10

LOOKING INWARDS

I said at the end of the last chapter that the ‘looking outwards’, transparency account of self-knowledge arose as a response to the idea that we acquire self-knowledge by looking inwards. For example, Evans prefaces his defence of his version of the transparency method by quoting something obscure that Wittgenstein is once reported to have said in a discussion in Oxford and observing that Wittgenstein was ‘trying to undermine the temptation to adopt a Cartesian position, by forcing us to look more closely at the nature of our knowledge of our own mental properties and, in particular, by forcing us to abandon the idea that it always involves an inward glance at the states and doings of something to which only the person himself has access’ (1982: 225). Evans doesn’t say very much about why we should abandon this idea, beyond pointing out that a person’s internal state ‘cannot in any sense become an object to him. (He is in it)’ (1982: 227). So my question in this chapter is this: what exactly is wrong with the ‘inward glance’ model of self-knowledge?

Although this model has had a fairly bad philosophical press in recent years, even its staunchest critics accept that there is something intuitive about the idea that we acquire self-knowledge by some form of inner perception, by looking inwards. It’s natural to suppose that a basic source of self-knowledge for humans is introspection, and the point of talking about knowing one’s own states of mind by means of an ‘inward glance’ is to suggest that introspection is a form of inner observation: we know our own states of mind by perceiving or observing them. Obviously introspection isn’t perception in the ordinary sense, and involves the exercise of ‘inner’ rather than ‘outer’ sense. Still, even if introspection isn’t literally perception you might think that perception is, as Armstrong puts it, ‘the model for introspection’ (1993: 95).

Saying that there is something intuitive about the idea that we acquire self-knowledge by some form of inner perception might not cut a whole lot of ice philosophically speaking. There are lots of examples of philosophy overturning our intuitions, and who is to say that the supposedly intuitive perceptual model of introspection won’t turn out to be another victim of philosophical reflection? What we need is arguments in support of the perceptual model, not intuitions. Okay, so here’s an argument: suppose you start out with the idea that knowledge of our own beliefs and other attitudes is normally direct or immediate. In this context, direct knowledge is non-inferential knowledge. As Boghossian puts it in a passage I first quoted in chapter 4, in the case of others I have no choice but to infer what they think from what they do or say but in my own case ‘inference is neither required nor relevant’ (1998: 150-1). In my own case inference is neither required nor relevant because ‘normally I know what I think – what I believe, desire, hope or expect- without appeal to supplementary evidence’ (1998: 151).

Suppose we label the premise that self-knowledge is normally non-inferential the immediacy premise. Then the next question is: how can the immediacy premise be correct? One possibility is that inference can be neither required nor relevant when it comes to self-knowledge because knowledge of one’s own thoughts is normally based on nothing. But this is hard to swallow. Maybe cognitively insubstantial self-knowledge – say the knowledge that I am here- can be based on nothing, but self-knowledge isn’t cognitively insubstantial and so can’t be based on nothing. That leaves only one option: self-knowledge can be immediate and yet not based on nothing if and only if it is a form of observational knowledge. Assuming that perceptual knowledge is immediate, then modelling introspection on perception is really the only effective way of securing the immediacy of introspective self-knowledge. I’ll refer to this as the Immediacy Argument for the ‘inner perception’ model of introspection.

Some proponents of TM won’t agree that thinking of self-knowledge as perceptual or as based on nothing are the only two ways of securing its immediacy. Rationalists like Moran think that TM is a source of immediate self-knowledge but I have argued that they are wrong about that; self-knowledge acquired by using TM is notably indirect, so if you want self-knowledge to come out as immediate then TM doesn’t look like a good way of getting what you want. Given the immediacy premise, it would appear that all roads lead to the perceptual model, and that’s the point of the Immediacy Argument.

