



## ARTICLE

DOI: 10.1057/s41599-017-0049-5

OPEN

# Overconfidence, self-knowledge, and self-improvement

Eylem Özaltun<sup>1</sup>

**ABSTRACT** Recently physician overconfidence has been considered as a major factor contributing to diagnostic error. A philosophical inquiry into overconfidence as a character vice has the promise of shedding light on how we can overcome this vice and potentially reduce diagnostic errors. In his recent work, Quassim Cassam conducts such an inquiry. This paper puts Cassam's work on physician overconfidence in the context of his theoretical work on self-knowledge and epistemic vices. It shows that physician overconfidence, considered as a major factor in diagnostic error, provides a significant real-life application of Cassam's accounts of self-knowledge and epistemic vices. It focuses on the features of these accounts that lead to the following result: self-knowledge is rarely, if ever, a remedy for physicians' overconfidence and the resulting diagnostic errors. By appealing to the same data Cassam cites regarding diagnostic error and physicians' overconfidence, it is shown that the more substantial—in the sense Cassam specifies—one's third-personal knowledge of oneself is, the less likely it is to be of any practical value qua *self-knowledge*. This paper defends the view that what Cassam calls 'trivial self-knowledge'—first-personal knowledge that has been the primary concern for philosophers—is crucial for any kind of self-knowledge to be instrumental for self-improvement. Since an agent acts from the standpoint that she is aware of herself trivially, it is argued that what Cassam calls 'substantial self-knowledge' has no practical value unless it is integrated with what he calls 'trivial self-knowledge'. In this way the paper explains why if what one learns about oneself from the third-person perspective is drastically different from what one takes to be true from the first-person perspective, one cannot act on this knowledge. Since the standpoint from which one experiences and acts intentionally are one and the same, the paper also explains why traumatic experiences (such as the death of a patient or of a loved one) can sometimes lead to fundamental change and self-improvement.

<sup>1</sup>Department of Philosophy, Koç University, Istanbul, Turkey. Correspondence and requests for materials should be addressed to E.Öz. (email: [eoaltun@ku.edu.tr](mailto:eoaltun@ku.edu.tr))

## Introduction

Philosophers do not care about particular facts a particular person knows about herself. Instead, they study the ways in which people know about themselves. One way is by collecting evidence about the object which happens to be oneself. Each person is a public object, hence can observe oneself just like she can observe other objects around her. She can look and see there is a blank page in front of her, and she can look in the mirror and see that she has more gray hair now than she had five years ago.

But there is another way in which a person can know about herself and only herself. Some facts are known not because one has evidence for them, but solely by virtue of these facts being true of oneself. Paradigmatic examples of the things one can know about oneself in this way are the facts about one's own mental life, or at least conscious aspects of one's mental life. For example, being in pain and knowing that oneself is in pain do not come apart. Although my having gray hair is true of me regardless of my knowing this fact, I cannot be in pain and not know it. This way of knowing oneself is not confined to knowledge of sensations or even to mental life. The deliberate thoughts, conscious beliefs and desires, but also, arguably, intentional actions have this peculiar feature: the agents who have these thoughts, beliefs, desires, or perform these actions necessarily know that they have these attitudes or perform these actions just by virtue of having these attitudes or performing these actions. Philosophical work on self-knowledge predominantly focuses on this second way of knowing oneself—that is, knowing by being a self-conscious being, or as it is sometimes called, “first-person knowledge of oneself”.

In a recent book (2014), Quassim Cassam criticizes the philosophical tradition that neglects the first way of knowing about ourselves, namely the type of self-knowledge which requires epistemic work, the type of knowledge we have about ourselves on the basis of evidence. His charge is that the philosophers who concern themselves solely with first-person knowledge do not pay due attention to the practical role that self-knowledge by evidence plays in the lives of ordinary human beings in their everyday behavior. As a result, most philosophical work on self-knowledge, he claims, has little or no moral or practical impact. This criticism is reflected in the names he uses to refer to these two types of knowledge one can have about oneself. He calls the knowledge which is of interest to philosophers—first-person knowledge of oneself—‘trivial self-knowledge’, and the knowledge of oneself by evidence—the type of knowledge which requires epistemic work—, ‘substantial self-knowledge’. I will give a more detailed analysis of Cassam's distinction in the following (See Sections ‘Knowledge of one's overconfidence as self-knowledge for humans’ and ‘Knowledge of one's overconfidence as substantial self-knowledge’), but first let me note what drives him in choosing these particular names to mark the distinction. In his book, he motivates the study of substantial self-knowledge by giving a *low road* account that shows its role in tackling the practical problems faced by ordinary people. Self-knowledge is valuable in so far as it is instrumental in making “a positive difference to one's overall happiness and well-being” (Cassam, 2014, p 223).

