. **Bradley**, who asserted that there was in reality only one thing, the "Absolute".
- The **transcendental idealism** of **Kant**, who claimed that there is an unknowable "noumenal" world beyond the "phenomenal" world known through sense-experience.

However, in the first half of the 20<sup>th</sup> century a view gained popularity among philosophers according to which these kinds of metaphysical claims are simply **speculative nonsense** with no **factual content**, which do not succeed in making any genuinely meaningful claim about the world. This view was known as **logical positivism** and originated in the 1920s and 30s in a small discussion group known as the **Vienna Circle**. It was later popularized throughout the English-speaking world via A. J. **Ayer's** book *Language, Truth, and Logic*.

### Verificationism

The central tenet of the **logical positivist** view was the so-called **verification principle**. This states that any meaningful sentence that is not *analytic* (true solely in virtue of its meaning) must be such that we can understand what kinds of experience would count towards the truth of what it says. Any non-analytic sentence such that we are unable to say what would count as "verifying it" in that sense is to be dismissed as meaningless – literally nonsense. So, for example, because we cannot understand what kind of experience would count as verifying Kant's claim that there exists a "noumenal" world beyond sense-experience, we have to dismiss that claim as nonsense. Rather more controversially, A. J. Ayer suggested that the claim "God exists" is neither true nor false but literally meaningless, because we could not explain what kind of experience would count as verifying it. (This enabled him to irritate those who wanted to label him as an "atheist" – of course Ayer could no more be an atheist than he could be a Christian, as according to him "God does not exist" is just as meaningless as "God exists".)

The logical positivists debated at great length how precisely to formulate the verification principle, but the most frequently cited version is that proposed by Ayer. This required only that we understand how to verify a sentence *in principle*, rather than requiring that we have the practical ability to verify it from our own situation – a modification necessary to allow sentences such as "there are no human beings on Andromeda" to count as "verifiable". It also said that experiences

could count as “verifying” a sentence if they counted as strong *evidence* towards the truth of that sentence; there was no requirement that we should know how to *prove* conclusively that a sentence is true. This modification was introduced to ensure that universal generalizations such as “all men are mortal” could count as verifiable even though it can never be proven absolutely, since it covers a potentially infinite number of cases. Verificationism posed a serious threat to traditional metaphysics; however, it has now fallen almost entirely out of fashion as a result of the strength of criticisms levelled against it:

First, it seems that the verification principle is self-refuting in much the same way as the sentence “this sentence is false”. The principle says that any non-analytic sentence which cannot be verified through experience is meaningless. But the sentence “any non-analytic sentence which cannot be verified through experience is meaningless” is apparently not analytic – it is not true solely in virtue of the meanings of the words involved – and it is hard to see what kinds of experience could count as verifying it. So the very sentence used to state the verification principle turns out to be literally meaningless! It seems the only way in which the statement of the verification principle could count as meaningful would be if it was *false*. Since the principle certainly seems meaningful, we seem to be justified in deciding that the principle is false.

Second, the verification principle conflicts with the view expressed by **Quine**, that it is a mistake to think of experience verifying, or counting as evidence for, single sentences within a theory. According to Quine, any experience is compatible with any single element of our theory provided we are willing to make the necessary adjustments to the rest of the theory to accommodate it. So it is a mistake to search for evidence which would “verify” a single sentence in the sense intended by Ayer; instead our theory gets “verified” or confirmed **holistically**, by how good a job it does predicting and explaining the totality of our experience. As Quine says,

“our statements about the external world face the tribunal of sense experience not individually but only as a corporate body”

If that is right, it is a mistake to criticize a sentence as meaningless because we cannot see what evidence would count towards verifying it in isolation; it may be that the sentence can only be “verified” by being part of a theory which does a good job overall.

*Exam tip:* the vast majority of A2 students write as though the verification principle is something that is obviously true – despite the fact that it has now been abandoned by the vast majority of philosophers. Remember to approach this principle (like everything else) with an attitude of *critical evaluation*.

*Extra material :* watch the interview with A. J. Ayer at <http://tinyurl.com/MageeAyer>

## Falsificationism

An alternative approach to the confirmation of theories was proposed by the philosopher of science Karl **Popper**. His view was that we should not demand that a theory should be capable of being “verified” by evidence that counted in its favour; instead the test should be that the theory is in principle capable of being “falsified” or “disconfirmed” by evidence that counts against it. This view is known as **falsificationism**. For example, I might have a theory that all swans are white. While it is not within my power to prove conclusively that this theory is true, it is certainly a possibility that I could encounter evidence that shows that my theory is *false* – just one encounter with a black swan would be enough to do that.

