

Critical notice of Quassim Cassam, *Self-Knowledge for Humans*, to appear in *Mind*.

Brie Gertler

Penultimate draft. For quotes, please consult the published version.

With this provocative book, Quassim Cassam aspires to reorient the philosophical study of self-knowledge so as to bring its methodology and subject matter into line with recognizably human concerns. He pursues this reorientation on two fronts. He proposes replacing what he sees as the field's standard subject, an ideally rational being he calls *Homo Philosophicus*, with a more realistic *Homo Sapiens*. And he proposes shifting the field's primary focus from 'narrow epistemological concerns' to issues reflecting 'what matters to humans', such as knowledge of one's own character and the moral significance of self-knowledge. Cassam also contributes to this field: he advances an inferentialist account of self-knowledge and a moderate instrumentalism about self-knowledge's value.

The book's signal virtue is its unwavering insistence that philosophical views about self-knowledge should be judged by their fidelity to what self-knowledge actually is: an untidy phenomenon in the lives of cognitively limited creatures. Cassam uses this realist standard to challenge extravagant claims about self-knowledge: that we have direct, infallible access to our attitudes; that our rationality guarantees a robust capacity for self-knowledge; and that self-knowledge is essential to an authentic, meaningful life. His accounts of self-knowledge and its value are models for avoiding the excesses he decries. On his view, self-knowledge is typically achieved through inference from diverse kinds of evidence, and is epistemically distinctive only in a comparatively modest way. It derives its value from the practical and moral goods it (sometimes) promotes.

Cassam's realist outlook is sensible and refreshing, and his effort to bring philosophical attention to neglected issues about self-knowledge is commendable. But I have reservations about the book's framing conceits, namely, that taking seriously how humans actually think amounts to a 'radical reorientation of the philosophy of self-knowledge' (11), and that the field's current focus on epistemological issues is indefensible. This way of framing the discussion may help to attract a wide audience. But from a scholarly perspective, the choice to structure the book as a polemic against mainstream theorizing about self-knowledge is unfortunate. It leads Cassam to focus his critical remarks on approaches to self-knowledge starkly opposed to his own, whereas critically evaluating closer competitors would yield greater philosophical payoff. This polemical structure also constrains Cassam's discussion of his own promising accounts of self-knowledge and its value, which deserve fuller development than they receive here. Still, the book is a real achievement, and makes important contributions to an impressive range of philosophical issues about self-knowledge.

I'll begin by discussing Cassam's proposal to dethrone *Homo Philosophicus* (Section 1). The bulk of my remarks will center on Cassam's case for shifting the field's focus from relatively technical epistemological questions to issues with wider appeal. I evaluate this case by closely examining his characterization of the contrast between 'trivial self-knowledge' and 'substantial self-knowledge' (Section 2), and considering his inferentialist account of substantial self-knowledge (Section 3).

1. *Homo Philosophicus* vs. *Homo Sapiens*

Of the book's fifteen chapters, six are devoted to challenging overly idealized conceptions of the rational thinker. The choice to discuss rationalism at such length is driven by a somewhat curious objective. Quoting from the book jacket: 'This book tries to do for philosophy what behavioural economics tries to do for economics', namely, to expose the myth of the ideally rational subject. (This comparison is repeated in Chapter 5 and elsewhere.) The analogy seems inapt. Rationalist assumptions are rejected by most self-knowledge theorists, and certainly lack the foundational influence in the field they (reportedly) once had in economics.

These chapters draw on the work of Daniel Kahneman, who helped to pioneer the shift to behavioural economics, to describe how our cognitive processes fall short of rational ideals. Our cognitive systems often favor economy over scrupulousness, and employ heuristics that sacrifice accuracy for efficiency. So our patterns of thought are less than optimally rational: beliefs persevere even after the original evidence for them is eliminated; we place greater weight on evidence confirming a belief than on counterevidence; etc. Cassam argues that these tendencies don't render *us* irrational. Since our cognitive and temporal resources are limited, favoring efficiency and economy is prudent.¹

Cassam cites this disparity, between our actual attitudes and those that an ideally rational being would have, in criticizing Richard Moran's rationalist transparency account of self-knowledge (Moran 2001). On that account, we can normally identify our attitudes by reflecting on our reasons, and thereby determining the attitudes we ought (rationally) to have. Cassam argues that, since we fall short of the rational ideal, our attitudes often diverge from those supported by our reasons. He acknowledges that deliberation about our reasons can *shape* our attitudes, but argues—persuasively, to my mind—that our agency, relative to our attitudes, does not explain self-knowledge.