In this paper, I will critically examine a case of substantial self-knowledge Cassam himself discusses in this collection (Cassam, 2017), namely the knowledge of one's own overconfidence. Cassam's analysis of overconfidence shows that substantial self-knowledge by itself does not bring about substantial self-change. I will investigate why this is so and conclude that substantial self-knowledge *qua self-knowledge* cannot effect self-change unless it is somehow incorporated into a person's trivial self-knowledge. Since a positive difference to one's overall happiness and well-being requires substantial self-improvement, I will also conclude

that that substantial self-knowledge *qua self-knowledge* has no practical value. This will also serve as a low road defense of the value of trivial self-knowledge.

## Knowledge of one's overconfidence as self-knowledge for humans

Cassam's substantial self-knowledge is self-knowledge for *humans* in the three senses of ‘human’ Cassam had in mind when he gave his book the title ‘Self-knowledge for Humans’. That is, humans are:

1. Creatures prone to epistemic vices especially when it comes to knowing themselves
2. Creatures who are less than perfectly rational
3. Creatures who face practical and moral problems in their everyday lives, not just when they are writing philosophy papers.

Consider an overconfident physician. This physician is human in all three of these respects. First of all, she has an epistemically vicious character trait. Overconfidence is a character trait which impedes the acquisition, retention and transfer of knowledge (Cassam, 2017). Having this character trait, the physician must be less than perfectly rational at least in matters require self-evaluation. Moreover, she is someone who faces the problem of diagnosing her patients, which is undeniably a practical problem of vital importance. It has been conjectured that physician overconfidence “is a major factor contributing to diagnostic error” (Berner and Graber, 2008, p S6). In that case, the fact that she is overconfident is relevant to the practical problems she deals with everyday, as it may lead to diagnostic errors. If this physician could discover that she is overconfident, that knowledge could potentially have many moral and practical consequences. Hence, that knowledge would be *substantial* in the sense of its first definition in the English dictionary: “of considerable importance, size and worth.”

## Knowledge of one's overconfidence as substantial self-knowledge

Cassam subjects the notion of substantial self-knowledge to conceptual scrutiny in order to determine what makes substantial self-knowledge distinct from trivial self-knowledge, and what makes it substantial. He comes up with the following (incomplete) list of characteristics (Cassam, 2014, pp 30–31):

- I. The fallibility condition
- II. The obstacle condition
- III. The self-conception condition
- IV. The challenge condition
- V. The corrigibility condition
- VI. The non-transparency condition
- VII. The evidence condition
- VIII. The cognitive effort condition
- IX. The indirectness condition
- X. The value condition

Not all of these characteristics must be satisfied for a piece of knowledge to be substantial, nor does Cassam give these conditions as necessary or sufficient conditions. But he claims that the more the knowledge in question satisfies these conditions, the more substantial it is.

I will not go into the details of each condition. Instead, I will take them in three groups. Group A consists of conditions I, IV, and V, and this group of conditions, when satisfied, ensures that the knowledge in question is genuine knowledge—that is, a claim about which the epistemic agent can be in error. The agent can be

challenged in her claim to knowledge, and even corrected by other epistemic agents. If one takes the paradigmatic example of trivial self-knowledge to be knowing one's own pain, one might think that trivial self-knowledge does not satisfy this group of characteristics. But the things that can be known trivially form a diverse group in which no single member has a claim to be paradigmatic (Donnellan, 1963). For example, it has been claimed, albeit with some controversy, that although knowledge of one's own actions and of one's own bodily posture is trivial, it satisfies all three criteria of Group A (Anscombe, 1963; Özalton, 2016). So it is not obvious whether Group A is comprised of characteristics which would differentiate substantial from trivial self-knowledge.

Group B consists of conditions VI, VII, VIII, and IX. Group B conditions differentiate substantial self-knowledge from trivial self-knowledge with respect to the ways in which it is acquired. This knowledge is not *merely* reflective, it is not immediate, and it is based on evidence. Acquisition of substantial self-knowledge requires epistemic work: cognitive effort is required. Most of our empirical knowledge would require such work and one might fail to complete this work successfully. Hence, if Group B characteristics are satisfied, then so are Group A.