From *Science as Falsification* - Karl Popper

The problem which troubled me ... was neither, "When is a theory true?" nor "When is a theory acceptable?" my problem was different. I wished to distinguish between science and pseudo-science...

I knew, of course, the most widely accepted answer to my problem: that science is distinguished from pseudoscience—or from "metaphysics"—by its empirical method, which is essentially inductive, proceeding from observation or experiment. But this did not satisfy me...

Among the theories which interested me Einstein's theory of relativity was no doubt by far the most important. The three others were Marx's theory of history, Freud's psycho-analysis, and Alfred Adler's so-called "individual psychology."

There was a lot of popular nonsense talked about these theories, and especially about relativity (as still happens even today), but I was fortunate in those who introduced me to the study of this theory. We all—the small circle of students to which I belong—were thrilled with the result of Eddington's eclipse observations which in 1919 brought the first important confirmation of Einstein's theory of gravitation. It was a great experience for us, and one which had a lasting influence on my intellectual development...

My problem perhaps first took the simple form, "What is wrong with Marxism, psycho-analysis, and individual psychology? Why are they so different from physical theories, from Newton's theory, and especially from the theory of relativity?" ...

The most characteristic element in this situation seemed to me the incessant stream of confirmations, of observations which "verified" the theories in question; and this point was constantly emphasize by their adherents. A Marxist could not open a newspaper without finding on every page confirming evidence for his interpretation of history; not only in the news, but also in its presentation — which revealed the class bias of the paper — and especially of course what the paper did not say. The Freudian analysts emphasized that their theories were constantly verified by their "clinical observations." ... it turned out that the theories in question were compatible with the most divergent human behaviour, so that it was practically impossible to describe any human behaviour that might not be claimed to be a verification of these theories.

These considerations led me in the winter of 1919-20 to conclusions which I may now

reformulate as follows:

- (1) It is easy to obtain confirmations, or verifications, for nearly every theory — if we look for confirmations.
  - (2) Confirmations should count only if they are the result of risky predictions; that is to say, if, unenlightened by the theory in question, we should have expected an event which was incompatible with the theory — an event which would have refuted the theory.
  - (3) Every "good" scientific theory is a prohibition: it forbids certain things to happen. The more a theory forbids, the better it is.
  - (4) A theory which is not refutable by any conceivable event is non-scientific. Irrefutability is not a virtue of a theory (as people often think) but a vice.
  - (5) Every genuine test of a theory is an attempt to falsify it, or to refute it. Testability is falsifiability; but there are degrees of testability: some theories are more testable, more exposed to refutation, than others; they take, as it were, greater risks...
- One can sum up all this by saying that the criterion of the scientific status of a theory is its falsifiability, or refutability, or testability."

Popper himself accepted that there could be "unfalsifiable" statements that were nevertheless meaningful; his test was designed precisely to distinguish "unscientific" (but meaningful) theories such as Marxism from testable "scientific" ones. However, this approach has been developed into a way of distinguishing meaningful from meaningless statements by Anthony **Flew**:

"Let us begin with a parable. It is a parable developed from a tale told by John Wisdom in his haunting and revolutionary article "Gods." [1] Once upon a time two explorers came upon a clearing in the jungle. In the clearing were growing many flowers and many weeds. One explorer says, "Some gardener must tend this plot." The other disagrees, "There is no gardener." So they pitch their tents and set a watch. No gardener is ever seen. "But perhaps he is an invisible gardener." So they set up a barbed-wire fence. They electrify it. They patrol with bloodhounds. (For they remember how H. G. Wells's *The Invisible Man* could be both smelt and touched though he could not be seen.) But no shrieks ever suggest that some intruder has received a shock. No movements of the wire ever betray an invisible climber. The bloodhounds never give cry. Yet still the Believer is not convinced. "But there is a gardener, invisible, intangible, insensible, to electric shocks, a gardener who has no scent and makes no sound, a gardener who comes secretly to look after the garden which he loves." At last the Sceptic despairs, "But what remains of your original assertion? Just how does what you call an invisible, intangible, eternally elusive gardener differ from an imaginary gardener or even from no gardener at all?" [...]

And in this, it seems to me, lies the peculiar danger, the endemic evil, of theological utterance. Take such utterances as "God has a plan," "God created the world," "God loves us as a father loves his children." They look at first sight very much like assertions, vast cosmological assertions. Of course, this is no sure sign that they either are, or are intended to be, assertions. But let us confine ourselves to the cases where those who utter such sentences intended them to express assertions. [...]