Another source of rationalism about self-knowledge is the idea that critical reasoning requires self-knowledge (Burge 1996, Shoemaker 1994). Cassam deploys an example of Christopher Peacocke's to dispute this idea (Peacocke 1998, 277). In the example, a subject's reflection on her evidence prompts her to revise a belief, yet the subject never thinks about her belief as such. The example shows that critical reasoning—reasoning that reflects a grasp of evidential relations—need not involve self-knowledge.

This conclusion, which seems to me entirely right, contributes to a larger theme in Cassam's sustained campaign against overestimating the role of self-knowledge in ordinary life. One needn't be *aware of* one's attitudes in order for them to play the kinds of salutary roles, in the meaningful life of a rational thinker, that inspire rationalistic approaches to self-knowledge. This point reappears in Cassam's discussion of the value of self-knowledge, as we will see below.

¹ Cassam takes particular issue with those, like Dan Ariely, who conclude from such phenomena that we are irrational. This dispute seems largely verbal.

2. Self-knowledge and ‘what matters to humans’

Mainstream philosophical study of self-knowledge is driven (at least proximally) primarily by epistemic concerns. For this reason, its discussions center on instances of self-knowledge that appear to be epistemically distinctive: to attain an especially high level of certainty, or to involve a special route to knowledge. Cassam says that the field’s epistemic orientation ‘to some extent justifies’ its narrow focus on what he calls ‘trivial’ cases, such as my knowledge that I’m now in pain or that I believe that it’s raining. But he argues that philosophers interested in self-knowledge should consider questions about self-knowledge beyond the epistemic, and should turn their attention to ‘substantial’ kinds of self-knowledge. In particular, we should consider the following questions.

- Is self-knowledge practically or morally valuable? If so, what kinds of self-knowledge are valuable, and why are they valuable?
- How does one achieve knowledge of one’s own character, values, and other ‘substantial’ traits? That is, what is the epistemology of substantial self-knowledge?

Few philosophers would deny that these questions are philosophically significant. More contentious is Cassam’s assertion that the choice to focus one’s research exclusively on the epistemology of trivial self-knowledge is *indefensible*.

There is no excuse for ignoring substantial self-knowledge. (47)

If, as a philosopher and a human being, you are interested in self-knowledge then you really should be interested in substantial self-knowledge; there is no excuse for only trying to account for trivial self-knowledge and its supposed privileges. (174)

This is a striking claim. After all, one can recognize that an issue is important without pursuing it oneself. A philosopher might choose to focus her research on metaphysical issues about personhood, while regarding bioethical issues about personhood as equally—or even more—important.

An understanding of trivial self-knowledge will have significant repercussions beyond this field. (Many who work on self-knowledge were initially drawn to the field because of its connections to other subdisciplines of philosophy.) The leading arguments for dualism rest on the idea that self-knowledge of sensations differs epistemically from perceptual knowledge: this connection between self-knowledge and mental ontology has prompted an explosion of literature on phenomenal concepts. Whether and how trivial self-knowledge differs from perceptual knowledge also bears on issues about perceptual justification, the ‘veil of perception’, epistemic foundationalism, and skepticism. And many philosophers take the viability of mental content externalism to depend on disputed questions about self-knowledge.

So why does Cassam think that focusing on trivial self-knowledge is indefensible? In a nutshell, he seems to think that this narrow focus will mislead us about, or blind us to, important facts about self-knowledge. We will mistakenly generalize, from what’s true

of trivial self-knowledge, to conclusions about self-knowledge as a whole; and we will fail to grasp the value of self-knowledge (46-7).