Group C consists of conditions II, III, and X, and these conditions distinguish substantial self-knowledge from trivial self-knowledge with respect to its content. Group C is peculiar in that in a regular epistemology course these conditions would hardly be mentioned: when we do epistemology, we abstract from the particular contents of our knowledge states. So it is worth stating these conditions explicitly (Cassam, 2014, pp 30–31):

II. Obstacle condition: ...for humans there are familiar and reasonably well understood obstacles to the acquisition of substantial self-knowledge. Such obstacles include repression, self-deception, bias, and embarrassment. Some of us find it hard to be honest with ourselves about our own limitations.

III. The self-conception condition: the existence of such obstacles to substantial self-knowledge is related to the fact that this kind of self-knowledge often 'tangles' with a person's self-conception. To know that you have a particular character you have to believe you have that character, it might be hard for you to believe that if your having that character is at odds with your self-conception.

X. The value condition: substantial self-knowledge matters to us in a practical or even a moral sense...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.

In satisfying Group C conditions, knowledge of *some* facts about oneself, such as one's character traits, values, aptitudes, emotions, abilities, or social standing, is more substantial than, for instance, knowledge of the color of one's socks. Unless she is a self-conscious teenager whose coolness is on the line if her socks are not in the fashionable color, the color of a person's socks is hardly a part of her self-conception as a morally accountable practical agent. Whether her socks are in a particular color makes no difference to what sorts of goals she will set for herself or how she will pursue them. Whereas if she comes to know that she is an influential member of her community, this would put some pressure on her to avoid slipping into derogatory or inflammatory language, for example. Now she needs to evaluate the consequences of her utterances in light of her knowledge of her influence.

This relevance to one's practical dealings is not the only reason why knowledge relevant to one's self-conception is more substantial than the knowledge of some other facts about oneself. After all, there are plenty of facts about *other* selves that would be relevant for setting my goals and how I carry them out. If I know that you are hungry, I will eat less and keep some of the lasagna I made for you. If I know that you are an overconfident physician, I will seek a second opinion on your diagnosis. In addition to practical relevance, the substantiality of knowledge relevant to one's self-conception has another source. There is a special sort of hardship in updating one's self-conception in the light of new evidence; there are obstacles which are specific to the acquisition of knowledge relevant to one's self-conception. And this is not because the facts relevant to one's self-conception are harder to uncover. Evidence of one's own character is not harder to come by than evidence of someone else's. And yet it is not uncommon that one fails to believe some obvious fact about oneself when the fact in question is at odds with one's self-conception. For such facts, in addition to the epistemic work of uncovering the truth, which is characterized by Group B conditions, there is an extra step of accepting the truth. Knowledge relevant to one's self-conception<sup>1</sup> is peculiar in this respect. Normally, when we are presented with evidence of some proposition, for instance that *my socks are green*, we are settled with the belief that *my socks are green*. And this settling is not up to us: having the evidence for *p*, other things being equal, is to believe that *p*. There is no gap between having the evidence for *p* and believing that *p*. Or so it must be for a rational being. But when *p* is relevant to one's self-conception, there is further resistance to overcome if one is to acquire knowledge. Hence, such knowledge is not only substantial in its practical impact, but also in its being a substantial cognitive achievement in the face of the adverse epistemic conditions peculiar to its content. The harder is it to accept the truth, the more substantial the cognitive achievement when one succeeds in doing so.

We have seen in the previous section that our very human physician's knowledge of her own overconfidence is substantial self-knowledge. Now in light of this discussion of Cassam's list, we see that her knowledge is *very* substantial, as it satisfies *all* the characteristics for substantiality in the list.

### Overconfidence as an epistemic vice

For an ideally rational being there should not be a gap between evidence and belief regardless what the belief is about. That there is such a gap for some self-knowledge is crucial for Cassam's recent work, "Vice Epistemology" (2016). In order to study substantial self-knowledge, we need a theory of knowledge for less than ideally rational beings. We need not only self-knowledge for humans but also an epistemology for humans, a study of epistemic vices.