I don’t much care for the Immediacy Argument, though it does at least have the merit of being an argument. There are basically two things wrong with it. Firstly, the immediacy premise is no good, despite its popularity with philosophers of self-knowledge. Secondly, it’s far from obvious that perceptual knowledge is non-inferential and, in this sense, immediate. If perceptual knowledge is based on inference then modelling introspection on perception won’t provide much of an explanation of the immediacy of self-knowledge. It’s good that the Immediacy Argument is no good because it so happens that there are also other, independent objections to the perceptual model. As we will see, this model is more robust than it is usually given credit for being, but is ultimately unacceptable.

If you accept the immediacy premise but reject both the perceptual model and the view that self-knowledge is based on nothing then, as Boghossian points out, you really are in trouble. When it comes to explaining knowledge of our own thoughts the only three options are that this knowledge is:

1. Based on inference or reasoning.
2. Based on inner observation.
3. Based on nothing.

If, like Boghossian, you end up rejecting all three options then all you’re left with is sceptical thesis that we can’t know our own minds. The sensible response to Boghossian’s trilemma is to question the immediacy premise and argue instead in favour of inferentialism, the view that knowledge of our own attitudes can be, and indeed is, a form of inferential knowledge. I will defend inferentialism in the next chapter. What I want to do in the rest of this chapter is to discuss some arguments against the inner observation model. Specifically, I want to take a look at what I’m going to call the neo-Humean argument against this model. This argument, versions of which you can find in Boghossian’s paper ‘Content and Self-Knowledge’ and in Sydney Shoemaker’s Royce Lectures, leads naturally to the inferentialism I ultimately want to defend.

This is my plan: I’ll start by setting out the orthodox way of representing introspection as perceptual, which Shoemaker labels the ‘object perception model of introspection’. Then I will look at various arguments against the object perception model, including the neo-Human argument. Although I’m sympathetic to some aspects of this argument, there is one element of it which is fishy but nevertheless common ground between some proponents and some opponents of the object perception model. This is the assumption that genuinely perceptual knowledge is non-inferential. As we have seen, proponents of the object perception model rely on this assumption in defence of the idea that the perceptual model of introspection has introspective self-knowledge come out as non-inferential. And as we will see, opponents rely on it when they argue that introspection can’t be perceptual because it is inferential. They are both wrong because ‘perceptual’ and ‘inferential’ aren’t contradictories. However, in the end I don’t think that we should accept the object perception model. Nor, as I will argue at the end of this chapter, should we accept the alternative perceptual model which Armstrong defends in A Materialist Theory of the Mind. The right view is that knowledge of our own beliefs and other standing attitudes is inferential and is not perceptual. In this chapter I will defend the second conjunct of this conjunction. I will defend the first conjunct in the next chapter.

The basic idea of the object perception model is that being introspectively aware of one of one’s own states of mind (beliefs, sensations, or experiences) is like being perceptually aware of a non-mental object: just as sense perception acquaints us with non-mental objects, so introspection acquaints with mental objects. Introspection, on this view, is a kind of internal searchlight that illuminates, and provides us with, knowledge of our own states of mind in the way that ordinary sense perception illuminates, and provides us with, knowledge of bits of non-mental reality. Of course there are differences between introspection and ordinary sense perception of objects but the claim is that they are sufficiently alike for it to make sense to think of introspection as a kind of perception.

How alike is ‘sufficiently alike’? It’s hard to answer this question in the abstract but consider these apparent differences between introspection and ordinary perception: when you perceive a non-mental object such as a tree or a table you use your sense organs - your eyes, for example- but there is no organ of introspection. Sense perception involves the occurrence of impressions that are distinct from the object of perception but ‘no one thinks that one is aware of beliefs and thoughts by having sensations or quasi-experiences of them’ (Shoemaker 1996: 207). The objects of ordinary sense perception exist whether or not they are perceived but introspectable mental states do not exist whether or not we are introspectively aware of them. These are all respects in which introspection and ordinary perception are not alike, so how can the object perception model possibly be correct?

