

Chapter 2

Knowledge and Justification

We ended Chapter 1 at an impasse, with the rationalist and the empiricist accusing one another of being superstitious. The rationalist accuses the empiricist of superstition because the latter believes that events in the world will continue in predictable, law-like fashion, without any basis for this beyond what has happened in the past. The empiricist accuses the rationalist of superstition because the latter believes in necessary and universal laws underpinning the regularities we observe, laws that can't be seen or touched or measured. Perhaps we can make progress by asking how we would go about justifying our beliefs; we may be able to adjudicate between the two by seeing which of them can give the more compelling justification for their views.

A famous approach to knowledge, one that goes back at least as far as Plato, is that *knowledge is justified true belief*. We each have many beliefs; that it will rain this afternoon, that there will be pizza for tea, that film stunt-men don't really die, and so on. But when it comes to the crunch some of these beliefs may prove to be false. There is no denying that I hold these beliefs and that if they do indeed turn out to be false, it can hardly be said that I knew whatever-it-is that turned out to be false. A belief is just a belief, it has to be *true* to be knowledge.

This suggests that knowledge is belief plus *x, y, z*, and so on, where *x, y, z*, and so on, are conditions that have to be met for a mere belief to be magically transformed into knowledge. The question then becomes, what has to be added to belief to get to knowledge? What extra ingredient is needed?

First of all, we will shovel “belief” off to one side. What we are interested in is the *content* of a belief. We aren’t interested in the fact that your friend believes that Midnight Blue will win the 3.30 at Epsom, we are interested in what it is that he believes; that Midnight Blue will win the 3.30 at Epsom. The having of beliefs is of interest to psychologists, from a philosophical perspective it is the content, what comes after the “that”, that we are interested in.

Given a possible content of belief, there are a number of attitudes you can take up to it. You can hope that it comes about, wish for it, fear it won’t happen, and so on. But these all seem to be just psychological attitudes. What we want to get at is what is involved in knowing that ...that what? Well, that your belief is true, of course! And so it seems that knowledge is inescapably bound up with truth. You can’t coherently say “I know this, but it may turn out to be false”, whereas you can perfectly coherently say “I believe this, but it may turn out to be false”. Anyone who comes out with the first of these shows that they don’t understand what is involved in knowing something to be the case.

A different way of putting this is to think how unsettling it is to find out that something you felt you knew to be true is in fact false. If one of your beliefs turns out to be false then you may well feel upset or angry or betrayed, but if something you think you know, something that you rely on, turns out to be false, this is far more disturbing. This is bound up with the sense in which knowledge is *authorising*. If someone says, “I believe this is the way to the beach”, then while it would be annoying to walk for miles and find it isn’t at all, it wouldn’t be grounds to push them in the sea when you do finally get there. But if someone says “I know the bus will be here at four o’clock” and it has been and gone by 3.50, you have legitimate grounds for complaint. It is wrong to make a knowledge claim without justification, without good reason. Anyone who makes such a claim without good grounds leaves themselves justifiably open to criticism.

So far we have two ingredients for knowledge; something that we can believe (the content of a belief, what comes after *that*), and the

truth of what it is that we believe. But how can we find out whether the content of a belief – a content, for short – is true? “Ah, truth”, said Pontius Pilate, and washed his hands. Truth is a notoriously slippery philosophical concept, and a viable account of it will assure its author of a place in the philosophical pantheon. Fortunately we don’t have to come up with an account of what truth is, rather what we need is a way of telling whether a given content is true, or is false. Much as you don’t need to know how a microwave works in order to cook with it.

Your friend reckons that Midnight Blue will win the 3.30 at Epsom. He suggests that you should put your money on Midnight Blue as the odds are good; perhaps 5:1, in which case you stand to make £100 back off a £20 stake. Should you bet? First of all you ask if he has put his money where his mouth is. If the answer is no, then you have reason not to. This is because knowledge, more so than belief, involves commitment. We show what we believe by acting on it; actions do indeed speak louder than words.

