

This is the author's copy of a review published in 2016 in *Dialectica*, vol. 70(1), pp. 113-119.

**Review of *Self-Knowledge for Humans* by Quassim Cassam (Oxford University Press, 2014)**

The thesis that self-knowledge is acquired through inference has had its advocates, but has never reached the level of a mainstream view in the philosophy of self-knowledge. In *Self-Knowledge for Humans* Quassim Cassam rises to the challenge of defending an 'inferentialist' theory of self-knowledge, but one which is largely restricted to explaining knowledge of our own 'standing attitudes': beliefs, desires, hopes, fears, etc., as well as the objects of 'substantial self-knowledge'. Written in an impeccably clear style, this book leaves the reader with the impression that inferentialism should be taken seriously. Below I outline the book's trajectory, mention its strengths, and raise some worries along the way.

Cassam distinguishes between *trivial* self-knowledge (e.g., knowing that you believe you're wearing socks), and more hard-to-get *substantial* self-knowledge, which includes cases of knowing your own character, aptitudes, emotions, and what makes you happy (e.g., knowing whether your feelings towards someone is love or a passing infatuation). He points out that substantial self-knowledge has been neglected by philosophers, who focus on trivial self-knowledge mainly because of its apparent epistemologically distinctive features (immediacy, first-person authority etc.). A useful list of ten characteristics of substantial self-knowledge is given (pp.30-32). Importantly, Cassam points out that no particular kind of self-knowledge, such as knowledge of one's beliefs or emotions, is trivial or substantial as a rule: 'it all depends on the content of the belief' (p.33). For instance, knowing that you want some vanilla rather than chocolate ice-cream would usually be of the trivial sort, while knowing that you want another child would be substantial (p.33).

**Critical section:** Chapters 1 to 10 are mainly critical. Though the inner-perception theory of self-knowledge gets targeted, Cassam's main prey is 'Rationalism' about self-knowledge, which he associates with the 'transparency method' (TM) explanation of how we acquire (some) self-knowledge, and also the thesis that self-knowledge of attitudes is normally immediate (i.e., non-inferential/not based on evidence). He mainly works with David Finkelstein's formulation of TM: 'The question of whether I believe that P is, for me, transparent to the question of what I ought rationally to believe—i.e. to the question of whether the reasons require me to believe that P. I can answer the former question by

answering the latter' (quoted on p.4), a formulation which can be adapted for other types of standing attitudes.

Cassam accuses Rationalism of implicitly relying on an idealised conception of humans as generally believing, desiring etc., what they rationally ought to (p.51). Why? Because 'TM says that you can determine what your attitude is in a given case by determining what it ought rationally to be. This requires the assumption that your beliefs and other attitudes *are* as they ought rationally to be' (p.83). This idealised conception he dubs *Homo philosophicus*, which contrasts with *Homo sapiens*, or how humans actually are. The difference between *Homo philosophicus* and *Homo sapiens* he calls 'the Disparity'.

According to Cassam, many contemporary philosophers influenced by Quine and Davidson believe that there can't be an extensive Disparity between us and *Homo philosophicus* because, as Bill Child puts it 'if a subject has attitudes at all, the relations amongst her attitudes, perceptions, and actions must be by and large rational' (quoted on p.62). Cassam takes Child to be saying that if 'we have beliefs and desires then by and large they must be as they ought to be' (p.62). But Cassam shows, partly by appealing to psychological studies, that our beliefs etc., frequently are not as they rationally ought to be. For instance, our beliefs are often influenced, unbeknownst to us, by non-rational factors, and we often rely on 'fast', heuristic thinking which can lead us into error.

