

Self-Knowledge A Beginner's Guide



Quassim Cassam

TABLE OF CONTENTS

Page 3:	Introduction
Page 4-5:	Getting Started
Page 6-8:	The Philosophical Picture of Self-Knowledge
Page 9-11:	The Specialness of Self-Knowledge
Page 12:	How Is Self-Knowledge Possible?
Page 13-14:	The Perceptual Model of Self-Knowledge
Page 15-16:	The Inferential Model of Self-Knowledge
Page 17-18:	Problems for Rationalism
Page 19-21:	Substantial Self-Knowledge
Page 22-23:	Knowing your Emotions
Page 24-25:	Knowing Why
Page 26-27:	Self-Ignorance
Page 28-29:	The Value of Self-Knowledge
Page 30:	The True Self
Page 31:	What Next?

INTRODUCTION

This beginner's guide is about the philosophy of self-knowledge. Philosophers aren't the only people who think and write about self-knowledge. Psychologists have plenty to say about it, and I will discuss psychological approaches to self-knowledge later in this guide. You can also gain insights about the nature and sources of self-knowledge by reading great literature, such as Proust's *Remembrance of Things Past* or Jane Austen's *Emma*. In comparison with psychological and literary approaches to self-knowledge, the philosophy of self-knowledge is dry and difficult. One reason is that philosophers have tended to concentrate on relatively trivial self-knowledge, as distinct from what I call 'substantial' self-knowledge. In this article I explain why philosophers find trivial self-knowledge so interesting, despite the fact that it isn't even what we would ordinarily call 'self-knowledge'. Substantial self-knowledge is much closer to self-knowledge in the everyday sense, and I'll suggest that philosophers should pay much more attention to self-knowledge in this sense.

The philosophy of self-knowledge I'll be discussing in this article is Western rather than Eastern philosophy. Self-knowledge is a major topic in Indian, Chinese and Islamic philosophy but my own expertise is in what is called the Western 'analytic' tradition. Many of the ideas in this guide are explained in greater detail in my book *Self-Knowledge for Humans*, published by Oxford University Press in 2014.

GETTING STARTED

What is self-knowledge? One natural and popular answer to this question is that it is knowledge of what is sometimes called your 'true self'. This is what I've been referring to as self-knowledge in the everyday sense. On this way of thinking, your true self is the real 'you', and is made up of your true character, values, desires, emotions and beliefs. It is what you are really like, as distinct from what other people, or indeed you, believe you are like. The true self is not the self as it appears to others or to itself but the self as it really is. The search for self-knowledge is, you may well think, the search for your true self.

This way of thinking assumes, of course, that there is such a thing as the 'true self' and that is something which some people have questioned. Leaving this issue to one side for the moment, another feature of the ordinary way of thinking about self-knowledge is the assumption that self-knowledge isn't easy to come by. To have knowledge of your true self would be a genuine 'cognitive' or intellectual achievement, and one which requires time and effort. And once you think of self-knowledge as effortful, there is a natural further question which arises: what is the point or value of self-knowledge? What good does it do you to have it and what would you lose by not having it?

It's often taken for granted that self-knowledge is worth having. Some people believe that self-knowledge is valuable because, to put it crudely, you will be happier with it than without it. But that's not obvious; maybe there are some truths about yourself you would be happier and better off not knowing. Perhaps, in that case, we need to look elsewhere to explain the value of self-knowledge. You might think of Socrates' suggestion that

self-knowledge is needed to live meaningfully or the related idea that you need self-knowledge to live 'authentically', that is, in a way that is true to yourself, to your true self. Whether such explanations of the value of self-knowledge are correct sounds like an excellent question for a philosopher of self-knowledge.