The object perception model has at its disposal two ways of dealing with such points. One is to argue that the alleged differences between introspection and perception aren’t genuine. The other is to argue that though the differences are genuine they aren’t significant enough to undermine the model. For example, it’s true that there is no organ of introspection but there can also be perception without an organ of perception; bodily perception (proprioception) is what tells you whether you are sitting or standing but there is no organ of bodily perception. It’s also false that the proper objects of introspective awareness aren’t independent of our introspective awareness of them: you can have a belief or desire or hope you aren’t introspectively aware of. Since our beliefs and other standing attitudes are not self-intimating, this leaves it open that our introspective awareness of them is quasi-perceptual. It’s true that we aren’t introspectively aware of our beliefs by having sensations of them, but this difference between introspection and perception isn’t important enough to undermine the object perception model.

No doubt there is more to be said about these issues but the discussion so far suggests that many of the standard objections to the object perception model aren’t decisive. There is, however, another objection to this model which looks much more threatening. What I have in mind is an objection implied by the neo-Humean argument, and it’s now time to say what this argument is. Boghossian’s version of the argument turns on what he calls an ‘apparently inevitable thesis about content’, the thesis that ‘the content of a thought is determined by its relational properties’ (1998: 149). Given this apparently inevitable thesis, ‘we could not know what we think merely by looking inwards’ because ‘what we would need to see, if we are to know by mere looking, is not there to be seen’ (ibid.). Hume gets into the picture on the assumption that the relational properties of a thought which determines its content are causal properties. This assumption causes a problem for the object perception model because it’s not possible to know a thought’s causal role directly: ‘the point derives from Hume’s observation that it is not possible to ascertain an item’s causal properties non-inferentially… discovering them requires observation of the item’s behaviour over time’ (1998: 162).

An example might help: suppose that what makes a thought a thought about vodka as opposed to a thought about gin is the thought’s causal properties: thoughts with causal role R1 are thoughts about vodka whereas thoughts with causal role R2 are about gin. So to know that my present thought is about vodka thought I would need to know that it has causal role R1 as opposed to causal role R2 but I couldn’t possibly know non-inferentially, by mere inspection, a thought’s causal role. A thought’s causal role is what is ‘not there to be seen’, and that’s why we could not know what we think by looking inwards. Shoemaker is making essentially the same point when he argues against the object perception model on the grounds that our mental concepts are concepts of states that occupy certain causal or functional roles: ‘we have been “taught by Hume” (I do not say that we have been correctly taught this) that any report that carries causal implications goes beyond what we can, strictly speaking, perceive to be the case’ (1996: 221).

Is the neo-Humean argument any good? The two key questions here are:

1. Is the content of a thought is determined by its relational properties?
2. Is it possible to ascertain an item’s relational properties perceptually?

The neo-Humean argument goes through only if the answer to the first question is ‘yes’ and the answer to the second question is ‘no’. In fact, the answer to both questions is ‘yes. Even if you can’t ascertain an item’s relational properties non-inferentially it doesn’t follow that you can’t ascertain an item’s relational properties perceptually. The point is not that the object perception model is okay after all but rather that the neo-Human argument isn’t decisive as it stands. What actually undermines this model is a variation on the neo-Humean argument which doesn’t require acceptance of “what Hume taught us”.

Let’s start with (a). The functionalist thesis that the content of a thought or belief is determined by its causal role is one way, but not the only way, of expressing the idea that its content is fixed by its relational properties. There are ‘externalist’ or ‘anti-individualistic’ accounts of content on which the answer to (a) is also ‘yes’. For example, the answer to (a) is ‘yes’ if a person’s thought contents are determined by the relations he bears to the physical or social environment. I haven’t committed myself to either externalism or functionalism but I have endorsed a dispositionalist view of belief and other attitudes. For example, in chapter 2 I described beliefs as ‘habits of thought’ and I said in the last chapter that whether you believe that P depends on whether you are disposed to think that P when the question arises, act as if P is true, and use P as a premise in reasoning. The dispositions that go with believing that P are multi-track and include what Schwitzgebel calls ‘phenomenal’ dispositions – dispositions to have certain sorts of conscious experiences - as well as cognitive and behavioural dispositions. Now the question is: what is the impact of this view of belief on (a)?