To comprehend these worries, we must understand how Cassam distinguishes between trivial and substantial self-knowledge. His official explication of this distinction consists in a list of ten characteristics he associates with substantial self-knowledge. These are not intended as individually necessary or sufficient conditions. Rather:

The point of saying that knowledge of, say, your own character is substantial is to indicate that it has at least some of the characteristics I have in mind. The more of these characteristics it has the more substantial it is.
(30)

Cassam's argument seems to be this. Insofar as substantial self-knowledge differs from trivial self-knowledge, in the ways indicated by these ten characteristics, restricting our attention to trivial self-knowledge threatens to leave us with a distorted and incomplete picture of self-knowledge.

So what kinds of characteristics determine whether an instance of self-knowledge is substantial? All but one of the ten characteristics are epistemic, in a broad sense; the remaining characteristic concerns value. Correspondingly, there are two potential dangers of limiting oneself to epistemic questions about trivial self-knowledge, and ignoring substantial self-knowledge and questions about value. One is epistemic, and one concerns value. I'll consider each in turn.

2.1 *First danger of focusing on trivial self-knowledge*

First Danger. We will generalize, from the epistemic features of trivial self-knowledge, to conclusions about the epistemology of self-knowledge more broadly. The resulting picture will be inaccurate, since substantial self-knowledge differs epistemically from trivial self-knowledge.

The epistemic characteristics associated with substantial self-knowledge include fallibility, corrigibility, indirectness, and requiring cognitive effort. These are, of course, converses of the epistemic features historically attributed to self-knowledge. That epistemic characteristics play such a central role in distinguishing substantial from trivial self-knowledge is somewhat puzzling. For a major theme of the book is that most self-knowledge, of every sort, lacks the special epistemic features historically attributed to it. Perhaps the most surprising epistemic characteristic on the list is this: *cannot generally be achieved by the rationalist transparency method*. But Cassam's arguments against that method have wide application. They suggest that even trivial self-knowledge—such as knowing that one wants a vodka martini (104)—cannot generally be achieved by use of that method. So the very success of those arguments implies that this characteristic will not help to distinguish substantial from trivial self-knowledge.

This puzzle arises for the other epistemic characteristics as well. Cassam argues that trivial self-knowledge typically lacks the infallibility, incorrigibility (etc.) historically attributed to it. For example, 'there is no immunity to error even when it comes to self-ascriptions of trivial attitudes' (44): I can wrongly think that I want chocolate ice cream only to realize, when my order arrives, that I really wanted vanilla. That case suggests that

knowledge of my flavor preferences, an example of trivial self-knowledge that recurs throughout the book, is not only fallible but also has most or all of the other epistemic characteristics associated with substantial self-knowledge. Cassam extends this skepticism even to judgments about sensations: one who has difficulty classifying a sensation as *pain* may be aided by considering the sensation's cause (164). This implies that even self-knowledge of a sensation may have (most of) the epistemic characteristics on the list: it may be fallible, corrigible, inferential, and based on evidence, and it may require cognitive effort.² (Insofar as these characteristics are intended to *define* substantial self-knowledge, as the quote from page 30 above indicates, the proper conclusion may be that self-knowledge of pain can be substantial. I consider this possibility below.)

Cassam's arguments for inferentialism about self-knowledge emphasize that (nearly) all self-knowledge exhibits another epistemic characteristic on the list, indirectness. In response to the argument that we seem sometimes to know our attitudes directly, without inference, he says that such cases likely involve 'unconscious inference' (144). More generally, he thinks we should favor a single epistemology of an attitude type such as desire (145), regardless of the attitude's content—and so regardless of whether knowledge of it qualifies as trivial or substantial. On his view, even access to (many or most) occurrent thoughts and feelings is inferential (165).

According to Cassam's arguments, then, nearly all self-knowledge lacks the impressive epistemic features historically attributed to self-knowledge, and trivial and substantial self-knowledge share a basic epistemology. These conclusions appear to minimize the detrimental effects of using instances of trivial self-knowledge as data for an epistemic account of self-knowledge generally. They thereby appear to defuse, or at least diminish, the first danger of focusing on trivial self-knowledge.