THE PHILOSOPHICAL PICTURE OF SELF-KNOWLEDGE

Surprisingly, perhaps, the issues I have just been describing have not been the focus of (Western) philosophical accounts of self-knowledge. Their focus, at least since the 17th century, has been on other things. Suppose you *believe* you are wearing socks and know that that's what you believe. Your belief that you are wearing socks is one of your current 'states of mind', and many philosophers would want to say that knowing that you believe you are wearing socks is a form of 'self-knowledge'. The same goes for knowing that you have a headache or that you want to go to see a movie this evening. These are all examples of 'self-knowledge' in the philosophical sense. Notice that in the socks example what is at issue isn't whether you know that you are wearing socks but whether and how you know that you believe you are wearing socks.

This view of self-knowledge might come as a surprise to you. For a start, it's hard to believe that knowing that you believe that you are wearing socks is knowledge of your true self, the real you. No doubt you have 'deeper' beliefs which are part of your true self but surely not the belief that you are wearing socks. Unlike the self-knowledge I started by describing, knowing that you believe you are wearing socks or that you have a headache or that you want to see a movie this evening look like pretty boring and trivial pieces of self-knowledge. On the face of it, such trivial self-knowledge isn't hard to acquire and doesn't represent much of a cognitive achievement. Nor does it seem particularly useful or valuable. What possible use is it to know that you believe you are wearing socks?

Thinking a bit more about such examples, it's natural to distinguish

between 'trivial' and 'substantial' self-knowledge. There is much more about this distinction in chapter 3 of *Self-Knowledge for Humans*, where I suggest that the difference between the two kinds of self-knowledge is a difference in degree and that many different considerations bear on whether a given piece of self-knowledge is 'trivial' or 'substantial'. Intuitively, knowing that you believe that you are wearing socks or that you feel like watching a movie are examples at the more trivial end of the spectrum. The suggestion here is not that that knowledge of your own beliefs and desires is *always* trivial but that it often is. In contrast, substantial self-knowledge includes knowledge of your own character, values, abilities and emotions. Examples might include: knowing that you are a kind person, that you aren't cut out for your current job, or that you harbour deep feelings of resentment towards a sibling. You might think that substantial self-knowledge is knowledge of your 'true' self but you don't have to think that way. The important point about substantial self-knowledge is that it represents a genuine cognitive achievement and has an obvious claim (which may or may not turn out to be right) to being regarded as valuable.

In these terms, much recent philosophical discussion has been about trivial rather than substantial self-knowledge. Why is that? Because many philosophers since Descartes in the 17th century have been struck by the *peculiarity* or *specialness* of trivial self-knowledge. They think that its specialness distinguishes trivial self-knowledge from substantial self-knowledge and from knowledge of other kinds. On this approach, the philosophical challenge is to explain how such 'special' self-knowledge is possible. Viewed from this perspective, substantial self-knowledge might still be *humanly* important, important to us as human beings, but is less interesting *philosophically* because it doesn't have the supposed specialness of trivial self-knowledge. Rather, it is

much more like any other knowledge that has nothing to do with the self.

The next question, therefore, is whether it is true that trivial self-knowledge is special. Even if it is true, that is not an excuse for neglecting substantial self-knowledge. Given the importance to us as human beings of substantial self-knowledge, you would have thought that philosophers of self-knowledge would have much more to say about it than they actually do. In *Self-Knowledge for Humans* I argue that philosophy tends to overestimate the specialness of trivial self-knowledge and underestimate the philosophical interest of substantial self-knowledge. I will come back to this. First, we need to get clear about the supposed specialness of some forms of self-knowledge.

THE SPECIALNESS OF SELF-KNOWLEDGE

Our beliefs about the world around us are fallible. You might believe it is raining even though it isn't. You might *believe* you are wearing socks even though you aren't. These beliefs aren't immune to error, even though mistakes about such matters might seem highly unlikely. But now consider the question whether you believe you *are* wearing socks, as distinct from the question whether you are wearing socks. Your belief that you are wearing socks could be mistaken, but it has seemed to many philosophers that your belief *that you believe* you are wearing socks couldn't be mistaken; you can't be wrong about what you yourself believe. The same goes for other states of mind. Suppose you think you have a headache. Can you be wrong about that? Surely not. You can be wrong about *why* your head hurts but you can't be wrong about *whether* your head hurts. One way of putting this would be to say that your knowledge of such states of mind is *infallible*.