Assuming your friend has put up his own money, do you trust his judgment? Does he form beliefs rashly, or does he consider the evidence and act accordingly? Before you put your money up you might want to ask him why he believes that this horse will win, what his reasons are for backing it. Because you would think that he has reasons, that he can offer a justification for his belief that it will win. After all he can hardly *know* that this horse will win, assuming the race isn’t fixed. But he must at least have good reason, and as he is advising you to put your money up it is reasonable in the circumstances for you to ask. If it’s going to cost you, you are entitled to a reasoned justification. It may be that this horse likes this course, has had a particularly good season, that there are no other strong contenders, and so on. What you would not expect is “I stuck a pin in the race card blindfold”, because that isn’t a reason at all.

On the other hand it could be that without knowing how it is done, your friend has an unerring knack for picking winners. He can reliably identify which horse will win even though when asked he can’t give a reasoned answer beyond “I just *know* it will win, I feel it in my bones”. Just as the clairvoyant claims contact with the beyond without knowing how it is done, so your friend just has a knack. Would you want to say that he has knowledge? In one sense of course

he has, but there is something fishy about this. What isn't on offer is any sort of justification beyond a gut feeling.

What we would expect, rather than gut feelings or pins in race cards, is a reasoned explanation. But however reasoned a response your friend can give of the likelihood of the statement "Midnight Blue will win the 3.30 at Epsom" coming true, you are still dealing with probabilities rather than certainties. There doesn't seem to be any way of attaining knowledge in such cases. But if there was a way of knowing rather than speculating about the future in such cases the world would be a very different place. We have, though, established that beliefs stand in need of an account, of a systematic set of reasons, before we should take them on board. We have made some progress towards filling out the concept of justification.

If knowledge is justified true belief,

- there has to be a content (what comes after *that*) that you believe,
- the content has to be true,
- you need reasons – justifications – to believe the content.

Foundationalism and Coherentism

Foundationalism and coherentism are two different approaches to justifying beliefs, that is, adding whatever-it-is that is needed to turn beliefs into knowledge. As a first stab, foundationalism begins with a class of basic beliefs which are regarded as so immediate or basic or primitive that they can't be sensibly challenged. Examples of such beliefs are "the sky is blue" or "you are now reading a book" (assuming you aren't having this read to you!). Such beliefs are often described as what is *given*, is presented immediately in perceiving the world in such a fashion that no thought or mental effort is needed. They record what is there, how the world around you immediately impinges on your senses. Given these, so the story goes, you can build a body of knowledge outwards and upwards, hence the "foundationalist" label.

Coherentism works rather differently. The coherentist argues that there is no sense in which we are straightforwardly passively acted on by the world around us, there is no "given". There is only what we have in our own heads, what we are aware of. The moment we think about our perceptions we are thinking about them and forming

beliefs about what they are of. You can no more have a “given” as the foundationalist claims, says the coherentist, as you can stand on your own shadow. We only have our beliefs, however these happen to be formed, and they are justified if they fit with – cohere with – one another in a systematic, interlocking fashion. Justification doesn’t come one belief at a time, as the foundationalist claims, rather it is a feature of a system of mutually supporting beliefs. If foundationalism is akin to putting up a building then coherentism is akin to a stone arch in which each stone is needed to support one another.

There are distinct differences between these two approaches, and you will need a firm grasp of these. With this sketch in place we will look more closely at the strengths and weaknesses of each position, and consider how they fit with empiricism and rationalism.

Foundationalism

The intuition that underlies foundationalism is that what is given in perception is indubitable. It’s just *there*, in front of you, beyond any reasonable doubt. If what is central to knowledge is certainty, what is beyond all rational doubt, then if we can locate what is certain and build on it we should be home and dry. What can be more obvious and immediate – that is, literally, where there are no intermediaries between us and what we seek knowledge of – than what is directly perceived?

This is the start of a paper called ‘Certainty’, delivered as a lecture by the Cambridge philosopher G.E. Moore (1873-1958, known for his forceful defences of common sense) in 1941:

I am at present, as you can all see, in a room and not in the open air; I am standing up, and not either sitting or lying down; I have clothes on, and am not absolutely naked; I am speaking in a fairly loud voice, and am not either singing or whispering or keeping quite silent; I have in my hand some sheets of paper with writing on them; there are a good many other people in the same room in which I am; and there are windows in that wall and a door in this one.