Now to assert that such and such is the case is necessarily equivalent to denying that such

and such is not the case. Suppose then that we are in doubt as to what someone who gives vent to an utterance is asserting, or suppose that, more radically, we are sceptical as to whether he is really asserting anything at all, one way of trying to understand (or perhaps to expose) his utterance is to attempt to find what he would regard as counting against, or as being incompatible with, its truth. For if the utterance is indeed an assertion, it will necessarily be equivalent to a denial of the negation of the assertion. And anything which would count against the assertion, or which would induce the speaker to withdraw it and to admit that it had been mistaken, must be part of (or the whole of) the meaning of the negation of that assertion. And to know the meaning of the negation of an assertion, is as near as makes no matter, to know the meaning of that assertion. And if there is nothing which a putative assertion denies then there is nothing which it asserts either: and so it is not really an assertion. When the Sceptic in the parable asked the Believer, "Just how does what you call an invisible, intangible, eternally elusive gardener differ from an imaginary gardener or even from no gardener at all?" he was suggesting that the Believer's earlier statement had been so eroded by qualification that it was no longer an assertion at all."

Flew, "Theology and Falsification"

*Question:* what is Flew's reason for saying that we understand a sentence only if we understand what we would "regard as counting against its truth"?

However, Flew's falsificationist approach faces the same objections as faced the verificationist: first, the sentence "All unfalsifiable sentences are meaningless" seems to be such that no possible experience could falsify it, suggesting (paradoxically) that the very statement of the falsificationist's principle is meaningless. Second, Quine's point about **holistic confirmation** still stands: there is no such thing as an experience that definitively *proves* a sentence is false, because any sentence within a theory can be preserved providing we make changes to other parts of the theory. Suppose I believe that all swans are white, then have the experience of seeing a black swan. Does this definitively falsify my theory? Not if I am prepared to believe that I am having a visual hallucination, or dreaming, or that there is a species of black animal which looks very much like a swan from a distance but isn't *really* a swan. So it seems that falsificationism has no more chance than verificationism did of showing that "metaphysical" statements are meaningless.

### **The impossibility of eliminating metaphysics**

An assumption of the preceding discussion is that there is a significant distinction in kind between the "metaphysical" statements of philosophy, and the "factual" statements of science. This assumption has been challenged, particularly in the work of **Quine**. His view is that

"Ontological questions [questions about what exists] are on a par with questions of natural science."

According to Quine, the difference between science and metaphysics is simply a matter of degree: scientific statements tend to be more directly related to empirical evidence, while philosophical theories tend to be confirmed more indirectly, by how well they "fit" with the rest of our theory. But there is no sharp division between the two: some scientific theories (for example the idea of

“multiverses”) are based almost entirely on thought-experiments rather than empirical evidence.

Moreover, Quine argued that all scientific theories have consequences for our metaphysics, since every theory has **ontological commitments**: i.e. whichever theory we accept commits us to the existence of entities required for the truth of the theory, just as modern physics commits us to the existence of electrons. Quine’s view is that, given that every scientific theory has metaphysical consequences, it is part of the job of philosophers to work out what those consequences are – for example, finding out whether the ontological commitments of a theory include universals, abstract mathematical objects such as sets and numbers, or even physical objects. Thus Quine suggests that metaphysics – far from being a separate branch of “nonsense” that no scientist should pay attention to – in fact forms part of a “total theory” which is **continuous** with science and is confirmed holistically by how well it makes sense of the totality of our experience. As he says:

“Our acceptance of an ontology, is, I think, similar in principle to our acceptance of a scientific theory, say a system of physics: we adopt, at least insofar as we are reasonable, the simplest conceptual scheme into which the disordered fragments of raw experience can be fitted and arranged. Our ontology is determined once we have fixed upon the over-all conceptual scheme which is to accommodate science in the broadest sense, and the considerations which determine a reasonable construction of any part of that conceptual scheme, for example, the biological or the physical part, are not different in kind from the considerations which determine a reasonable construction of the whole.”

Quine, *On What there Is*, p.16-17

*Homework:* Read Chapter 1 of A. J. Ayer’s *Language, Truth, and Logic*, and answer the following questions:

- (1) The metaphysician would deny “that his assertions were ultimately based on the evidence of his senses”. What would he say instead?
- (2) What have those trying to describe “a reality transcending the limits of all possible sense-experience” ultimately produced?
- (3) What is “the criterion of verifiability”?
- (4) Suppose that a sentence is not “literally significant”. What other kind of significance can it have for a person?
- (5) What is the difference between “practical verifiability” and “verifiability in principle”?
- (6) What is “conclusive verifiability”? Why does Ayer reject it as part of his criterion (give examples)?
- (7) What is the nature of the debates between (i) monists and pluralists? (ii) (metaphysical) realists and idealists?
- (8) What is a “metaphysical sentence”, according to Ayer?
- (9) What does “the postulation of real non-existent entities” result from?
- (10) In what way are poets better than metaphysicians?