Let us look more closely at how a physician's knowledge of her own overconfidence, of her "tendency to believe that she knows more than she does," (Cassam, 2017) satisfies the obstacle condition. What is in the way of acquisition of the knowledge of one's own overconfidence? One's overconfidence itself! And it is not just because no one, but especially not overconfident people, wants to think of themselves as overconfident. Any vice can be hard to own, but overconfidence is not just any vice. It is an epistemic vice, a character trait which is an impediment to effective inquiry. Being an epistemic vice, overconfidence is an obstacle to any inquiry, and in particular to self-inquiry.

But overconfidence is not just any epistemic vice, either. It is a *stealthy* epistemic vice: it especially impedes conscious critical reflection and hence specifically impedes its own detection. That is, it is stealthy because it is *self-concealing* (Cassam, 2015, p 22).

The obstacle in the way of discovering one's own overconfidence is due to this self-concealing character of overconfidence. Being overconfident significantly diminishes one's ability to detect one's own overconfidence. Hence, it is very hard to know one's overconfidence. Consequently, this knowledge is substantial self-knowledge. In general, with respect to the obstacle condition, the stealthier an epistemic vice, the more substantial self-knowledge of it would be.

Since overconfidence is self-concealing, it is unlikely, if not impossible, that our physician will come to know her overconfidence by self-reflection. But there are other ways. One way is receiving feedback about one's performance. Of course, overconfidence might block the reception of feedback as well as self-reflection. But feedback might be presented in ways that could circumvent the initial resistance. For example, suppose the administration of the hospital presents a blind performance report for a certain 'Dr. Sure' to our overconfident physician, which clearly shows that Dr. Sure's diagnostic accuracy is below the hospital average. They then ask her what she would recommend for Dr. Sure to improve his/her diagnostic accuracy. Our physician would list some extra measures to double-check the diagnoses this unidentified physician makes—presumably the measures she thinks she herself does not need. Then it might be revealed that Dr. Sure is in fact herself. Now natural expectation would be for our overconfident physician to conclude that she herself is overconfident. But as we have seen, there might be resistance not only to the acquisition of evidence but also to drawing such conclusions pertaining to oneself. But suppose that our overconfident physician reasoned as follows: Dr. Sure is overconfident, I am Dr. Sure, so *I* am overconfident. Should we expect that she will implement the strategies she recommended for Dr. Sure? Would the substantial knowledge she gained make substantial difference in her conduct?

### Substantial self-knowledge and (self)improvement

Given the way in which Cassam motivated the study of substantial self-knowledge as opposed to the trivial one, which is the concern of philosophers, we expect the answer to be 'yes'. We expect the knowledge of one's own overconfidence to satisfy the value condition and make "a positive difference to one's overall happiness and well-being" (Cassam, 2014, p 223). Cassam considers this expectation that "self-knowledge is a remedy for physician overconfidence and the diagnostic errors to which it gives rise" (Borrell-Carrió and Epstein, 2004; Croskerry et al., 2013). But he tells us that this suggestion is "underestimating the obstacles to self-knowledge while overestimating its benefits" (Cassam, 2017). And in the course of the paper in this volume, he cites empirical evidence that shows that even after being shown and accepting that they are not as accurate as they take themselves to be in general, the subjects might go on to be as confident as they were before about the accuracy of their future judgments (Kahneman and Tversky, 1996, p 588). Our overconfident physician, even after concluding "I am overconfident", would probably go on to be as confident as she was before. Cassam concludes "Self-knowledge does not automatically lead to self-improvement" (Cassam, 2017). This implies that substantial knowledge of one's own stealthy epistemic vices does not satisfy the value condition automatically.

Although skeptical about the efficacy of self-knowledge in bringing about substantial change, Cassam is not pessimistic: he thinks we can improve on diagnostic error. He proposes forcing functions as a promising alternative to the self-improvement that would be brought about by the self-knowledge of overconfidence, were this knowledge efficacious. Forcing functions are mechanisms/procedures which are built in the design of a system with a

view to minimizing error. The crucial feature of these functions is that they work on the agents of the system from the outside; they are in a way forced on them. For example, the institution, in our case the hospital, can implement a series of procedures which will minimize diagnostic error due to overconfidence. The overconfident physician might be forced to follow these procedures by the administration as a general policy, perhaps not even knowing why she is following them. In this way the hospital can achieve a reduction of diagnostic error. For Cassam, this crucial feature of forcing functions is their main attraction: they take the control of the epistemic vice out of the hands of the individuals who have those vices.