Nowadays, even philosophers who think that trivial self-knowledge is special tend to be sceptical about the idea that it is infallible. What they say is that although genuine mistakes about our own states of mind are possible such mistakes are abnormal, and that there is a presumption that our beliefs about what we believe or want or feel aren't mistaken. One way of putting this is to say that some self-knowledge is *authoritative*, even if it isn't strictly infallible.

In addition, it's often claimed that what I'm calling 'trivial' self-knowledge is special in that it isn't normally based on behavioural or other evidence. You don't need to observe your own behaviour to know that you believe you are wearing socks or that you want to go to the cinema this evening. These are things you know 'immediately', rather than by inference. You

don't normally need to work out that you believe you are wearing socks, and evidence doesn't come into it. This *immediacy* of some self-knowledge is another respect in which it is special – or so many philosophers have thought.

If trivial self-knowledge is special in some or all of these ways then it is markedly different from substantial self-knowledge, as well as knowledge of minds other than our own. Suppose you think of knowledge of your own character traits as an example of substantial self-knowledge, and let's suppose that kindness is one such trait. Are you infallible about whether you are a kind person? Clearly not. Your sincerely held belief that you are kind might be mistaken. Are you authoritative about your own character traits, in the sense that there is a presumption that your beliefs about your own character traits aren't mistaken? Surely not. We all like to think well of ourselves, and this threatens the presumption that we know what we are really like. In addition, to know your own character you need evidence, including behavioural evidence, and this means that knowledge of your own character is also not immediate.

The same goes for knowledge of minds other than your own. You need evidence – usually behavioural evidence - to know what someone else thinks or feels, and your beliefs about such matters are neither infallible nor authoritative in the way that self-knowledge is authoritative. So there is what Paul Boghossian describes as a 'profound asymmetry between the way in which I know my own thoughts and the way in which I may know the thoughts of others'. A philosophical account of self-knowledge must surely acknowledge and explain this asymmetry.

To sum up, I've explained the philosophical interest in trivial self-knowledge on the basis of the supposed specialness of this type of knowledge rather than its importance. This isn't to say that philosophers who focus on trivial self-knowledge think that it is *unimportant*. For example, it has been argued by Sydney

Shoemaker that what I've been calling 'trivial' self-knowledge is necessary for rationality. Whether this is actually so is not a straightforward matter and turns on very difficult questions about the nature of rationality. Suffice to say that the thesis that self-knowledge is a precondition of rationality remains highly controversial.

HOW IS SELF-KNOWLEDGE POSSIBLE?

Suppose you have followed the discussion so far and buy the idea that trivial self-knowledge is philosophically interesting on account of its specialness. Then the natural follow up questions include: how is authoritative and immediate self-knowledge possible? What are the limits of such self-knowledge and what explains its authority and immediacy? The first of these questions seems pressing, at least for philosophers, because most of our knowledge isn't authoritative or immediate in the way that some self-knowledge is (supposedly) authoritative and immediate. So there is something here which needs explaining; it's not enough to point out that some self-knowledge is special. We also want to understand how this kind of self-knowledge is possible.

One possibility would be to explain how authoritative and immediate self-knowledge is possible by identifying its sources, that is, by figuring out how we get it. One suggestion is that we know our own minds by introspection, where this is thought of as a form of inner perception. This is the *perceptual* model of self-knowledge. Another possibility is that we know our own minds by inference or reasoning. This is the *inferential* model of self-knowledge. Each of these models has many different variants, and they aren't mutually exclusive. If perception involves inference, then saying that self-knowledge is perceptual is compatible with saying that it is inferential. But for present purposes let's consider these models separately.