The answer to this question is ‘very little’. For a start, dispositional properties are still broadly relational, in which case thinking that the content of a belief or thought is determined by its dispositional properties about belief is not at odds with thinking that it is determined by its relational properties. The sense in which dispositional properties are relational is that they can’t be described without describing other things (Langton 2006). In the case of beliefs, the other things include certain sorts of behaviour and conscious experiences. Functionalism and the dispositionalism I’ve been presupposing are different forms of Boghossian’s ‘apparently inevitable thesis about content’, and it might seem that the neo-Humean argument works fine either way: surely it isn’t possible to detect an item’s dispositional properties perceptually any more than it is possible to detect an item’s causal properties perceptually (cf. Goldman 2006: 248). So the answer to (b) is still ‘no’

When a philosopher says ‘surely P’ it’s generally a good idea to watch out, and this is just such a case. It’s actually far from obvious that functional or dispositional properties can’t be ascertained perceptually. For example, you might think that colour and shape are causal or dispositional properties of objects but this doesn’t imply that you can’t know by seeing that an object is square or red. Reports of shape and colour certainly carry causal implications but such reports don’t go beyond what we can, strictly speaking, perceive to be the case. This is one reason for thinking that “what Hume taught us” isn’t right; just because the content of a thought is determined by its relational properties it doesn’t follow that we couldn’t know what we think merely by looking inwards, or that what we would need to see, if we are to know by mere looking, is not there to be seen.

Here’s an even better example that also suggests that the correct answer to (b) is in fact ‘yes’:

Being a dime is not an intrinsic property of an object. For something to be a dime it must bear a number of complicated relations to its economic and social environment. And yet we seem often able to tell that something is a dime purely observationally, by mere inspection of its intrinsic properties (Boghossian 1998: 162-3).

Here is now Boghossian thinks that the neo-Humean argument should deal with this potential counterexample:

The reason an extrinsic property seems, in this case, ascertainable by mere inspection, is due to the fact that possession of that property is correlated with possession of an intrinsic property that is ascertainable by mere inspection. The reason the coin’s dimehood seems detectable by mere inspection derives from the fact that it having the value in question is neatly encoded in several of its purely intrinsic properties: in the phrase “ten cents” that is inscribed on it, and in several properties of its size, shape, and design characteristics (1998: 163).

What this means, Boghossian concludes, is that ‘the process by which we know the coin’s value is not really inspection, it’s inference: you have to deduce that the coin is worth ten cents from your knowledge of its intrinsic properties plus your knowledge of how those intrinsic properties are correlated with possession of monetary value’ (1998: 163-4).

Let’s just grant that the process by which we know the coin’s value is inference. In talking about ‘inference’ in this context we obviously aren’t talking about something we are conscious of doing. The inferences in question require no attention; they are automatic and effortless. How it is supposed to follow that the process by which we know the coin’s value is ‘not really inspection’? It follows only on the assumption that genuine perception is non-inferential but why should we accept that assumption? It’s no good arguing that perception can’t be inferential because we aren’t conscious of inferring when we perceive; that would also be a reason for questioning the idea that we know a coin’s value by inference. The issue is whether, apart from whether it feels like perception involves inference (it doesn’t), there are nevertheless good explanatory reasons for supposing that genuinely perceptual knowledge is, or can be, inferential. The natural view here is that while there might be a sense in which knowledge of a dime’s monetary value is inferential, there is also a perfectly straightforward and intuitive sense in which you can see that it’s a dime.