## Week 12: Objective knowledge vs. relativism

### “Objective” and “Absolute” vs. “Relative” and “Subjective”

Knowledge (and truth) that is “objective” and “absolute” is often contrasted with knowledge that is “relative” or “subjective”. Knowledge is said to be **subjective** if its correctness depends on the perspective or point of view of the individual knower; if the correctness depends on the standards current within the knower’s community, then this knowledge would be better described as **relative** rather than merely **subjective**. If we’re being precise, **objective** means “not subjective”, and **absolute** means “not relative”. However, in practice the term “objective” is used to describe anything that is neither relative nor subjective, and this is a convention I’ll follow.

The mainstream discussion of knowledge, truth, and justification presupposes that all three are **objective** notions – i.e. that whether a subject’s beliefs are true, are justified, or count as knowledge, is *not* relative to or dependent on the standards current within her society or culture. Defenders of this “objective” conception of knowledge point out that the concept of knowledge is intended to capture something like getting something right for the right reasons, and there must be a distinction between merely *thinking* you have got it right, and *really* getting it right. But the idea of “really” getting it right is the idea of getting it right *no matter what anyone thinks*, which is the idea of objective truth – truth which does not depend for its correctness on the opinions or values of the people around you.

However, this seemingly intuitive view has been challenged from a variety of perspectives – within postmodernist literary and cultural theory, within sociology, as well as from within philosophy itself. It is now commonplace to hear cultural theorists talking about knowledge as something that is “socially located”, treating “knowing” as something that depends on the cultural standards relevant within a community. Indeed, some theorists allow that “knowledge” in one culture at one time may be in direct contradiction with “knowledge” in other cultures and times, saying that people in medieval Europe “knew” that the Sun goes round the Earth while we “know” that the solar system works the other way round.

### Global relativism and subjectivism about truth

The most straightforward way to deny objectivity is to claim that all truth is relative, dependent for its correctness on the world-view of a particular society, or even to claim that all truth is subjective, depending on the perspective of the individual believer: we should no longer talk about things being “true” absolutely, but only about things being “true for me” or “true for you”. This view is associated with the sophist **Protagoras**, whose reportedly claimed that

“man is the measure of all things, of things that are that they are, and that things that are not, that they are not”.

According to Protagoras, for something to be *perceived* as true is for something to *be* true for that individual. This view was criticized by Plato in his dialogue *Theaetetus*:

“*Socrates*: Then you were quite right in affirming that knowledge is only perception; and the meaning turns out to be the same, whether with Homer and Heracleitus, and all that company, you say that all is motion and flux, or with the great sage Protagoras, that man is the measure of all things; or with Theaetetus, that, given these premises, perception is

knowledge.

...

Soc. Shall I tell you, Theodorus, what amazes me in your acquaintance Protagoras?

Theod. What is it?

Soc. I am charmed with his doctrine, that what appears to be, is true to each one, but I wonder that he did not begin his book on Truth with a declaration that a pig or a dog-faced baboon, or some other yet stranger monster which has sensation, is the measure of all things; then he might have shown a magnificent contempt for our opinion of him by informing us at the outset that while we were reverencing him like a God for his wisdom he was no better than a tadpole, not to speak of his fellow-men-would not this have produced an over-powering effect? For if truth is only sensation, and no man can discern another's feelings better than he, or has any superior right to determine whether his opinion is true or false, but each, as we have several times repeated, is to himself the sole judge, and everything that he judges is true and right, why, my friend, should Protagoras be preferred to the place of wisdom and instruction, and deserve to be well paid, and we poor ignoramuses have to go to him, if each one is the measure of his own wisdom? Must he not be talking *ad captandum* in all this? I say nothing of the ridiculous predicament in which my own midwifery and the whole art of dialectic is placed; for the attempt to supervise or refute the notions or opinions of others would be a tedious and enormous piece of folly, if to each man his own are right; and this must be the case if Protagoras Truth is the real truth, and the philosopher is not merely amusing himself by giving oracles out of the shrine of his book."

The criticism put forward by Plato here is simple but damaging: if Protagoras is right, and everyone's opinion is "true for them", then why should anyone listen to anyone else's opinion; and in particular, why should anyone listen to Protagoras' opinion that "man is the measure of all things"? If what Protagoras says is true, it is at best "true for him" but need not be "true for me" – unless I also happen to believe that it is true.