THE PERCEPTUAL MODEL OF SELF-KNOWLEDGE

The perceptual model of self-knowledge was endorsed by John Locke in the 17th century and, with complications, by Immanuel Kant in the 18th century. A distinguished 20th century proponent of the perceptual model was the Australian philosopher D. M. Armstrong in his book *A Materialist Theory of Mind*, originally published in 1968. Usually when philosophers talk about perception they mean *sense* perception, that is, seeing, hearing, touching, tasting and smelling. Introspection, on this account, is a kind of inner seeing, seeing with what is usually referred to as the 'mind's eye'. You know your own beliefs, desires, feelings and so on by introspecting them, by seeing with your mind's eye that you have certain specific beliefs, desires, feelings and so on.

This model is good at accounting for the immediacy of self-knowledge but less good at accounting for its authority. If you know that you are wearing socks by seeing that you are wearing socks then you might think (although this is controversial) that your knowledge is 'immediate' rather than 'inferential'. So if you can "perceive" that you *believe* you are wearing socks or that you have a headache then the result self-knowledge – the knowledge that you believe that you are wearing socks or that you have a headache- will presumably also be 'immediate'.

The perceptual model is less good at accounting for the authority of self-knowledge unless you think that perceptual knowledge is authoritative. There are many other objections to the perceptual model which all turn on the idea that the means by which we know our own minds are radically

disanalogous to perception. Introspection, the objection goes, is nothing like perception, and our own states of mind aren't there to be perceived in anything like the way things like socks are there to be perceived. Sydney Shoemaker is an influential recent critic of the perceptual model in his 1996 book *The First-Person Perspective and Other Essays*.

THE INFERENTIAL MODEL OF SELF-KNOWLEDGE

The inferential model of self-knowledge says that we come to know our beliefs, desires and feelings by inference. Inference from what? From various types of evidence that are available to us. The philosopher Gilbert Ryle and the psychologist Darryl Bem both claim that we rely on behavioural evidence. Ryle defends this view in his famous book *The Concept of Mind* (1949). This view has been widely denounced by philosophers of self-knowledge on the grounds that you don't need to observe your own behaviour to know how you feel or what you think. However, this leaves open the possibility that self-knowledge is acquired by inference from other types of evidence, including psychological evidence. This is the view defended in *Self-Knowledge for Humans*, chapters 11 and 12.

A variant on the inferential model is the view that self-knowledge is acquired by employing what is called the 'Transparency Method'. This says, very roughly, that a way to establish whether you believe that P (that you are wearing socks) is to ask yourself whether you ought rationally to have this belief. If the answer to the latter question is 'yes', then you can conclude that you do believe that P. In drawing this conclusion you are relying on the assumption that what you do believe is determined by what you ought rationally to believe. This is what is sometimes called a Rationalist conception of self-knowledge, and its main proponent is the Harvard philosopher Richard Moran. His book *Authority and Estrangement* (2001) is well worth reading. A notable critic of Moran is David Finkelstein in his book *Expression and the Inner*, published in 2003.

The inferential model is at odds with the idea that self-knowledge is normally immediate. Since there is nothing special about

inferential knowledge *per se*, and our knowledge of minds other than our own is also inferential, proponents of the inferential model tend to be sceptical about the specialness of self-knowledge; they doubt that self-knowledge is different in kind from knowledge of other minds. However, Richard Moran argues (mistakenly in my view) that his Rationalist approach can account for the authority and immediacy of self-knowledge. There is more about this issue in chapter 9 of *Self-Knowledge for Humans*. It's worth adding that Rationalists don't regard themselves as inferentialists, though it's hard to see how the Transparency Method can deliver non-inferential self-knowledge. This remains a tricky issue, partly because the distinction between inferential and non-inferential knowledge is so obscure.

PROBLEMS FOR RATIONALISM

An additional complication for Rationalism is that human beings are imperfectly rational, and that there is therefore no guarantee that our beliefs, desires, fears and so on are as they ought rationally to be. Suppose the question arises whether you want to go to the gym this afternoon. Since you have been advised by your doctor that you need to take more exercise, going to the gym might be what you *ought* rationally to want to do. Unfortunately, going to the gym this afternoon might still be the last thing you *actually* want to do. If that is correct, then asking yourself whether you ought rationally to want to go to the gym looks like a bad way of figuring out whether you do want to go, just as asking yourself whether you ought rationally to fear the spider in your bathtub looks like a bad way of establishing whether you do fear it. Maybe you are perfectly well aware that you have no reason to fear it but you still fear it.