Moore, ‘Certainty’, p.171

Moore doesn’t say that he merely *believes* the facts he has stated. He claims that the sentences in which he has expressed them are *true*,

and that he is *certain* that they are so. He says that he *knows* that what he says is true. So we have examples of what is given immediately in perception, that seem to be beyond any reasonable doubt, and consequently can be justified as knowledge. From some such cases, can we build up a body of knowledge about our world, with the certainty that goes with our initial bits of knowledge transmitted to the rest?

The structure of the foundationalist approach is this:

- a class of basic beliefs that are individually immediately true,
- derived beliefs whose justification relies on one or more basic beliefs.

It matters that each basic belief is true independently of any other, because otherwise we end up with the holistic, systematic approach that is characteristic of coherentism (as we will see later in this chapter) rather than foundationalism. Note also the connection between basic beliefs and perception, because this links up with empiricism and has consequences for scepticism (Chapter 3) and our knowledge of the external world (Chapter 4).

The most obvious way to account for our knowledge of basic beliefs is, as we have seen, to think of them as directly and immediately given in experience. There is a very obvious sense in which we don't control what we see; I have little control over what is in front of me when I open my eyes and it is hard to ignore the road-mender outside with the jack-hammer. There is a sense in which we are *passive* perceivers of the world around us. Foundationalism relies heavily on this sense of passivity.

With these basic beliefs in place, so the story goes, we aim to build on them. Beliefs arising from direct perception don't take us very far. So we seek more complex beliefs derived from our basic beliefs, derived in such a way that the truth of our basic beliefs is passed on to our more complex beliefs; just as the foundations of a house ensure the stability of the structure built upon them. This is easy enough to state in the abstract but harder to pin down in detail. There is no obvious way in which the certainty that attaches to what is immediately "given" can be transmitted to what is further removed from perception, for example, " $\sqrt{2} = 1.414$ " or "the patient's immune system was damaged by the drug she was given during the trial".

Fortunately, perhaps, we already have a view on knowledge that is based on perception, observation and experience, because this is the basis of the empiricism. So we can shed some more light on this by looking at the fit between foundationalism and empiricism.

Foundationalism and Empiricism

The empiricist claims that all our knowledge is based on experience, of the world without us and of our selves within us. Experience tells us that we form beliefs and that we act on these beliefs, both in my own case and in the case of the people I observe around me. (Notice how “experience” has become a noun, a thing with a life of its own.) This being so, we must have some mental capacities to form and work with the basic beliefs we get out of perception, because this is what it is to be rational. It is by exercising these capacities that we construct our body of knowledge about the world.

The capacity to form and work with beliefs is based in mental capacities to recognise regularities, to realise that, for example, you’ve been here before or that you’ve seen this before. Experience enables us to map out where we are in time and in space, and to predict more or less what will happen next. There is nothing mysterious about these capacities in the sense that we can see ourselves and others operating in these sorts of ways; the evidence is all around us. This gives us the justification we need to claim that we have the relevant mental capacities.

This also explains how we can form *systems* of beliefs. We develop theories about ourselves and about the world around us, taking our experiences (and other people’s experiences, if you are prepared to accept their testimony and learn from their successes and failures) as datum points. The structures of these theories derive from our mental capacities, as evidenced in our behaviour. It may be a very long way from observing that chewing senna pods makes you rush to the loo or that chewing the bark of certain trees (containing quinine) relieves the symptoms of people who have fallen ill after being bitten by mosquitoes (victims of malaria) to “the patient’s immune system was damaged by the drug they were given during the trial”. Nevertheless, so the foundationalist claims, in principle we could trace back every step. There is nothing mysterious involved, it is all based on a long history of medical observations and experiments.

Equally in cases where we can't establish a connection we assume that there is one, only we haven't been able to find it. In the case of aircraft crashes we send in teams of investigators to establish what happened. In rare cases, when no cause can be found that fits with our existing body of theory, we assume that we haven't found the cause. We don't assume that there is some mysterious and essentially inexplicable cause – divine intervention, perhaps, or the work of a malicious spirit. We keep the file open and hope that an explanation will be forthcoming sometime in the future. In this sense foundationalism fits with the fallibilist attitude that permeates empiricism. It is only the basic beliefs that are given, that are indubitable. What we build on this is theoretical in the sense that it is tentative and more or less speculative, and may be revised in the future.