Perhaps Cassam regards the danger of an exclusive focus on the trivial as follows. Trivial self-knowledge *appears* infallible, direct (etc), and this may mislead us into thinking that it actually has those epistemic features; if we use the trivial as a model for self-knowledge generally, we will mistakenly conclude that substantial self-knowledge also has them.³ However, this interpretation is difficult to square with Cassam's observation that philosophers focus on mundane cases precisely because they think that these cases have special epistemic features, features that less mundane kinds of self-knowledge may lack.

There is one epistemic characteristic that, compatibly with Cassam's other arguments, may distinguish substantial from trivial self-knowledge: substantial self-

² Similar worries apply to another epistemic characteristic, namely that substantial self-knowledge is sometimes impeded by familiar obstacles, such as self-deception or bias. In criticizing the tendency to contrive the thinking subject as an ideally rational *Homo Philosophicus*, Cassam marshalled psychological evidence showing that our thought about even mundane matters is often shaped by factors, like confirmation bias, that stem from heuristics that value efficiency over accuracy. On the standard interpretation, the fraternity case is precisely a case of expectation bias: the blind-folded pledge who expects that his hands will be burned with cigarettes believes he's experiencing a sharp pain; in fact, he is experiencing a cold sensation, since his hands are touched with ice.

³ I'm grateful to Eric Schwitzgebel for suggesting this interpretation.

knowledge is more likely to be impeded by obstacles involving one's self-conception. It is easy to imagine such obstacles hindering self-knowledge of character or deep-seated attitudes, such as knowledge that one is avaricious or racist. Might they also thwart trivial self-knowledge? I suppose an emotional investment in seeing oneself as a devotee of chocolate could blind someone to the fact that he prefers vanilla. Perhaps Cassam would say that in this case, coming to recognize that one prefers vanilla *would* be substantial self-knowledge.

This brings us to the possibility mentioned earlier. Perhaps the upshot of Cassam's arguments that even mundane self-knowledge is fallible, indirect (etc), is that such self-knowledge is substantial rather than trivial. This reading, which takes seriously the idea that these (mostly epistemic) characteristics *define* the trivial-substantial contrast, draws support from the following passage.

Whereas trivial self-knowledge can seem interesting because it is so *easy* to get, what is striking about substantial self-knowledge is that it can be so *hard* to get. It is the *elusiveness* of substantial self-knowledge, and the resultant threat of self-ignorance, which accounts for some (but not all) of its philosophical interest. (47)

But in that case, the best strategy for a philosopher whose interest in self-knowledge centers on its epistemic distinctiveness—the fact that some kinds of self-knowledge are 'so *easy* to get'—is precisely to focus exclusively on trivial self-knowledge.⁴

Let me sum up the discussion of this first danger. That the trivial-substantial contrast is drawn mainly along epistemic lines is difficult to reconcile with the book's larger emphasis on the idea that trivial and substantial self-knowledge have similar epistemic features and share a single basic epistemology.⁵ If these are epistemically similar, the danger of using trivial self-knowledge as a basis for a general epistemology of self-knowledge is limited. If, on the other hand, the trivial-substantial contrast is largely defined by epistemic differences, it is eminently reasonable for a philosopher moved by epistemic concerns to focus on trivial self-knowledge.

2.2 *Second danger of focusing on trivial self-knowledge*

Here is the second danger of focusing exclusively on the epistemology of trivial self-knowledge.

Second danger. The resulting accounts of self-knowledge will neglect or misrepresent its value, because substantial self-knowledge is valuable in a way (or to an extent) that trivial self-knowledge is not.

⁴ Just to be clear: Cassam plainly thinks that some self-knowledge is trivial self-knowledge.

⁵ I'm not claiming that Cassam's arguments don't allow for *any* epistemic distinction between trivial and substantial self-knowledge. But given that he takes pains to underscore the epistemic similarities between these, it is puzzling that epistemic characteristics play such a central role in *defining* substantial self-knowledge.