The best case for the Transparency Method is knowledge of our own beliefs. If you recognize that you ought rationally to believe that you are wearing socks (because all the evidence points that way) then it's probably safe to conclude that you do believe you are wearing socks. Even so, the possibility can't be ruled out that you continue to hold on to certain beliefs which matter to you in the face of overwhelming undermining or contrary evidence. This phenomenon, which psychologists call 'belief perseverance', causes problems for the idea that you can know your own beliefs by reflecting on what you have reason to believe.

The point here is that, as fallible human beings, we don't always think or want or fear what we ought rationally, even by our own lights, to think or want or fear. A further problem for

Rationalism is that it is often easier to know what our attitudes are than to know what they ought rationally to be. Rationalism is, at best, an account of self-knowledge for a mythical *homo philosophicus*, whose beliefs and other attitudes are always as they ought rationally to be. Rationalism looks less good as an account of self-knowledge for humans. This point is developed at length in *Self-Knowledge for Humans*.

SUBSTANTIAL SELF-KNOWLEDGE

I've now said just about as much as I'm going to say about the 'trivial' self-knowledge which has been the focus of so much philosophical attention. I've said that philosophers have been interested in trivial self-knowledge because it looks special, and I've considered various different accounts of how such self-knowledge is possible. On some of these accounts, such as the inferentialist account, trivial self-knowledge is less special than you might think, and this makes it less excusable that philosophers spend so much time on trivial self-knowledge and so little time on the more humanly important varieties of substantial self-knowledge. My examples of substantial self-knowledge are knowledge of your own character, values, abilities and emotions. I've said that the difference between substantial and trivial self-knowledge is one of degree but I haven't said enough about the basis on which any self-knowledge should be regarded as substantial.

Take the case of knowledge of your own character traits. What makes your knowledge of your own character traits (e.g. kindness, gullibility, fastidiousness etc.) more substantial than your knowledge that you believe you are wearing socks is that your beliefs about your own character traits are more prone to error, and more obviously open to challenge or correction by others. Your nearest and dearest might have a deeper insight into your character than you do, and that is because there are familiar obstacles to knowing your own character, such as repression, self-deception, bias and embarrassment. You need evidence to know your character, and normally much greater mental effort is required to know what kind of person you are than to know that you believe you are wearing socks. Substantial self-knowledge is tricky because, as Eric Schwitzgebel points out, it tangles with a person's

self-conception. Finally, you might think that knowledge of your own character or other varieties of substantial self-knowledge is much more valuable in a practical or even a moral sense than your knowledge that you believe you are wearing socks.

How, then, is substantial self-knowledge possible? Perhaps the answer will be different for different examples of substantial self-knowledge. Perhaps the basis on which you know you are kind is different from the basis on which you know that you harbour deep feelings of resentment towards a sibling. Nevertheless, what does seem reasonably clear is that neither the perceptual model nor the Transparency Method can account for substantial self-knowledge. You can't know that you are kind through introspection or inner perception and you can't know whether you have deep feelings of resentment towards a sibling by considering whether you ought rationally to have such feelings. When it comes to substantial self-knowledge, inferentialism is the only game in town.

From what do you infer your own character traits or values or emotions? It's often assumed that you infer such things exclusively from your own behaviour, but this is wildly implausible. Behavioural evidence certainly comes into it but so do other kinds of evidence. You might also infer your character and values from how you think, what you care about, and how certain things make you feel. This is a reflection of the fact that character traits and values aren't just dispositions to act. To infer a character trait from your thoughts, feelings and what you care about, you need to see the evidence in your possession as evidence that you have that trait, and this means that you need a 'theory' or understanding of the trait in question. On this account, substantial self-knowledge is acquired by theory-mediated inferences from the psychological and other evidence available to you. This means that substantial self-knowledge requires a degree of intellectual sophistication. You can be manipulative or fastidious without understanding what

these things are but *knowing* that you are manipulative or fastidious is a different matter. There is much more about all this in chapter 13 of *Self-Knowledge for Humans*, where I discuss the example of knowing that you are fastidious.