The picture is this; what is given in perception is indubitable. Think of these as a set of building blocks. Reflecting on (in the sense of “reflection” discussed in the section on ideas in Chapter 1, above) these ideas lead us to construct theories by ordering and arranging these basic beliefs. If these theories don't work we can go back to the drawing board and rearrange the basic beliefs, to generate new theories. This is a key aspect of foundationalism; there is what is given, *then* there is what we build out of it. In other cases we might discover new basic beliefs, in which case we have a new set of blocks to work with.

To the two points in the previous section we can add these:

- theories about the world based on systematising our basic and derived beliefs in accord with our mental capacities to establish relations between ideas given in sensation,
- a preparedness to revise these theories (fallibilism) by re-ordering our beliefs on the basis of new evidence (new basic beliefs or successes/failures of previous orderings of derived beliefs).

Objections to Foundationalism

The classic objection to foundationalism is the “infinite regress” argument. If a belief is justified by a further belief, and so on, then either we can follow up a chain of beliefs forever or it must end with something that isn't a belief. If it does go on forever then no belief is ever justified so foundationalism is bust as a theory of knowledge. If

on the other hand it does come to an end then whatever the terminus is, it can't be a belief. So we need to give a better account of what it is that is foundational.

Since foundationalism assumes passivity in perception the idea is that we begin what is "given", with what is immediately present in perception. Whatever is so given is certain and indubitable. But in order to bring whatever-it-is that we get immediately in perception to mind, in order to think about it at all, we must make some sense of it. In other words, we have to *conceptualise* it. To do this we have to think of whatever-it-is as a something-or-other; as animal, vegetable or mineral, and so on. In order to talk about whatever-it-is we have to classify it and find words that we can use to refer to it. At the very least I can choose some arbitrary name, that *this* (whatever-it-is, pointing to it) is a *chumba*. But I can't say anything at all about it to you unless I also induct you into the naming game. I point to it in your presence and say "*this* is a chumba" and now that you know what it is that I'm referring to, we can go on and have a conversation about it.

We might also go on to have an argument about it. I say it's blue, you say it's green. I say it's shiny and it smells nice, whereas you think it's smooth and sleek but smells a bit ripe. Which one of us is right? Which one of us has truly got hold of this chumba as it is in itself? If the two of us can't even agree on this, what possibility is there of establishing an objective standard of correctness?

Is all I can say is that this is how it looks or smells or feels to me? But if this is where we end up we are a long way from the idea we started with, that the world as I get it passively in perception is somehow given to me in a clear and indubitable fashion. The certainty we sought seems to have evaporated in a haze of "looks to me" talk and this can't bear the weight that foundationalism seeks to ascribe to it.

This, then, is the "infinite regress" argument:

- either a chain of beliefs goes on forever, in which case foundationalism lacks a basis for justification,
- or it terminates in beliefs of the form "this is how it appears to me", and these are too subjective and uncertain to bear the weight the theory requires them to bear.

It is a philosophical irony that what is supposed to be the paradigm of objective scientific practice – look for the evidence, for the facts

– ends up with subjective reports about how it seems to me. Of course we could appeal to the idea of an ideal observer, who is as carefully calibrated as the instruments he or she is supposed to work with. But how would we choose such a person, and how can we tell that they do indeed perceive the world exactly as it is? The idea is that objectivity resides in what we are studying, not in ourselves. In the words of the American philosopher Wilfred Sellars, the “given” that empiricists are so fond of is a “myth” that can’t withstand close scrutiny.

If we pursue a chain of beliefs from what we can think of as a higher-level statement (“the patient’s immune system was damaged by the drug she was given during the trial”, say, or “miscarriages of justice damage public faith in the police and in the legal system”) down to statements that are intended to catch what is immediately given in sensation (“it looks red to me”, perhaps, or “it smelt like rotting pig manure”) only to find that the content of these latter statements fails to connect us in a suitable fashion with the world around us (aren’t adequately “given”, to use the terminology), then both foundationalism and empiricism look to be in trouble. In the case of foundationalism we haven’t got the foundations we need, and in the case of empiricism the idea of direct, immediate evidence looks a bit shaky. Perhaps we should take a look at the rival theory, coherentism.