There is a whole lot to be said about the role of inference in perception, certainly far more than I can possibly say here. To save time, I’m just going to state dogmatically that I’m pretty much in agreement with the view that perception involves inference. As Harman argues in his book Thought, one indication that this is so is that the inferential approach provides the best explanation of perceptual ‘Gettier cases’ such as the following: you come to believe by looking that there is a candle directly in front of you. There is a candle there but it’s obscured by a mirror which is showing you the reflection of a candle off to one side. Your belief that there is a candle directly in front of you is true and justified but not knowledge. Why not? Because you infer that it looks to you as if there is a candle there because there is a candle there, but this explanation is false; it only looks to you as if there is a candle there because you can see the reflection in the mirror of a candle that is not directly in front of you.

In addition to explaining perceptual Gettier cases, the inferential approach also seems to accord with the best scientific understanding of perception. For example, you determine how far away a perceived object is by means of visual cues from which you infer the location of the object: ‘given those cues the perceiver infers that objects are in those places’ (Harman 1973: 175). It’s true that we sometimes say things like “I don’t infer that P, I can see that P” but this can be read as shorthand for “I don’t just infer that P, I can see that P”. As it happens, the justification for reading the first statement this way is itself inferential: given that the best total explanatory account of perception implies that it involves inference, the basis for saying that perception involves inference is inference to the best explanation.

Suppose this is all correct. This spells trouble for the neo-Humean argument because it deprives it of one of its key premises. The argument goes: the content of a thought is fixed by its relational properties; we can only know our own thoughts perceptually if we can know them non-inferentially; our knowledge of our own thoughts can’t be non-inferential if the content of a thought is determined by its relational properties; therefore our knowledge of our own thoughts can’t be perceptual or modelled on perception. It’s the second premise of this argument which now seems dubious. For all that the neo-Humean argument proves, it would be open to someone who wants to hang on to the object perception model to insist that we can know our own thoughts by ‘looking’ even if it’s not possible to ascertain a thought’s content-determining relational properties non-inferentially.

This might strike someone like Boghossian as a pretty hollow victory for the object perception model. After all, this model was first introduced as an alternative to an inferential conception of self-knowledge, as a way of explaining how inference can be neither required nor relevant when it comes to knowledge of one’s own thoughts. If it turns out that the object perception model is itself committed to regarding self-knowledge as inferential then it is in no position to accommodate the immediacy premise, but without this premise the Immediacy Argument for the ‘inner perception’ model of introspection also lapses. This is what I meant when I said that both proponents and opponents of the object perception model rely on the same dubious assumption that ‘perceptual’ and ‘inferential’ are contradictories; neither the Immediacy Argument nor the neo-Humean Argument can do without the assumption.

Given that this is where we have ended up the obvious next step would be to question the immediacy premise. What causes all the trouble is the assumption that inference is neither required nor relevant for knowledge of one’s own thoughts. I’m going to argue in the next chapter that the immediacy premise has much less going for it that it generally assumed, and is only as popular as it is because critics of inferentialism are usually attacking a straw man. Going back to Boghossian’s trilemma, this means that I’m going for a version of option 1: knowledge of our own thoughts is based on inference or reasoning. What I’ve been arguing in this chapter is that this doesn’t rule out option 2: knowledge of our own thoughts is based on inner observation. Be that as it may, I think that we should still reject option 2, if not for the exact reasons given by the neo-Humean argument then for closely related reasons.

The point is this: suppose that you know that you believe your socks are stripy, and we’re trying to model the introspective awareness on which this knowledge is based on object perception. When you perceive that your socks are stripy you do so by perceiving your socks but you aren’t introspectively aware that you believe your socks are stripy by being aware of the belief that your socks are stripy. In perception you are normally aware of a multiplicity of objects, and there is such a thing as singling out an object and distinguishing it from others by its perceived properties and spatial location. In contrast, even if you are introspectively aware that you believe that your socks are stripy, this isn’t a matter of singling this belief out and distinguishing it from other beliefs you are also introspectively aware of. Propositional attitudes aren’t ‘objects’ waiting to be ‘singled out’ on the basis of introspectively available information about their relational and non-relational properties.