When our overconfident physician is not in charge of her own vice, the knowledge of her vice is used to manipulate her behavior for the common good. But this is third-personal knowledge of the manipulated agent by another agent, in this case, the institution. It is the institution's knowledge of physician overconfidence that is effective in implementing procedures to reduce diagnostic errors resulting from that overconfidence. The practical value of this knowledge, however, does not tell us anything about the practical value of substantial self-knowledge of one's own epistemic vices. There is no knowledge here that brings about a substantial change in the physician's character or in her actions qua self-knowledge. But then Cassam's examination of the obstacles in the way of acquiring self-knowledge of stealthy epistemic vices seems to have this surprising upshot: The more substantial self-knowledge is with respect to obstacle conditions, the less likely that it can be of any practical value qua self-knowledge!

### Substantial self-knowledge as evidential/third-personal

We have seen in Section 'Knowledge of one's overconfidence as substantial self-knowledge' that, for Cassam, substantial self-knowledge always requires epistemic work. Each characteristic in Group B highlights an aspect of this requirement. Since overconfidence is a stealthy epistemic vice which impedes self-critical reflection, the knowledge of it is even less direct than the knowledge of a less stealthy vice, such as carelessness, would be. In our example of the previous section, our overconfident physician acquired the knowledge of her overconfidence in three steps. At the first step, she acquired evidential knowledge of Dr. Sure's overconfidence. At the second step, she acquired evidential knowledge of her identity as Dr. Sure. And then at the third step, she acquired inferential evidential knowledge of her own overconfidence. So this knowledge is doubly third-personal, so to speak: at the first step, it is third personal knowledge of another; at the second step, it is third-personal knowledge of oneself. The more steps in acquisition, the more substantial the knowledge acquired. The first step is required for the agent to detach herself from the object of her knowledge and thus bypass the epistemic vice. We deliberately created a case where our overconfident physician could approach the data impartially by presenting the evidence as if it were insignificant to her self-conception. In fact, in any knowledge one acquires about oneself via being presented evidence about oneself, there is an element of detachment, a setting oneself apart from what one examines. We can see the first step as this detachment being pushed to its limit, where the awareness of the identity of the known and the knower is completely lost. And then we can see the consequence of detachment for the practical significance of the knowledge acquired thanks to it: What I come to know about myself as the knowledge of another might not get a grip on me as a practical agent.

According to Cassam, detachment from the object known, as required for receiving evidence about oneself, is the source of significant, non-trivial knowledge of oneself. I may know I have



chest pains without epistemic work that requires detachment, but only after a medical examination would I know whether I have a heart condition. So Cassam is right. But knowledge of my heart condition would not make a difference in my conduct were I not the one who is suffering chest pains, who has shortness of breath when playing with her toddler son, who fears that she will die before seeing him grown up. And all these are given to me in my experience directly; I know all these trivially. The knowledge of my heart condition is integrated with all these other trivially known attitudes and that is why it is not just any fact about me which I must acknowledge on pain of irrationality, just as our overconfident physician had to acknowledge that she is overconfident once she realized that she is Dr. Sure. I know I have a heart condition in being scared, in regretting, in taking my pills, in changing my diet, in the bitterness I fight against. Substantial self-knowledge can be efficacious *qua* self-knowledge only when it is integrated with the attitudes, states, and actions of the agent which are available to her trivially. That is, only when it is available to the agent from the first-person standpoint, the standpoint from which she acts, does substantial knowledge matter.

### Substantial self-knowledge via traumatic experience

We have seen how third-personal knowledge of oneself requires the further work of integration into the standpoint from which one acts in order to be of some practical value. The stealthy epistemic vices are challenging cases of self-knowledge, not just because, as Cassam nicely shows, they are self-concealing. Since they infiltrate the capacities for self-reflection, they render the agent incapable of facilitating the required integration deliberately. We need our reflective capacities not only to gain knowledge about ourselves, but also to unify our cognitive lives so that different standpoints from which we comprehend ourselves can feed into a cohesive self-conception. That is why even after acquiring the knowledge of her own vice against all odds, Dr. Sure remains overconfident. If the knowledge of her own overconfidence had been incorporated into her self-conception, it would have been immediately transforming. In the heartfelt utterance of “I am overconfident”, one already becomes less so.