Coherentism

The essence of coherentism is that a belief can only be justified by another belief. We still have chains of beliefs but these go round in circles, the idea being that chains of beliefs hang together, thereby supporting one another. Justification arises from mutual support.

A city is a good example of a mutually self-supporting structure. There has to be an infrastructure of roads and houses and public buildings, and public utilities (gas, water, electricity, telecoms). The people who build and maintain this infrastructure need to be fed and clothed, so there are shops and manufacturers, and farmers, bakers, butchers, and so on. The whole somehow works together with, generally, not much by way of central planning and direction. It arises out of a spirit of co-operation and those who refuse to co-operate tend to be excluded or imprisoned. The city is almost an organic unity, akin perhaps to a horse or even a person. The seventeenth-century English political

philosopher Thomas Hobbes presents such a view of society in his *Leviathan*, with the king as the head and the rest of society fulfilling the other functions needed to sustain the whole. Hence the old-fashioned saying that everyone should know their place in society.

The picture that coherentism offers of our knowledge is that of a systematic set of beliefs. Each belief finds a place within the system, so it has to cohere with or be consistent with other beliefs. Of course it is still the case that we perceive things, but what we get in perception is not “given”. It is a joint product of perception and conceptualisation, with the inescapable fact that we interpret what we see. We don’t just see something, we see it *as* a whatever-it-is. Try looking up at whatever you can see in front of you, and try and drain it of all conceptual content. Can you really see nothing more than different patches of colour? Can you see a flat surface, as distance away from you is presumably a matter of interpretation, something you learn to appreciate? Yet this is the picture the empiricist works with. Which seems more plausible to you?

This raises an interesting question. If the coherentist is right, how can you ever see anything other than how you expect to see it? Clearly our theories about the world do change, but how is this to come about? The foundationalist can claim that you can relearn how to interpret the basic data, just as, whether you realise it or not, you learnt how to in the first place. For the coherentist, though, if someone makes an observation that conflicts with accepted theory, then either the observation has to be explained away or the theory has to be adapted to incorporate it. Before Australia was discovered classifications of birds listed swans as white. When black swans were discovered in Australia, the classification had to be changed to adapt to this. This sounds fine in theory but in practice the coherentist faces a continuing difficulty, that of integrating perceptual data into her approach. We will come back to this in the next section, on objections to coherentism.

The characteristics of coherentism are these:

- holism – our body of theory is taken as a whole, and alterations to one aspect of the theory may lead to adaptations elsewhere in the body of theory,
- consistency – beliefs must be consistent with one another.

It isn’t consistent to believe, for example, that all swans are white *and* that there are black swans in Australia. Something has to give.

How would the empiricist/foundationalist react to the discovery of black swans? In principle, one response would be to say, *this* bird (a black swan) and birds like it, we will call *swains*. They are similar to white swans but they aren't swans. (Bearing in mind the foundationalist belief in the independence of truths from one another.) By contrast the coherentist has already perceived it as a swan, albeit a non-standard one (the wrong colour) because within her system of beliefs it has many similarities to birds classified as swans. Hence to maintain her classifications she is under far more pressure to revise the classification of swans. Inventing a new category is a very unattractive option for the coherentist. In truth there are powerful reasons for both the empiricist/foundationalist and the coherentist to extend the classification of swans. The point of this example is to distinguish *being given a fact* (a large black feathered bird) and *seeing something as* a whatever-it-is (as a non-standard swan). This is where the difference lies.