This points to another difference between introspection and perception: as Shoemaker notes, ‘perception of objects standardly involves perception of their intrinsic, nonrelational properties’ (1996: 205). When it comes to beliefs and other attitudes, it isn’t clear what their ‘intrinsic, nonrelational properties’ are, let alone what it would be for introspection to involve awareness of such properties. Suppose you are some kind of physicalist and think that the ‘intrinsic properties’ of mental states are physico-chemical. Then you can hardly fail to notice that you aren’t introspectively aware of such properties. Perception yields detailed awareness of the intrinsic properties of objects whereas introspection provides us with little information about what physicalism regards the intrinsic nature of mental events.

These objections to the object perception model are in the same general ballpark as the neo-Humean argument, though the emphasis is somewhat different. That argument starts with the idea that the content of a thought is determined by its relational properties and then objects to the object perception model on the basis that it’s not possible to ascertain an item’s causal properties non-inferentially and hence perceptually. But even if perceptual knowledge can be inferential this is of no help to the ‘inner perception’ model of introspection unless in introspective awareness of a belief there is an “item” whose intrinsic properties you are aware of and from which you can infer that it has certain relational properties. The problem is that there is no such “item” or “object” of awareness, and this is the point at which the analogy with Boghossian’s dime breaks down completely. When you see that the coin in your hand is a dime you are implicitly inferring its monetary value from its intrinsic properties because its monetary value is correlated with its intrinsic properties. In the case of a thought or belief, you aren’t aware of the intrinsic properties of a mental “item” and even if you were that wouldn’t get you anywhere because, “facts about a thought token’s content are not correlated with any of that token’s purely intrinsic properties” (Boghossian 1998: 163).

Let’s see where we’ve got to: I’ve argued that as far as knowledge of our own beliefs and thoughts is concerned the object perception model is no good, even if it has turned out to be a tougher nut to crack than one might be expected. However, before giving up on the idea that introspection is a form of inner perception it’s worth point out that the object perception model is only one version of the perceptual model and there are others which avoid quite a few of the difficulties the object perception gets into. As I’ve already indicated, what I have in mind is Armstrong’s ‘broad’ perceptual model of introspection. This is by far the best bet for someone who is really serious about the immediacy premise so I’d like to end this chapter by briefly explaining why, in the end, the broad perceptual model isn’t the way to go.

Armstrong’s basic idea is really simple, and depends on his background conception of knowledge and perception more generally. Here’s the basic story: perception is the acquiring of beliefs or information about the current state of the organism’s body and environment. A belief is a dispositional state and, like other mental concepts, the concept of a belief is the concept of a mental state is the ‘concept of a state of a person apt for bringing about a certain sort of behaviour’ (1993: 82). Perceptions are events rather than dispositional states; they are acquirings of beliefs, and the major source of our non-inferential knowledge. Non-inferential knowledge is in turn a matter of reliability. You know that P non-inferentially just if, in the absence of reasons for believing that P, you do believe that P and your belief is empirically sufficient for the truth of P (1993: 189). For example, if by using your eyes you acquire the belief that you are wearing stripy socks then your belief is knowledge just if it is an empirical truth that your wearing socks is a necessary condition of your acquiring that belief. If your belief is reliable in this sense then you have acquired knowledge (1994: 237).

The idea that mental states are dispositional might seem difficult to reconcile with the claim that our introspective awareness of them is quasi-perceptual. If dispositional properties are relational then introspective awareness of a mental state would be direct awareness of an ‘abstract, relational state of affairs’; in effect, it would be ‘direct awareness of counter-factual truths’ (1993: 96), and the worry is that this isn’t possible. However, on an information-flow view of perception there is no difficulty. Just as perception is the acquiring of information or misinformation about our environment, so introspection is ‘the getting of information or misinformation about the current state of our mind’ (1993: 326). Introspectively acquired belief about your own states of mind are self-knowledge as long as they are reliable. It’s the picture of introspection as an internal searchlight that lights up the mind that causes all the trouble for the perceptual model. If introspection is simply the getting of beliefs, there is no reason to deny that these beliefs can be about one’s current states of mind, dispositional or otherwise.