There may be an ambiguity with 'rational' causing confusion here. One can detect a notion of rationality in use in philosophy which is much weaker than the ordinary notion of rationality which Cassam operates with. Sometimes by an attitude being rational philosophers just mean that it's based on reasons (whether good *or bad*). Alternatively, by 'rational relations' Child may simply mean relations of *consistency* between attitudes. This contrasts with a stronger notion of rationality, where a rational belief is one formed according to proper epistemic canons, with careful, unbiased, or dispassionate reflection, with an awareness of basic statistical principles, common fallacies, etc. And a person's beliefs can simultaneously be rational in the weak sense but not in the strong.

Consider, for instance, Cassam's case of Oliver, a 'conspiracy nut' who believes the collapse of the World Trade Centre towers on 9/11 was caused by explosives planted by government agents (pp.24-25). Though Oliver's belief may not be as it rationally ought to be, in that it was formed due to a bias to believe conspiracy theories and gullibility (p.25), it may be rational in the weak sense of being based on (perhaps shoddy) reasons and being mostly consistent with his other beliefs. It is part of a consistent, if warped, world-view. Of course, there may be inconsistency or tension between this belief and another of Oliver's

commitments. But such inconsistencies exist usually because they are either rationalized or remain unnoticed (Socrates, as we know, was adept at exposing such unnoticed internal inconsistencies). And it is indeed difficult to imagine massive or conspicuous inconsistency between a subject's beliefs, actions, intentions etc.; if Oliver later said things which flatly contradicted his account of 9/11, we would wonder, 'How can he believe that and also what he said before?'

So it's not clear that evidence of a large Disparity undermines Child's point. Anyhow, does TM really require 'the assumption that your beliefs and other attitudes *are* as they ought rationally to be'? The Rationalist could say no, since Oliver could use TM and gain knowledge of what he believes, even though that belief is not one he rationally ought to have. He sets about reviewing the evidence concerning 9/11 and concludes that the evidence points to government involvement. From this he can know, according to TM, that he believes the government were involved. But this is compatible with the possibility that his review of the evidence was biased, objectively speaking.

Cassam is aware of this move by the Rationalist: in such cases, 'you still believe what you ought rationally to believe *by your own lights* so you can still determine what you believe in such cases by determining what *you think* you ought rationally to believe' (p.55, emphasis added; also see pp.108-109). This would be to argue for what he calls Compatibilism: the view that Rationalism can live with a large Disparity between *Homo philosophicus* and *Homo sapiens*.

The key question for Cassam then becomes: whenever you think the reasons indicate that P, and hence that you ought rationally to believe that P, do you actually believe that P? In that case, in knowing that the former holds, you could know that the latter holds. That these go hand-in-hand is, indeed, *prima facie* quite plausible. After all, what more do we need to feel entitled to believe that P, then to find that this is what the evidence indicates?

Cassam argues that the idea that the attitude we have matches with the attitude we think we rationally should have is least convincing for attitudes other than belief. Someone may fear spiders while knowing that there is no reason to fear spiders. Someone may want another martini while knowing that he 'shouldn't want' another martini (p.110). Such people cannot determine that they fear/desire *x* by determining whether they rationally ought to fear/desire *x*. Thus the explanatory scope of TM, at least under this formulation, is seriously limited for Cassam.

The self-knowledge of belief, which TM was first developed to explain, is trickier since the idea that one could believe something while thinking that one has no good reason to

believe it, or even that the evidence suggests the contrary, is more puzzling (p.107). Cassam nevertheless thinks that this is possible, and so believing the reasons indicate that P does not necessarily mean that one believes that P. Readers should be interested to see his case for this (pp.22-23).

If this phenomenon is possible however, then presumably it's rare enough, so Cassam doesn't deny that TM can be a route to self-knowledge, for belief at any rate, if believing that P *normally* goes together with believing that the reasons indicate that P/that one should believe that P. His point is that it's only one way among many, since there can be other ways of knowing what we believe. Furthermore, it doesn't have the special status Richard Moran accords it in giving us *immediate* self-knowledge of one's beliefs (p.111). For Cassam understands TM as follows. We judge that the evidence/reasons indicate that P. Moreover, we know we judged this (incidentally, a piece of self-knowledge TM takes for granted (p.118)). We then *infer* from this that we believe that P, guided by the 'Rationality Assumption': that what you believe is generally in line with what you think the reasons require you to believe.