A related criticism is the so-called "paradox of relativism": that the relativist's claim "everything is relative" is self-defeating, because the relativist is putting this claim forward as an *absolute* truth while at the same time denying that there are any such absolute truths. It cannot be that the relativist is proposing that "everything is relative" is only true relative to his *own* particular set of standards – for the relativist wants to claim that there are no objective standards for correctness in *any* culture; otherwise the relativist would be committed to saying that the standards for correctness are culturally relative in cultures that believed in culturally relative standards, but that there are objective standards for correctness that apply in cultures that believe in objectively correct standards, and this just seems incoherent – either there *are* objectively correct standards for truth that apply in every situation, regardless of culture, or there aren't!

A similar relativist view to Protagoras' is found in **Nietzsche's** writings from the end of the 19<sup>th</sup> century. He proposed a view called **perspectivism**, according to which it is impossible for any thinker to escape the "prejudices" inherent in their own perspective or point of view, and so there is no such thing as an "objective" truth – only truth from one or another perspective.



“That which causes philosophers to be regarded half-distrustfully and half-mockingly, is not the oft-repeated discovery how innocent they are – how often and easily they make mistakes and lose their way, in short, how childish and childlike they are – but that there is not enough honest dealing with them, whereas they all raise a loud and virtuous outcry when the problem of truthfulness is even hinted at in the remotest manner. They all pose as though their real opinions had been discovered and attained through the self-evolving of a cold, pure, divinely indifferent dialectic (in contrast to all sorts of mystics, who, fairer and foolisher, talk of “inspiration”), whereas, in fact, a prejudiced proposition, idea, or “suggestion”, which is generally their heart's desire abstracted and refined, is defended by them with arguments sought out after the event. They are all advocates who do not wish to be regarded as such, generally astute defenders, also, of their prejudices, which they dub ‘truths’ ”

Nietzsche, *Beyond Good and Evil*

Nietzsche was undoubtedly correct to point out that much of what we believe to be true is shaped by unconscious prejudices which it may be impossible to overcome to arrive at a truly unbiased world-view. But commentators have questioned the validity of his reasoning: does it really follow, from the fact that our beliefs are often the product of our unconscious prejudices, that there is no such thing as an objective standard of truth against which those beliefs should be assessed as correct or incorrect? The relativist might respond that it makes no sense to talk about an ‘objective’ standard of truth if our prejudices make it impossible for us to reach that truth. However, the defender of objective truth can reply that the idea of truths that it is impossible for us to reach – truths which go beyond the limit of our ways of knowing the world – is simply a consequence of the belief in a world that exists independently of human thought: as long as we believe that there is a world which does not depend on the activity of human minds for its existence, we should be willing to recognize that there will be “objective” truths about the way that world is – truths that hold regardless of whether anyone believes them, or whether anyone is even capable of coming to find out about them.

### **Culturally located knowledge**

It seems that the idea of a global relativism or subjectivism about truth faces serious problems. Perhaps a more plausible approach is to say, not that the standard for *truth* varies from culture to culture, but rather than the standards for *justification*, *rationality*, and *knowledge* are relative to the social norms and expectations of particular cultures.

“For the relativist there is no sense attached to the idea that some standards or beliefs are really rational as distinct from merely locally accepted as such. Because he thinks there are no context-free or super-cultural norms of rationality he does not see rationally and irrationally held beliefs as making up two distinct and qualitatively different classes of things”

Barnes & Bloor, “Relativism, Rationalism, and the Sociology of Knowledge”, pp.27-8

If Barnes and Bloor are right, and there is no universal or “objective” norm of rationality, then it should follow that there is no objective notion of justification, as long as we hold on to the idea that there is a connection between what we are justified in believing in the circumstances and what we rationally *ought* to believe in those circumstances. If having justified beliefs is a matter of forming our beliefs rationally, and there is no objective standard of rationality, then there is no objective

standard determining whether our beliefs are justified; they can only be justified or unjustified according to the **internal criteria** of whatever belief system is current in our culture.

At this point, it is useful to distinguish two questions. One is whether, as a matter of fact, different societies do have different standards of rationality “encoded” within their cultures. The other is whether, faced with a disagreement about standards of rationality, there is any way to decide which of the rival standards is “objectively correct”. If – as it happens – all human societies fundamentally agree about what it is to be rational, then there is no reason to think that the standard of rational belief is culturally located. This provides the material for a powerful response to the relativist, since it is plausible that there are standards of rationality which are shared between cultures. For example, people from different cultures can agree that it is usually rational to do what you believe will bring about the outcome you prefer, all things considered, and agree that it is usually irrational to perform an action which you believe will have only undesirable consequences and no good consequences. These kind of examples suggest that there is a shared, universal standard of rationality which does not depend on culturally located values as the relativist had suggested.