Among the apparent advantages of Armstrong’s picture, one is that it seems to explain how introspective awareness of one’s mental states can be perceptual in a way that bypasses many of the objections to the object perception model. The issue of whether inferential knowledge can be perceptual doesn’t arise because for Armstrong one’s knowledge that one believes that P can be non-inferential even if the content of one’s belief is determined by its dispositional properties. On the issue of whether introspection tells us anything about the intrinsic nature of the mental, Armstrong is happy to accept that it does not. He thinks that mental states are in fact ‘physico-chemical states of the brain’ (1993: 90) but that there are perfectly good biological reasons why ‘introspection should give us such meagre information about the intrinsic nature of mental events’ (1993: 99). The main reason is that having such information is of little value in the conduct of life. So we are not in the position of having to infer a belief’s dispositional properties from our introspective awareness of its intrinsic properties; there is no such awareness and no such inference. We are directly aware of the dispositional properties of our beliefs and that is why Armstrong’s model appears to be in a much better position than the object perception model to regard self-knowledge as genuinely ‘immediate’.

Still, there are good reasons not to settle for Armstrong’s view. The first point is ad hominem: comparing introspection with perception only makes it plausible that introspective knowledge is non-inferential if perceptual knowledge is non-inferential, but Armstrong himself ends up arguing that a lot of ordinary perceptual knowledge is inferential. Even when it comes to something as simple as seeing a cat’s head or a sheet of paper ‘there is a concealed element of inference’ from ‘a certain group of properties of objects that may be called visual properties’, including colour, shape and size (1993: 234). What we see ‘immediately’ is that there is a thing with certain visual properties before us. This, ‘by an automatic and instantaneous inference, produces the further belief that there is a cat’s head or a sheet of paper before us’ (1993: 235). However, if only a special subclass of perceptual knowledge is non-inferential, then what exactly is the argument for modelling introspective knowledge on non-inferential rather than on inferential perceptual knowledge? This is just the view I’ve been exploring, and Armstrong doesn’t show that there isn’t a concealed element of inference in introspection, just as there is perception. The fact that we aren’t aware of any inference in introspection is of little significance because Armstrong is perfectly happy with the idea that inference can be unconscious (see 1993: 194-8). This isn’t an argument against thinking of introspection as perceptual but it is an argument against thinking of introspective knowledge as immediate.

A more substantive worry about Armstrong’s view is that it makes introspection out to be fundamentally no different, epistemologically or phenemologically, from clairvoyance. What I mean by clairvoyance is the kind of thing Laurence BonJour has talked about over the years. For example, there is the case of Norman who, for no apparent reason, finds himself with beliefs about the President’s whereabouts; Norman has neither seen nor heard that the President is in New York but believes that the President is in New York, and his belief is reliable. Even if it’s right to describe Norman as ‘knowing’ that the President is in New York his knowledge is very different from ordinary perceptual knowledge. When you know that the President is in New York by seeing him in New York you are aware of the President, and your knowing requires a degree of cognitive effort, even if it’s only the minimal effort of looking and paying attention to where the President is. In contrast, Norman is not aware of the President and his belief is not the result of any cognitive effort on his part. The belief that the President is in New York simply comes to him; he has no justification for believing that the President is in New York and no idea why he believes this or how he could possibly know where the President is. That’s why some critics have concluded that what Norman has is not really knowledge, properly so-called, or that it is only knowledge in a secondary or derivative sense.

In what sense is introspection, as Armstrong thinks of it, like clairvoyance? Suppose that, in addition to believing that the President is in New York, Norman also has the second-order belief that he believes that the President is in New York. If introspection is simply the ‘getting of beliefs’ by means of a reliable mechanism then the way that Norman ‘gets’ his second-order belief is similar to the way he ‘gets’ his first-order belief about the President’s whereabouts. No effort is required and he just finds himself believing things about his beliefs. Could this be how we know our own beliefs? Aaron Zimmerman thinks not:

We are not in the position of the infallible psychic who just finds herself believing things about the future for no good reason; we simply do not find ourselves believing that we believe some things and not others (2006: 349).