This worry stems from the final characteristic Cassam associates with substantial self-knowledge, the ‘Value Condition’:

[S]ubstantial self-knowledge *matters* in a practical or even a moral sense. ...
[N]ot knowing what will make you happy can result in your making bad choices, and we think of some forms of self-ignorance not just as cognitive but also as moral defects. Being unkind is bad in itself but made morally worse if it is combined with the belief that one is kind. (31-32)

His main examples of substantial self-knowledge are knowledge of one’s own character, values, and substantial attitudes—where substantial attitudes include the desire for a child and racist beliefs. Is such self-knowledge practically or morally significant, in the sense expressed in the Value Condition? He addresses this question in the concluding chapter of the book, where he again casts a skeptical eye on lofty claims about self-knowledge.

Cassam’s target here is the idea that self-knowledge is intrinsically valuable because it is indispensable for an authentic or unified life. He allows that authenticity and unity may be valuable, and that self-knowledge has derivative value insofar as it facilitates these. But he contends that self-knowledge is not strictly required for authenticity and unity. I can be authentic, and live a life that expresses my character and values, without being guided by *knowledge of* my character or values: for instance, generosity (as a character trait and/or a value) may guide my actions and decisions without my thinking of myself *as* generous, or *as* valuing generosity. This point nicely complements his earlier grounds for denying that critical reasoning requires awareness of one’s attitudes, and strengthens his larger case against inflated conceptions of the significance of higher-order reflection.

Cassam allows that self-knowledge is sometimes valuable. (Though he cites psychological research suggesting that moderate self-*ignorance*—usually, a slightly more positive self-conception than what is warranted—can promote well-being.) When self-knowledge is valuable, he thinks, its value is instrumental. ‘Self-knowledge derives whatever value it has from the value of what it makes possible’ (227). For instance, realizing that I’m timid has instrumental value if this realization prompts me to take an assertiveness course. This moderate position about self-knowledge’s value is clearly in keeping with Cassam’s larger outlook. I find it sensible and extremely plausible.

Let’s consider what this moderate position means for the Value Condition. If even substantial self-knowledge has only instrumental value, why think that ‘Being unkind is bad in itself but made morally worse if it is combined with the belief that one is kind’ (32)? Perhaps the unkind person’s mistaken belief is morally problematic because it decreases the likelihood that she will strive to become kinder. Of course, it decreases that likelihood only on the assumption that she *values* kindness. And surely knowing that one is unkind, and celebrating that fact, is at least as morally problematic as mistakenly believing oneself to be kind (while valuing kindness). Realizing that one is kind, when one values unkindness, may have negative moral consequences, as it may lead one to strive to become less kind. As Cassam observes, ‘We clearly don’t want people like Stalin and Hitler to be true to themselves’ (223).

The moral value of self-knowledge's consequences will thus depend on the knower's own motivations or values. This is arguably a feature of knowledge generally. Since knowledge typically facilitates the achievement of one's goals, the instrumental moral value of any kind of knowledge will depend upon the knower's motivations or values.⁶

So is substantial self-knowledge generally more valuable than trivial self-knowledge, as the Value Condition says?

Trivial self-knowledge has obvious practical value: e.g., it enables me to order the proper dessert. (Here, I'm assuming that knowledge of sensations and mundane attitudes qualify as trivial self-knowledge.) And it is sometimes morally valuable. In a visit to the emergency room, knowledge that I'm feeling pain rather than nausea could save my life. If I'm a subject in a research study investigating side effects of a new medication, that knowledge may carry morally significant benefits for others.