The second question was whether we could find a way of distinguishing between rival standards of rationality or justification to determine which is objectively correct. Here it seems that the defender of objectivity has a harder task – for any attempt I make to justify my standards of rationality over someone else’s standards will have to appeal to *my* standards, and it is precisely my standards that my opponent does not accept. However, there may be a possibility of reaching an “objective” defence of a community’s standards for rationality and justification by showing that some standards of rationality have more *practical* use than others. For example, a standard of rationality which said that it is rational to hold contradictory beliefs, rather than saying (as our standard does) that it is rational to *avoid* contradictory beliefs, would be a standard of rationality which made it impossible for anyone in the community to achieve very much: someone who actively sought out contradictory beliefs on every subject would never know what to do, as he would never know which of his contradictory beliefs to follow on a given occasion; if I simultaneously believe that it is raining and that it is not raining, how do I decide whether to take an umbrella with me? This kind of example suggests that it may be possible to defend standards of rational belief-formation such as “avoid contradictory beliefs” from a point of view that is not “located” within our own cultural system, by looking at whether those standards have positive or negative practical consequences.

Finally, when we encounter sociologists and post-modern theorists who put forward relativistic theories about “socially located knowledge”, we should be aware of the possibility that they are using the word “knowledge” to mean something completely different from the knowledge that is debated by epistemologists and which we have been discussing in this course. This is something that Barnes & Bloor take pains to stress in the article quoted above, saying that

“We refer to any collectively accepted system of belief as ‘knowledge’. Philosophers usually adopt a different terminological convention” (*op. cit.* p.22).

Moreover, Jean-Jacques **Lyotard**, sometimes described as the “father of postmodernism”, describes knowledge as a “kind of discourse”, arguing that “the nature of knowledge cannot survive unchanged” in the modern technological world, and suggests that knowledge has become an

“informational commodity” to be computerized, and bought and sold.<sup>2</sup> This conception of knowledge – as a kind of information to be bought and sold – is so radically at odds with the standard philosophical conception of knowledge as a certain kind of relation between a knowing subject and a proposition which is the content of this knowledge, that it seems that epistemologists and post-modernists are **talking past one another** – i.e. they think they are debating one thing, knowledge, but really their conceptions are so different that they are in fact talking about completely different things.

### **Science, relativism, and conceptual schemes**

It used to be taken for granted that science is a **cumulative** discipline, in which the insights of scientists from previous generations are gradually added to by new discoveries made by current researchers. But in 1962 Thomas **Kuhn** published *The Structure of Scientific Revolutions*, in which he argued that science proceeds through periods of stagnation and minimal progress separated by moments of radical change – which are known as **scientific revolutions** or **paradigm shifts** – in which old beliefs are overturned and new ways of doing science are established. However, Kuhn took the view that the grounds on which new paradigms overturned the old were not always purely rational: sometimes a shift in paradigm is dictated by the greater problem-solving ability of the new paradigm, but sometimes “personal and inarticulate aesthetic considerations” (p.158) are the only way of deciding between old and new paradigms. Moreover, although Kuhn thought it likely that new paradigms would have a greater problem-solving ability than previous paradigms, he rejected the idea that the progress of switching from one paradigm from the next will get us closer to an objective “truth”:

“ the list of problems solved by science and the precision of individual problem-solutions will grow and grow... In the sciences there need not be progress of another sort. We may, to be more precise, have to relinquish the notion, explicit or implicit, that changes of paradigm carry scientists and those who learn from them closer and closer to the truth.” (p.170)

Kuhn *denied* that his views were a kind of relativism, arguing that each new paradigm was objectively “better” than the preceding one. However, he has often been interpreted as a kind of relativist, on the grounds that the sense in which one paradigm is “better” than another is not the “objective” issue of whether the paradigm gets us closer to the truth, but rather depends on sociological factors such as the values current within the community of scientific enquirers and which problems those scientists find it important to solve.