One consequence of this account of substantial self-knowledge is that it presupposes access to the various kinds of evidence, including psychological evidence, from which you infer your character traits, values and so on. Is your access to that evidence also inferential, and, if so, is that a problem? The answers to these questions are 'yes' and 'no'. Suppose you need knowledge of your own thoughts in order to know what kind of person you are. As Peter Carruthers argues in his book *The Opacity of Mind* (2011) access to our own thoughts is interpretational (we have to interpret them to know them) and in this sense inferential. This is only a problem for inferentialism if you assume that inferential self-knowledge must ultimately be grounded in knowledge that isn't inferential. I argue against this assumption in chapter 12 of *Self-Knowledge for Humans*. As far as the inferential model is concerned, it's inference all the way.

Before moving on, I need to acknowledge that the whole question of how you know your own character traits only arises if there are character traits. Inspired by the work of social psychologists, Gilbert Harman argues that 'there is no reason at all to believe in character traits as ordinarily conceived'. This isn't the place for a discussion of scepticism about character. Suffice to say that Harman's arguments for this startling claim are far from decisive and that, in any case, knowledge of one's own character traits is only one example of substantial self-knowledge. It's now time to look at some others.

KNOWING YOUR EMOTIONS

I've described knowledge of your own emotions as substantial self-knowledge but you might think that that can't possibly be right. Surely it doesn't require cognitive *effort* to know that you love someone, it's simply obvious to you. In addition, if you sincerely believe that you love them then no one else can show that your belief is mistaken. You just know, and there is no need for inference. On this account, you have certain strong emotions your knowledge of which looks far from substantial, and love is a case in point.

It isn't an objection to inferentialism *per se* that self-knowledge of emotions isn't always substantial. Perhaps it is sometimes substantial and sometimes not, depending on the emotion in question. Having said that, it's worth pointing out that it isn't always obvious how you feel about someone. Sometimes cognitive effort plainly *is* required, and your friends may know better than you do how you feel about someone. You might think it's love but they know you well enough to know that it's just a passing infatuation. In such cases, knowledge of your own emotion is a matter of inference and interpretation. Proust gives a famous example of this. Marcel thinks that he is indifferent to Albertine, but his anguish on discovering that she has gone tells him otherwise. On an inferentialist interpretation, Marcel comes to know that he loves Albertine by inference from his anguish on hearing that she has gone.

This approach is criticised by Martha Nussbaum in her book *Love's Knowledge*. Her target is what she calls 'intellectualism', the view that knowledge of whether one loves another person can best be attained by a 'detached, unemotional, exact intellectual scrutiny of one's condition,

conducted in the way a scientist would conduct a piece of research'. For Nussbaum, 'knowledge of the heart must come from the heart'. Thus, Marcel's anguish is itself a piece of self-knowing rather than a piece of psychological evidence from which he can infer that he loves Albertine. In Nussbaum's words, 'love is not a structure of the heart waiting to be discovered'.

I argue against Nussbaum's view of love's knowledge in chapter 13 of *Self-Knowledge for Humans*. I maintain that the intellectualism which Nussbaum criticizes is in fact correct, and that Marcel's anguish is, just as inferentialism claims, evidence of his underlying emotional state. Regardless of whether I am right about this, Nussbaum's discussion is a wonderful illustration of how deep, rich and interesting the philosophy of self-knowledge can be when its practitioners put their minds to it. It isn't, or shouldn't be, all about how you know that you believe you are wearing socks.