I guess Moran would resist this interpretation of TM, but another one which Cassam mentions seems more in the spirit of his thinking. Assuming that judging that the reasons indicate that P amounts to judging that P (p.112 & p.116; Cassam grants this for argument's sake), then comes 'the crucial Rationalist move: the relationship between judging that P and believing that P isn't evidential; rather, your judging that P *constitutes* your believing that P' (p.112). This way no inference need take place (if judging and believing are categorically distinct like Cassam says, one being a mental action and the other a state, perhaps the view should be that judging constitutes the *onset* of believing, or something similar). He spells out the Rationalist's reasons for this 'constitutive view', but the move ultimately fails, because examples show that judging and believing can 'come apart'. In Cassam's view, Christopher Peacocke illustrated this with his example of a professor who judges that foreign undergraduate degrees are just as good as her country's degrees, but where it becomes clear, from decisions she makes on hiring and from her recommendations, that she doesn't really believe this (p.117).

Does this case show such a coming apart? Cassam's view of belief might suggest otherwise. This view is 'broadly dispositional. Whether you actually 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' (p.118). Now the professor in Peacocke's case apparently satisfies *one* stated criterion for believing that foreign degrees are as good as her country's: she thought/judged that this is so when the question arose. Thus we can't say *without*

*reservation* that she doesn't believe this. But neither can we say, without reservation, that she does believe this, since she fails to act as if this is true (Cassam's second criterion). What this case might show, then, is a borderline case of belief (though it may be too underdescribed for definite conclusions; would she make that judgement any time the question arose, for instance?). If Cassam partly *defines* believing that P in terms of a disposition to judge/think that P, he may have a hard time demonstrating a clean dissociation between judging and believing.

Though there is room for dispute on certain points, Cassam raises important challenges for transparency views of self-knowledge, many of which I have not mentioned here.

**Constructive Section:** The remaining chapters (11 to 15) are mainly constructive, where Cassam develops his inferentialist theory of self-knowledge, of standing attitudes in particular. Inferentialism is the view that inference is a 'key source' of such knowledge (p.137). Put that cautiously, it doesn't sound like the 'deeply unpopular view' (p.viii) he says it is, but what he seems to mean is that we *normally* gain self-knowledge of our beliefs, desires, etc., from inference (p.144), though in less restrained moments he suggests that, arguably, we *always* gain such knowledge inferentially (pp.144-145). Thus his scepticism towards the idea of immediate self-knowledge has grown since previous writings, where he was willing to say that some self-knowledge of propositional attitudes is immediate (Cassam 2010: 565).

A strength of Cassam's inferentialism is its emphasis on the variety of evidential sources relevant to inferring our mental states. These include overt behaviour, but also 'internal promptings' such as passing thoughts and feelings, imaginings, and fantasies, and also somatic phenomena (e.g., increased heart-rate indicating fear (pp.205-206)). Cassam's main illustration is the case (taken from Krista Lawlor) of Katherine, who discovers/infers that she wants another child on the basis of feeling envy when an acquaintance reveals her pregnancy, and her imaginings and fantasies, among other things (p.143).

As Cassam shows, inferentialism provides a useful framework for explaining how self-deception about our own minds and self-ignorance can occur, phenomena which seem real (chap.14). He outlines ways in which such inferences can go wrong, and investigates obstacles to self-knowledge. He can also agree that we can be in a privileged position to know our own minds, since we have a 'special access' to our internal promptings that others lack (p.150).