But the morally significant consequences of trivial self-knowledge are probably sparse in comparison to those of substantial self-knowledge. (Insofar as the trivial-substantial contrast is defined partly in terms of value, substantial self-knowledge generally has greater value than trivial self-knowledge by definition.) Character traits and values involve stable dispositions, likely to shape morally significant behavior. By contrast, much trivial self-knowledge is knowledge of states that are short-lived: fleeting sensations, passing thoughts, momentary urges. And trivial self-knowledge of more stable states, such as banal geographical beliefs, is perhaps unlikely to shape morally significant behavior. Whether the moral value of substantial self-knowledge is positive or negative will largely depend on the subject's own motivations and values.⁷

Granting, then, that substantial self-knowledge has greater value—especially greater instrumental moral value—than trivial self-knowledge, does this establish that 'there is no excuse' for focusing one's research on trivial self-knowledge? I don't think it does. For as even this brief discussion about the moral significance of self-knowledge demonstrates, in addressing the value of self-knowledge we move quickly into the ethical realm. My point is not that subdisciplinary boundaries should be strictly enforced. Rather, it's that the moral significance of self-knowledge essentially depends on ethical issues—the value of authenticity, whether normative beliefs are intrinsically motivating, and even the truth of consequentialism. It is natural that those self-knowledge theorists who have limited expertise in ethics and moral psychology (including me) would not take up these issues. Of course, interest in the value of self-knowledge might inspire a philosopher to educate herself in the relevant fields. But ceding this topic to those with expertise in ethics is one legitimate way to recognize its importance.

This understanding of the dialectical situation defuses the second danger. Entrusting questions about self-knowledge's value to those well-versed in ethics and

⁶ For anti-Humeans about moral motivation, knowledge of moral truths may be an exception here.

⁷ I include the qualification 'largely' to accommodate a point of Cassam's: that self-knowledge may contribute to authenticity and a unified life, which may be valuable. This point necessitates the qualification since authenticity and unity may be valuable independent of one's motivations or values.

moral psychology is one way of recognizing the importance of those questions. As long as self-knowledge theorists are clear about their projects, the focus on trivial self-knowledge will not have the effect of downplaying the importance of ethical issues about self-knowledge or misrepresenting self-knowledge's value.

3. Cassam's inferentialist account of self-knowledge

Cassam's account of self-knowledge is mainly concerned with substantial self-knowledge, but he has interesting things to say about some kinds of trivial self-knowledge as well. He argues that most self-knowledge of any stripe is inferential, where he takes this to mean that it is 'based on evidence'.⁸

His account of substantial self-knowledge of standing attitudes draws on work by Krista Lawlor and by Peter Carruthers. Lawlor argues that such knowledge sometimes rests on 'inference from internal promptings' (Lawlor 2009). For example, one might infer that she desires another child from feeling envy upon learning that an acquaintance is pregnant, and finding herself fondly remembering the feel of a newborn in her arms. Cassam maintains that this kind of inference from evidence is 'the normal way' to know one's standing attitudes.⁹

The resulting knowledge is doubly inferential, in that access to these internal promptings is itself inferential. To identify what I'm feeling as *envy*, I must exploit my knowledge of contextual factors: what topics have recently occupied me, and perhaps my behavior. (This picture of interpretive access is partly drawn from Carruthers; see Carruthers 2011.)

Self-knowledge of character traits is similarly inferential.¹⁰ In one example, Woody comes to realize that he is fastidious by the following route: when imagining various untidy scenes, he feels dismay and irritation. These internal promptings presumably require interpretation, so the conclusion that what he's feeling is *dismay at a mess* is inferential. 'On the basis of his thoughts, imaginings, and emotions Woody is in a position to conclude that he cares about cleanliness and attention to detail' (177). But that conclusion does not suffice to establish that he is fastidious, according to Cassam:

⁸ Cassam largely ignores epistemically externalist views about self-knowledge. Most notably, he does not address the externalist versions of the transparency account (advanced by Byrne and Fernández), although he mentions them in a footnote. This is notable because these versions, which are not rationalist, are immune from the objections he presses against the rationalist version of transparency. He addresses externalism primarily as part of his criticism of the inner sense view: and there, his main objection to externalism is that it is incompatible with the larger analogy to perception driving the inner sense view (134-36).

⁹ Lawlor presents her view, not as a general account of how we normally know our attitudes, but as a counter-example to various rationalist claims, e.g. that self-knowledge achieved through inference from evidence is 'alienated'.

¹⁰ Cassam describes self-knowledge of character as 'doubly inferential'; however, the Woody example seems to involve three inferential steps.