KNOWING WHY

Knowing *what* you want or believe or feel is one thing. Knowing *why* you believe what you believe or want what you want is a different matter. Much of the philosophical literature on self-knowledge is about 'knowing what' but 'knowing why' is no less interesting. In the case of beliefs, Rationalists focus on reasons: to know why you believe you are wearing socks is to know your reasons for believing you are wearing socks, that is, the reasons for which you have that belief. The parallel here is with actions: to know why you acted as you did (e.g. put on your socks) is to know the reason or reasons for which you acted as you did. There is more about this in Jonathan Dancy's book *Practical Reality*. For present purposes, the point is that as long as you have access to your reasons you have an answer to the 'why' question. The challenge is then to explain how you know your own reasons.

But now consider the case in which your reasons are no good and in which what you believe isn't what you ought rationally to believe. In such cases, it is natural to think that why you believe what you believe has less to do with your reasons and more to do with other factors. For example, some social psychologists explain belief in conspiracy theories by reference to a 'conspiracy mentality', a 'general propensity towards conspiratorial thinking' (Imhoff and Bruder). Yet the individual conspiracy theorist neither knows nor believes that he believes his pet conspiracy theories because he has this or any other intellectual character trait. In this sense he doesn't know why he believes his conspiracy theories.

This issue isn't just of academic interest. *Why Do People Believe Weird Things?* is the title of a book by Michael Shermer, and Shermer's question is an important one. For example, what do we make of 9/11 conspiracy theorists who believe that the 9/11 attacks weren't the work of al Qaeda or, to take an more extreme example, people who believe they were once abducted by aliens? In such cases, explanations in terms of

reasons only get you so far. The 9/11 conspiracy theories has his reasons, but there is a range of alternative explanations in terms of his character, situation or psychology. To the extent that such explanations are unknown by the conspiracy theorist, and would be rejected by him, he doesn't know why he believes what he believes.

The extent to which peoples' beliefs and actions can ever be explained by reference to their character traits is controversial. In *The Person and the Situation* Lee Ross and Richard Nisbett discuss what they call the 'fundamental attribution error', the error of inflating the significance of character traits in the explanation of behaviour and underestimating the significance of situational factors. Explaining beliefs by reference to intellectual character traits might be a form of this error but this doesn't affect the point I am making here: whether we place the emphasis on character traits or on situational factors, the influence of non-rational factors on our beliefs and other attitudes is something to which we are often blind. Such blindness is a form of self-ignorance, which is the other side of self-knowledge. This raises a more general question: just how self-ignorant are we?

SELF-IGNORANCE

It might come as a surprise that philosophers of self-knowledge generally say little about self-ignorance (though Eric Schwitzgebel is a notable exception). This is partly the legacy of Descartes, who is thought to have thought that our own states of mind are *self-intimating*. What this means is that if you believe that you are wearing socks, and have the concept of belief, then you can't fail to know that you believe you are wearing socks. Similarly, if you want something then you can't fail to know that you want it, and if you hope for something you can't fail to know that you hope for it. Self-ignorance is ruled out: you can't fail to know your own mind unless there is something wrong with you.

Freud's work makes this view of self-ignorance very hard to accept, and Nietzsche is someone else who emphasizes the prevalence of self-ignorance. More recently in philosophy, the idea that mental states are self-intimating (or 'luminous') has been criticized by Timothy Williamson in his book *Knowledge and Its Limits*. The upshot is that, far from being abnormal or unusual, self-ignorance is a pervasive feature of our lives. It isn't just a matter of not knowing *why* our beliefs or other attitudes are as they are. There is also the possibility of not knowing *what* one believes or wants or fears. Self-ignorance in this sense can be motivated but needn't be. For example, your failure to acknowledge a particular attitude of yours might be motivated by an unconscious desire to avoid the psychic pain or distress that you would suffer if you were to acknowledge that attitude. On the other hand, your self-ignorance might also be explained by a lack of evidence, or your inability to grasp the significance of the evidence available to you. Either way, you might fail to know your own attitude, just as you might fail to know your own

character or abilities. For humans, self-ignorance is always on the cards.