Zimmerman’s point here is primarily phenomenological rather than epistemological but there is also an underlying epistemological worry about Armstrong’s model of self-knowledge. Just as Norman has no justification for believing the President is in New York, and no idea why he believes this, so he has no justification for believing that he believes the President is in New York and no idea why he believes that he believes the President is in New York. If there is anything to this comparison then the upshot is that what Norman gets from introspection is not self-knowledge properly so-called or only self-knowledge in a secondary or derivative sense.

If you are drawn to the broad perceptual model of introspection could you just bite the bullet in response to such objections and argue that self-knowledge is in the relevant respects like clairvoyance? After all, if you ask Norman whether he believes that the President is in New York he has no trouble answering your question but he would be hard pushed to say how he knows what he believes. The belief that he believes the President is in New York just comes to him when the question arises but he has no reason for believing that the President is in New York. Doesn’t this mean that Norman is in the position of finding himself believing that he believes some things and not others? He doesn’t know how he knows that he believes the President is in New York, any more than he knows how he knows that the President is in New York. Yet, at least on an externalist understanding of knowledge, he still knows that the President is in New York and knows that he believes that the President is in New York. His knowledge in both cases in non-inferential, and externalism denies that in order to know non-inferentially that P you have to have a reason for believing that P.

The problem with this line of thinking, apart from its reliance on a pretty full-on form of externalism, is that it ends up undermining the idea that introspection should be modelled on perception. As I’ve argued, perceptual knowledge is strikingly different from clairvoyant “knowledge” so it doesn’t exactly help the perceptual model to say that self-knowledge is like clairvoyance. If perceptual knowledge that P requires you to be conscious that P as well as a degree of cognitive effort then the net effect of admitting that Norman’s knowledge that he believes the President is in New York is epistemologically and phenomenologically on a par with his knowledge that the President is in New York is to suggest that his self-knowledge is not perceptual. It lacks some key features of ordinary perceptual knowledge and only comes out as “perceptual” on an impoverished view of perceptual knowledge which has very little going for it.

It’s also worth pointing out that there are significant chunks of self-knowledge which clearly don’t fit the broad perceptual model, on the assumption that this model is prepared to live with the accusation that it treats introspection as akin to clairvoyance. For example, if the question arises whether you truly believe that all races are equal or whether you want another child you may know that the answer is ‘yes’ because you have thought about these things before. But if you haven’t thought about them before then knowing what you believe or want in such cases isn’t a matter of finding yourself with the yen to believe that you believe that all races are equal or that you want another child. To take such questions seriously is to be prepared to do some thinking; it isn’t like putting money into a slot machine and waiting to see what comes out. Self-knowledge can be a hard-won cognitive achievement, and the broad perceptual model doesn’t have anything to say about this.

Cases in which self-knowledge is a hard-won cognitive achievement are cases of what I have called substantial self-knowledge. From the fact that the broad perceptual model is ill-equipped to account for substantial self-knowledge it doesn’t follow that it is ill-equipped to account for trivial self-knowledge. There is always the possibility of a hybrid approach which sees self-knowledge, even knowledge of one’s own attitudes, as being sometimes perceptual and sometimes inferential. However, given all the objections to the broad perceptual model I have been discussing it would be worth considering another possibility: what I have in mind is the possibility that even supposedly “trivial” self-knowledge is not perceptual but is at least implicitly inferential. I’ve emphasised that your knowledge that P can be inferential even if you aren’t conscious of inferring that P. Could it not be, then, that when it comes to knowing what you want or believe or hope inference is always involved, even in straightforward cases, and that various attempts to account for the immediacy of self-knowledge are doomed at the outset because self-knowledge is not in fact immediate? This is the inferentialist approach to self-knowledge which I’m going to discuss in the next chapter.