Kuhn’s view have been developed by Hilary **Putnam**. He argues that a change in scientific paradigm results in a change in the *meaning* of the words used – claiming that although Newton defined “momentum” as meaning “mass times velocity”, Einstein’s Special Theory of Relativity shows that momentum does *not* equal exactly mass times velocity but merely approximates to it. Thus, according to Putnam, a change in scientific paradigm results in a change in **conceptual scheme** – a change in the concepts we use to describe the world. Putnam goes further: there is no description of the world that is *independent* of a conceptual scheme, and the ways we describe the world are true only relative to the system of concepts we adopt. Even words like “object” and “entity” have

---

<sup>2</sup> Lyotard, *The Postmodern Condition: a Report on Knowledge*, pp.3-5

different uses depending on which conceptual scheme we choose:

“The suggestion . . . is that what is (by commonsense standards) the same situation can be described in many different ways, depending on how we use the words. The situation does not itself legislate how words like “object,” “entity,” and “exist” must be used. What is wrong with the notion of objects existing “independently” of conceptual schemes is that there are no standards for the use of even the logical notions apart from conceptual choices.” Putnam, *Representation and Reality*, p.114

Further viewing: Putnam is interviewed about science and conceptual schemes at <http://tinyurl.com/PutnamPhilS>

However, although Putnam describes his view as a kind of **conceptual relativity**, it is not clear that it is a genuine kind of relativism. According to him, different conceptual schemes result in different, equally true descriptions of reality – but, once the meaning of the words we use in stating our theory is fixed as picking out one of these sets of concepts rather than another, the only thing that determines whether our statements are true is objectively existing reality; it is not the case, according to his view, that whether our concepts apply to reality depends on the social norms and cultural beliefs prevalent in our society, as the genuine relativist claims. Putnam may respond that the *choice* of a conceptual scheme is relative to a society, but that is beside the point: it would not make sense to argue that every truth stated in French is true only relative to French society, on the grounds that the choice of language is culturally determined; similarly it does not make sense that what is true within one conceptual scheme is true only relative to a society because it is the society that determines the choice of conceptual scheme. Once we have chosen a conceptual scheme, what determines whether those concepts apply is the nature of objectively existing reality; social customs and agreements within a culture play absolutely no role in determining how those concepts apply.

*Homework:* “Is objective knowledge possible?”

## Week 13: The implications of relativism

### Moral relativism

Up to now we have ignored one of the most widespread forms of relativism: this is moral relativism, the view that truths about morality – what we ought or ought not to do in a certain situation – are relative to the culture and society we live in. The claim that moral *truth* is relative to a society goes hand in hand with the denial that there is such a thing as **objective moral knowledge** – obviously, if the truth of your moral beliefs is relative to the society you inhabit, then so too is the correctness of your claims to moral *knowledge*. One of the reasons for the apparent popularity of moral relativism is that people who call themselves “moral relativists” misunderstand the nature and commitments of that view; in particular they do not realise the problem that this kind of relativism can leave us unable to *act* on the basis of our moral beliefs, leading to a kind of **inertia in human activity**. The best way to see this is to go through some of the common justifications for moral relativism.

First, people often espouse moral relativist views out of a desire to be **tolerant** of other cultures’ belief systems and practices. If I observe society X sanctioning an action which I consider morally repugnant, I may adopt a moral relativist stance, saying that what they are doing is morally fine within their culture and belief system, but would not be morally acceptable within my own. This enables me to claim to be “tolerant” of divergent moral opinions. However, that’s not the whole story. Consider the case of a person within society X who does something that is fine according to the standards current in *my* society, but morally repugnant within society X – say perhaps she wears clothes that would be fine in London but are judged as morally offensive according to the standards of society X. The moral relativist is committed to saying that the moral truth for you is determined by the society and culture you inhabit; and this person inhabits society X. So what the moral relativist has to say about this person is that she has done something that is definitely morally *wrong* (even though I might personally find nothing objectionable about it). What’s more, if the moral system current within society X sanctions severe punishment for this “transgression”, I should refrain from criticizing it – after all, the person they are punishing has done something that is genuinely “wrong” within their own moral system. So adopting moral relativism enables us to be tolerant of other societies’ moral systems at the cost of being *intolerant* of people within those societies who transgress against those systems.

Another reason which may be behind many people’s acceptance of moral relativism is a confusion between **descriptive** and **normative** uses of phrases like “morally acceptable”. To say that an action is “morally acceptable” in the *descriptive* sense is merely to say that most people within the culture *believe* it to be morally acceptable – it is a descriptive statement in the sense that it *describes* moral beliefs and attitudes within a culture. But that is not the same thing as saying that this action is “morally acceptable” in the *normative* sense which expresses a genuine moral judgement about the action in question. For example, I might say (descriptively), that crucifixion was “morally acceptable” in ancient Rome, meaning only that most people then thought it was morally okay. But that isn’t the same thing as saying that crucifixion was morally acceptable in the *normative* sense that means that all the people who thought that crucifixion was morally okay were *right* to have that belief – that there really *was* nothing wrong with crucifying people at that stage in history. Of

course, to be a moral relativist you have to believe that crucifixion was morally acceptable in the *normative* sense, not merely the descriptive – i.e. that what they did really was morally fine.