Once the prevalence of self-ignorance is acknowledged it is then an interesting question whether and how self-ignorance can be overcome. Are talking therapies the answer or would something more prosaic be as effective? The key is to identify the obstacles to self-knowledge and the appropriate means of overcoming them. Self-observation, social interaction and the arts all have a part to play, but you won't get far by trying to introspect your own states of mind. There is more about all this in chapter 14 of *Self-Knowledge for Humans*.

THE VALUE OF SELF-KNOWLEDGE

Since overcoming self-ignorance can require effort it would be natural to wonder whether it is worth the effort. What is so good about self-knowledge and so bad about self-ignorance? You could try taking the high road and argue that you need self-knowledge for your life to be meaningful or authentic. For example, to be authentic is to be true to yourself but how can you be true to yourself unless you know yourself? Such 'high road' explanations of the value of self-knowledge are seductive but flawed. To be true to yourself is, among other things, to behave in ways that are consistent with, and a reflection of, your 'true self' but why do you need to *know* yourself in order to *be* yourself? Why isn't it enough that your behaviour *is* a reflection of your true self, regardless of whether you know or believe that this is the case? Indeed, focusing on your true self, and on what you need to do to be true to it, might end up making your behaviour *less* authentic than it would otherwise be.

An alternative 'low road' explanation of self-knowledge says that its value doesn't derive from the value of 'high' ideals such as authenticity and the meaning of life. Instead, the value of self-knowledge is practical. You will be happier and more effective in your life if you know yourself. There is some evidence of this in research by psychologists. Elizabeth Dunn and Timothy Wilson report that while mild self-illusions can be beneficial by motivating self-improvement, extreme self-illusions can undermine well-being. On the whole, self-knowledge is better for you than self-ignorance. This assumes that the alternative to knowing yourself is having *false* beliefs about yourself. But what if the alternative is having *no* beliefs about yourself, a completely unreflective existence in which you never think about, and so have no illusions about, your own

character, abilities emotions, and so on? Would you necessarily be less happy than you would be with self-knowledge? Perhaps this question is academic since humans can't help thinking about themselves and seeking self-knowledge.

The distinction between 'high road' and 'low road' explanations of the value of self-knowledge is drawn and developed in chapter 15 of *Self-Knowledge for Humans*, which makes the case for low road explanations.

THE TRUE SELF

I started by saying self-knowledge in the everyday sense is knowledge of your true self. What should we make of this idea now? There is the view that there is no such thing as the 'true' self and that the self in this sense is an illusion. Bruce Hood talks about this in his book *The Self Illusion*. The illusory self is what Derek Parfit calls a 'separately existing entity', that is, something that is distinct from one's brain and body and the series of mental and physical events that make up one's life. An example of such an entity might be a soul, which is what Descartes thought each of us fundamentally is. The self in this sense is indeed an illusion but it doesn't follow that the *self* is an illusion. Instead of saying that there is no self we can say that we aren't separately existing entities. Each 'self' is no more and no less than the sum of his or her thoughts, actions, attitudes, emotions, abilities, values and physical characteristics.

This is an extension of Hume's view that what is called the 'self' is a bundle of sensations, and it still allows for the idea of the *true* self. Your true self is simply made up of your *true* or *actual* character traits, attitudes, emotions, abilities and so on. To know these things about yourself is to know your true self, the real you, and the self-knowledge that is at issue here is substantial. Given the reality of self-illusions, false beliefs about one's self, it is perfectly legitimate to talk about the 'true self' or the self as it really is, as distinct from the self as it appears to itself or to others.

Once you think of self-knowledge in the way that I have been recommending, it should be evident that it is, or can be, hard to acquire. The substantial self-knowledge which is the core of knowledge of the 'true self' is the self-knowledge we actually care about. Whether and how such self-knowledge is possible is, or should be, a fundamental question for the philosophy of self-knowledge.

WHAT NEXT?

If you have read this far and still want to know more about the philosophy of self-knowledge you should now go to the Resources Section of this website, which suggests further reading on each of the topics I have discussed in this Beginner's Guide. You might also want to look at my entry on Self-Knowledge in [Oxford Bibliographies Online](#).