One aspect of coherentism is that individuals, whether we think in terms of individual beliefs or individual parts of our city or the individual limbs and organs of an animal, aren't the same at all if they are separated from the whole. A leg of a horse, once separated from the rest of the animal, can no longer do what it did before. You can't have a baker unless she has flour, yeast, warm water, and a working oven. In terms of belief, beliefs are as they are because they have a place within the theory. A theory about whatever-it-is can be thought of as the story we tell ourselves about whatever-it-is. We have theories in physics and chemistry and the other sciences which tell us about the world around us, and about what we can and cannot do. A theory is a complex body of beliefs each of which is justified by the success (or failure, for that matter) of the theory as a whole. We don't test beliefs individually, we test theories as a whole. If the meteorologist predicts that it will snow next week, and the weather turns out fine, accounting for the failure of this prediction involves studying a mass of physical data. If all these accord with the prediction, it would be necessary to consider many aspects of a huge body of scientific theory about how weather systems respond to changing variables (land and sea temperatures, cloud cover, solar activity, and volcanic activity, as the smoke and hot particles emitted by volcanoes have an impact on the weather). The coherentist seems to have a better story to tell here than the foundationalist.

Different organisations look more or less hierarchical by nature and consequently more or less likely to fit the foundationalist model. An

army certainly looks to be about as hierarchical as it gets. But even here there is obviously a high degree of interdependence; a general can't fight a battle without subordinate officers and soldiers. As many companies have found to their cost not much gets done if you have too many chiefs and not enough indians. Here again, the coherentist looks to have a good story.

Given the parallels between foundationalism and empiricism, does something similar hold between coherentism and rationalism? We saw in the last chapter that rationalism in its extreme form involves spinning out a story about the world from a single fact – seeing the world in a grain of sand. This is clearly based on a strong notion of interdependence, so at first glance there certainly seems to be a link here. We can get a better hold on this by looking at one of the classic objections to coherentism.

Objections to Coherentism

The classic objection is the so-called “consistent fairy story” line. Our set of beliefs about the world may meet the criteria set out in the previous section (holistic and consistent) and yet be entirely false. Just as a novel imagining life in Europe after a Second World War won by the Nazis could be entirely consistent (entirely plausible) but entirely false.

This objection plays on the idea that coherentism operates only at the level of beliefs, that somehow it doesn't reach out and make contact with the world in the way that foundationalism's basic beliefs are supposed to. This is a deep problem for coherentism, that of doing justice to perception and the evident sense of passivity we feel when the world impinges on our senses. But instead of trying to “make contact” the coherentist can point to success, and say that since our present body of theory is an obviously successful basis on which to achieve what we set out to do it must, by and large, be correct. The “consistent fairy story” objection relies on considering only consistency as a property of a body of beliefs while ignoring the practical application of those beliefs.

What this latter move shows, though, is that there is a gap between coherentism and rationalism, because in defending coherentism we have looked for support to the world around us and not at what we

have in our own minds. Because it is sticking to what we have in our own minds – spinning out a story about the world around us by reason alone – that gives the “consistent fairy story” objection its bite. And on top of this we’ve actually brought coherentism into line with empiricism, by bringing in our experience of the world around us. To be honest, though, what we’ve done is skirt around the philosophical elephant in the room; how it comes about that our thought, that mysterious activity that seems to go on between the ears and behind the eyes, is of the solid world we encounter, more or less painfully, every day. Accounting for this just is the problem.

We have now moved away from rationalism in the direction of the fallibilism that characterises empiricism, because we’re allowing our experience of the world to lead to revisions in our body of theory about the world. Reading between the lines of Chapter 1 you can see that consistency and holism fit with the rationalist approach. Intelligible thought about the contents of my own mind has to be characterised by consistency, or my thoughts would be unintelligible; I can’t rationally have the thoughts that I am happy and that I am unhappy simultaneously. Rationalist adherence to the principle of sufficient reason is based on the view that the world is systematic and intelligible, and this forces holism.

As we saw in Chapter 1 the consistency and holism that rationalism goes in for isn’t supposed to arise from experience. Rather it is supposed to come out of reflecting on what we have in our minds, as a result of the exercise of reason. So while there are some parallels between rationalism and coherentism, in the end it seems that coherentism isn’t inconsistent with empiricism. And because coherentism fits better with a view of ourselves as beings that try to make sense of the world around us by telling stories – by devising, testing and adapting the theories that embody these stories – it looks like we should adopt as our theory of knowledge some combination of empiricism and coherentism. As long as we can reconcile the fact that empiricism begins with what we get in experience, and coherentism with the idea that a belief can only be justified by another belief because nothing is given straightforwardly in experience. This we will come back to in Chapter 4, below.