Since overconfidence itself, by not only being self-concealing but also infiltrating the capacity for self-reflection, stands in the way of an integrated knowledge of one's own overconfidence, a third-personally acquired knowledge of one's own overconfidence cannot be of practical value *qua* self-knowledge. It seems only a crisis or watershed event of some sort can do the trick: a lightning strike to the core, an immediate, direct, non-reflective, non-evidential illumination, something that bypasses all the rational capacities which are infiltrated by the vice itself:

At times, however, a single experience, which can be shared in such moments, can change thinking. This happens especially if the event is emotionally laden because we tend to be particularly influenced by emotionally arousing experiences. For example, if a physician misdiagnoses a headache as benign and the patient subsequently dies from a subarachnoid hemorrhage, the powerful impact of this experience might produce long-standing changes in the physician's approach towards the diagnosis of patients presenting with headache (Croskerry et al., 2013, p ii 66).

Traumatic experiences sometimes have the power to bring about fundamental changes. Cassam also mentions this in passing (Cassam, 2017), but he does not attempt to give an account of why traumatic experiences might have such power. Traumatic experiences might have such power because they are sudden, intense, direct, involuntary, unreflective. One knows one's experiences trivially, even when the content of the experience is

not trivial at all. Traumatic experiences can confront us with substantial content, and with the immediacy of a punch to the stomach. They deliver their message not from the perspective of one observing oneself as an other, but rather from the perspective of one who suffers. They do not leave room for detachment. But, like trivial self-knowledge, they do not satisfy any of the Group B characteristics for substantiality either. In undergoing such an experience, I might come to know that I am overconfident trivially, although the fact that I am overconfident is not trivial at all. Traumatic experiences *can* penetrate to the core of one's self-conception *not by reflective integration but by brute force*, so to speak. Hence, their message can transform one's self-image by bypassing the rational capacities which are infiltrated by the epistemic vice in question. The standpoint from which one receives the message is the standpoint from which one acts intentionally, for example, in the case of the physician, the standpoint from which she makes her diagnosis. Hence, such message has immediate practical impact. That is, unlike Cassam's conception of substantial self-knowledge of one's overconfidence, trivial knowledge of one's overconfidence via traumatic experience satisfies the value condition *qua* self-knowledge.

It is worth emphasizing that here I only explain why in *some* cases a traumatic experience can bring about substantial change where explicit knowledge of the vice gained third-personally cannot. My point is that in order for substantial change to occur the agent must be able to comprehend the knowledge of her vice from the standpoint from which she acts intentionally. Since this is also the standpoint from which one knows oneself trivially, this point speaks for the practical value of trivial self-knowledge. However, since the way in which traumatic experiences work on the agent is by bypassing her rational capacities, the impact of this experience is not under rational control. Hence, there is no guarantee that any particular traumatic experience will cause a change or that this change will be for the better. Here, I have only explained why a traumatic experience is capable of causing substantial change in the event that it actually does.

### Value of trivial self-knowledge

What Cassam calls trivial self-knowledge is what is studied by most philosophers. One of Cassam's examples of trivial self-knowledge is the knowledge that I believe that I am wearing socks.<sup>2</sup> No doubt, whether I believe that I am wearing socks is one of the least interesting things I or anyone can know about me. So Cassam rightly wonders “why should anyone care why or how I know that I believe that I am wearing socks?” (Cassam, 2014, p 29). Surely Cassam is right in that the content of this knowledge lacks significance. Moreover, in general there is not much danger of getting this type of fact wrong. Hence, why bother to study it?

The interest philosophers have in trivial self-knowledge is not in the particular facts we know trivially. Most of the time, these facts are not substantial in the sense of the first English dictionary definition of the word: of considerable importance, size, or worth. But the particular way in which we know these facts is revealing of the type of beings we are. How we know these trivial things is the key to understanding what it means to be a self-conscious being. Or so most of the philosophers who study self-knowledge think. Hence, this type of knowledge is substantial in the sense of the second definition we find in the dictionary: it concerns the essentials of something. Trivial self-knowledge is revealing of the essence of human beings. Consequently, the way in which trivial self-knowledge is known is also substantial in the first sense: what could be more important than understanding the most distinctive feature of the human condition? So our interest in trivial self-knowledge is metaphysical, and in pursuing this interest we do philosophy of mind, and not primarily epistemology (Boyle, 2015).