A similar kind of linguistic confusion arises when people endorse moral relativism because they do not believe in **moral absolutes**. The reason goes: I don't believe in moral absolutes, so I must be a moral relativist! But this is a mistake which results from a confusion about the meaning of "moral absolutes". A **moral absolutist** in this sense believes that the moral truth can be stated in simple principles which do not allow exceptions which take account of the circumstances of the particular case. For example, **Kant** argued that lying is always wrong, even when the result of telling the truth will be the brutal murder of a close friend. The crucial thing to remember here is that *denying that moral truths can be stated in simple exceptionless principles does not make you a moral relativist!* A moral relativist believes that the morally right thing to do depends on the views of the society you inhabit – but you can reject "moral absolutes" without accepting *that*. For example, you might be a **utilitarian** who claims that the right thing to do is fixed by the specific circumstances of the situation – namely which course of action will result in the most overall happiness. Utilitarians reject moral absolutes, but they are not moral relativists, because they claim that the right thing to do does not depend on the beliefs and values of your culture. Alternatively, you could be a **moral particularist**, who believes that the objectively "right" course of action is determined by the details of the specific circumstances, but rejects the idea that morality can be stated in terms of general rules or principles that guide action.

Moreover, there are reasons for thinking that moral relativism will result in moral stagnation and the lack of development of a society. The first is that moral relativists, in denying that any one system of morality can be objectively "better" than another, have great difficulty making sense of the idea that there can be such a thing as **moral progress**. Faced with the same society at different stages of its development, the moral relativist is obliged to say that there is nothing to make the society's later value system objectively any better than its earlier one. So they cannot say that a *change* in the moral values of a society is an improvement. For the moral relativist, for example, it genuinely was *wrong* for women to demand the right to vote at the start of the suffragette movement, because it was *wrong* according to the prevalent moral standards of the time. What's more, the development into a society in which it is morally *right* for women to have the vote cannot be described as an "improvement" by the relativist, because – according to him – there is no objective standard by which one set of moral standards can be counted as superior to another.

A second reason for thinking that moral relativism prevents progress is that it tends to prevent **intervention** to address apparent injustices within other societies. Suppose that you are trying to decide whether to volunteer to teach in a girls' school in a country where the prevalent morality condemns the education of women as immoral. If you are a moral relativist, you should not go: if that society says that the education of women is immoral, then – according to the moral relativist – the education of women *is* immoral within that society, and by going to help women become educated within that society you would be helping to achieve something that is genuinely morally wrong. So again it seems that moral relativism promotes **inertia in human activity**, preventing people from intervening in genuinely useful ways.

### **Is relativism scepticism in disguise?**

On the face of it, there is a clear distinction between relativism and scepticism: while the sceptic claims that knowledge is *impossible* in a certain area of human enquiry (e.g. knowledge of the external world), the relativist claims that knowledge is possible, but *socially located*, dependent for its correctness on the cultural practices and beliefs prevalent in the society as a whole. So the relativist says that knowledge is possible, and the sceptic says that knowledge is impossible. What clearer distinction could there be than that?

However, both relativism about knowledge and scepticism spring from a common cause, and can be seen as two rival answers to the same problem. The sceptic is motivated by the difficulty of moving from how things *appear* to us to knowledge of how things really *are* “out there in reality”; given the possibility that we might be in a sceptical scenario – for example a brain in a vat – it seems that there can be no objective justification for believing what we do about the outside world, and so the sceptic comes to believe that knowledge of the external world is impossible. In that sense, the relativist’s proposal can be seen as an answer to the same fundamental problem. The relativist may also be motivated by the difficulty of providing a justification of our beliefs about reality which shows that they are ultimately correct – but instead of declaring that knowledge is impossible, the relativist offers a different account of what knowledge is, to show that we can still have knowledge even if we do not succeed in objectively justifying our beliefs about external reality. According to the relativist, our beliefs can count as knowledge if they are licensed by the standards prevalent in the culture that we inhabit; never mind that these standards might fall short of the absolute certainty required by Descartes and others – what counts for knowledge is that we meet the standards for knowledge (and truth) set up by the beliefs of our community, and we are fortunate enough to inhabit a community where those standards are generally lax enough to permit us to “know” things even if we cannot justify ourselves in a way that would satisfy a committed sceptic.

*Extension reading:* Boghossian, *Fear of Knowledge*, ch. 1-3