One might have reservations about Cassam speaking of ‘evidence’ here. We often think of evidence as something distinct from what it’s evidence of (e.g., footprints, and someone having passed by). However, the sorts of evidential base which in Cassam’s view we use to infer our desires, emotions etc., often don’t seem distinct from or contingently related to these phenomena. Missing her when she’s gone (p.180), feeling excited about seeing her again, feeling envy when someone else wins her interest, waking up daily with the thought of her, may all be a good basis for Romeo to ‘infer’ that he loves Juliet. But these things are *part of what it is* to be in love with someone (for what are we to say love is, if it’s distinct from these manifestations?). Ditto with many of the internal promptings Cassam mentions. This should not be a problem for inferentialism however, so long as we enlarge our notion of evidence to encompass cases of inferring something from that which *partly constitutes* it (a case of inferring a whole from a part). Moreover, this might explain why self-knowledge is sometimes inferential and sometimes immediate: sometimes we become aware of a part and must infer the whole, whereas for simpler objects of self-knowledge (e.g., sensations) we can be directly aware of the whole.

The case of Katherine shows an example of substantial self-knowledge, and Cassam makes a strong case that substantial self-knowledge is normally inferential, developing an inferentialist theory more sophisticated than the view that such inferences are always made from behavioural evidence, and outlining important differences between types of substantial self-knowledge (chap.13). This attention given to substantial self-knowledge is most welcome. But what about trivial self-knowledge? Cassam acknowledges that in many such cases it doesn’t *appear* that our self-knowledge is inferential. It seems like we just know ‘immediately’ what we believe, desire etc. But the inferentialist may ‘maintain that we know [such attitudes] by unconscious inference’ (p.145). This, however, is an *ad hoc* hypothesis if made merely to render the appearances consistent with inferentialism, without supplying independent evidence for such unconscious inferences. Cassam also suggests that your knowledge that P is inferential ‘if your justification for believing that P comes in part from your having justification to believe other, supporting propositions’ (p.139). To make good on the claim that trivial self-knowledge of attitudes is normally inferential, then, one should be able to take normal cases and specify what these supporting propositions are. What, for instance, could the supporting proposition be which I know is true, from which I unconsciously infer that I believe I’m wearing socks? The book is short on details in this respect.

Cassam pushes his view further still, arguing that self-knowledge is also inferential and ‘interpretive’ on the level of the ‘internal promptings’, the occurrent mental events from which we infer our standing attitudes. Katherine, for instance, infers that she wants another child partly from noticing that she feels a yearning for another child (p.163; one might question how distinct yearning is from wanting, but let’s put that aside). Cassam then suggests that the fact that she feels a yearning for another child is ‘to some extent’ also inferred, from her ‘background knowledge’, or knowledge of ‘contextual factors’ (p.163). The idea seems to be this: Katherine feels something with the character of a yearning, and from the fact that ‘the question whether to have another child has been on her mind recently’ (p.166), she infers that it is a yearning for another child.

I was not convinced we use such background knowledge to identify our occurrent feelings. First, such background evidence can underdetermine multiple ‘interpretations’. Suppose Katherine’s husband has been pressuring her for another child. In that case, the possibilities that her yearning is for another child, or is for her husband to stop pressuring her, might both be compatible with her background evidence: that this issue was on her mind. Yet she might still have no difficulty knowing what her yearning is for. Second, sometimes feelings enter our minds which are unrelated, or only loosely related, to the recent context. Suppose that while thinking about having another child, Katherine gets a sudden yearning for some tea. Yet she would hardly mistake it for the yearning for another child just because she had been preoccupied with that topic.

I have not had space to summarize other aspects of the book, such as the rich discussion of the value of self-knowledge, the powerful critique of the inner perception theory, various other challenges to Rationalism, and numerous sharp observations made along the way. On the whole, it is an able, readable defence of an important and relatively fresh perspective on self-knowledge, though some might argue he pushes the idea too far. I look forward to the debate it will stimulate.

## **References**

Cassam, Q. (2010) ‘How We Know What We Think’, *Revue de Métaphysique et de Morale*, 68, 553-569.