‘he might wonder whether he cares enough about tidiness and attention to detail to make him fastidious.’ (177). So the step from knowing that he *cares about tidiness* to knowing that he is fastidious is itself inferential.

There is psychological evidence that this sort of process sometimes works. Studies by Schultheiss and Brunstein (1999) suggest that ‘people can detect their nonconscious dispositions and motives by vividly imagining a future situation and attending to how it would make them feel’ (Wilson and Dunn 2004).

When we succeed in knowing our substantial attitudes and character traits, on the basis of evidence, such knowledge is indeed ‘hard-earned’ (177). And it may be quite rare. As psychologists are fond of telling us, self-assessments of character traits are prone to well-documented biases. When it comes to predicting actions stemming from deep-seated, unconscious tendencies and motivations—as opposed to those controlled by conscious deliberation—spouses and close friends are typically more accurate than the subject herself.¹¹ This is because, as Cassam observes, emotional investments in one’s self-conception sometimes impede clear-eyed appraisal of one’s own character or attitudes. But there is also reason to think that the ‘evidence’ provided by conscious reflection on one’s own attitudes or motives is not a reliable guide to deep-seated character traits, which are largely unconscious. (For useful overviews, see Wilson 2002, Wilson and Dunn 2004.) While internal promptings can shed light on attitudes and emotions, self-understanding of more deeply-seated drives and proclivities is often improved by *ignoring* such evidence, in favor of evidence about one’s own past behavior and, even, information about one’s peers (Gilbert et. al. 2009). Cassam’s view accommodates this point, by including behavioral evidence in the inferential basis for substantial self-knowledge. Yet as he also notes, problems loom even there: the self-attribution of attitudes on the basis of behavior is subject to well-established biases. Most famously, we tend to confabulate the reasons for our actions (Nisbett and Wilson 1977); and we tend to ignore the influence of external factors, which often shape our behavior more directly than stable traits do (Bem 1972).

In discussing Cassam’s claims about substantial self-knowledge—how it is achieved, and our general accuracy—I’ve invoked empirical psychological studies, just as Cassam himself does. This is for good reason, as such issues cannot be resolved by introspection, conceptual reflection, and thought experiments. The research methods of social psychology are best suited to answering many of the questions Cassam urges us to consider. These include: How reliable are our judgments about our own character traits? What are the special obstacles to substantial self-knowledge? Are we generally good at affective forecasting, that is, ‘knowing what will make you happy?’ (32). The choice to leave such questions to those with expertise in psychological research methods—experimental design, statistical analysis, etc.—seems eminently reasonable. To echo my earlier point about the value of self-knowledge: entrusting the investigation of these questions to those with the relevant expertise is one way of registering their importance.

¹¹ The Schultheiss and Brunstein studies mentioned targeted subjects’ attitudes about playing a competitive video game. These attitudes are probably less emotionally loaded than, say, attitudes about having a child.

Conclusion

The topic of self-knowledge crosses both intra-disciplinary boundaries (such as that dividing epistemology from ethics) and inter-disciplinary boundaries (such as that dividing philosophy from psychology). Perhaps we should rethink these boundaries. Still, one who is interested in self-knowledge can reasonably choose questions about self-knowledge that suit her specific interests, and can reasonably employ the methods most appropriate to answering these questions. An epistemologist may legitimately focus her work on the epistemic dimensions of self-knowledge, and an ethicist may legitimately focus his work on its moral dimensions. Similarly, so long as researchers conscientiously attempt to match their methodology to the questions that interest them (or, as the case may be, to match their questions to their preferred methodology), it is entirely legitimate for some to conduct research about self-knowledge by carrying out empirical psychological studies; for others to conduct it from the armchair; for others to illuminate self-knowledge through other means, such as writing poetry; and for still others to employ some combination of these methods.

My reservations about Cassam's criticisms of mainstream theorizing notwithstanding, I highly recommend this book. Its powerful critical arguments and refreshingly level-headed approach to a wide range of questions about self-knowledge make it a must-read for specialists. Because the writing is accessible and highly engaging, philosophers in other fields and non-philosophers will profit from it as well.