But our current analysis points to some further interest: trivial self-knowledge, in addition to its relevance in understanding the human condition in the abstract, is needed for the self-understanding of individual human beings in the ordinary sense. It is relevant to their grasp of their character traits, motives, biases, and desires in a way that will enable them to change, reinforce, or abandon them, and so deliberately shape themselves and the world around them. By reflecting on the cases where substantial self-knowledge does not get a grip on the agent to facilitate self-improvement, we have arrived at a *low road* defense of trivial self-knowledge. Substantial knowledge of oneself can only be efficacious self-knowledge when it is integrated with the attitudes of the agent she knows trivially, because only then can it get a grip on her at the standpoint from which she acts. An agent can act intentionally, as opposed to acting under manipulation or self-manipulation (e.g., when one has self-imposed forcing functions), only when the attitudes she takes to be salient to the situation at hand, including her own epistemic standing—for example, her overconfidence—are available to her trivially.

### Conclusion

From Cassam we learnt the value of the information we can gain about ourselves by observing ourselves from third person point of view. He also taught us that the content of this type of knowledge shows us where we need improvement. But we also learnt from him not to overestimate the benefits of substantial self-knowledge. In this paper, I have argued that we should also learn not to underestimate the benefits of trivial self-knowledge. Substantial knowledge matters in the practical and moral sense through trivial self-knowledge. It turns out that the low road defense Cassam gives for substantial of self-knowledge works as a defense of the value of trivial self-knowledge.

Is substantial self-improvement possible? My answer is yes, it is possible. I think, even in the cases of stealthy epistemic vices such as physician overconfidence, we can do better than implementing forcing functions or adopting strategies which take away control from the agent who needs improvement. We should aim higher than damage control; we should aim to make a positive difference in the agent's overall happiness and well-being. In order to achieve this aim, I suggest we reflect on how traumatic experiences facilitate positive transformation, and investigate how we can achieve the integration of substantial knowledge about oneself with one's self-conception, so that its content will be available to the agent trivially in her actions and deliberations. Now that we see that integration is the key, we can devise strategies which aim for integration. An understanding of the salient features of traumatic experiences might suggest methods we have not considered before, the methods which will not only reduce diagnostic errors, but also cure overconfidence.

Only when agents mean well and do well in their trivially known actions, which are based on their trivially known attitudes, and only when the agents do better as an expression of their substantial self-improvement, will we live in a world which is substantially better than this one. That is why, as Socrates taught, when we want to connect self-knowledge with leading a morally right life, trivial self-knowledge takes center stage.

Received: 9 March 2017 Accepted: 21 November 2017

Published online: 05 December 2017

### Notes

- 1 Here we are focusing on self-knowledge; but knowledge of others in so far as it is relevant to and at odds with one's self-conception exhibits the same peculiarity. It might be almost as hard for us to accept the faults of our family members, friends, or fellow citizens as it is to accept our own.
- 2 Other examples Cassam gives: knowing that you are in pain, knowing that you believe it is raining. I think all his examples are knowledge of one's mental states or sensations. I did not come across an example of knowledge of one's own intentional actions.

### References

- Anscombe GEM (1963) *Intention*. Cornell University Press, Ithaca
- Berner E, Graber M (2008) Overconfidence as a cause of diagnostic error in medicine. *Am J Med* 121:S2–S23
- Borrell-Carrió F, Epstein R (2004) Preventing errors in clinical practice: a call for self-awareness. *Ann Fam Med* 2:310–316
- Boyle M (2015) Critical study: cassam on self-knowledge for humans. *Eur J Philos* 23(2):337–348
- Cassam Q (2014) *Self-knowledge for humans*. Oxford University Press, Oxford
- Cassam Q (2015) Stealthy vices. *Social Epistemol Rev Reply Collect* 4:19–25
- Cassam Q (2016) Vice epistemology. *Monist* 99:159–180
- Cassam Q (2017) Diagnostic error, overconfidence and self-knowledge. *Pal Commun* 3:17025, <https://www.nature.com/articles/palcomms201725>
- Croskerry P et al. (2013) Cognitive debiasing 2: impediments and strategies for change. *BMJ Qual Saf* 22:ii65–ii72
- Donnellan K (1963) Knowing what i am doing. *J Philos* 60:401–409
- Kahneman D, Tversky A (1996) On the reality of cognitive illusions. *Psychol Rev* 103:582–591
- Özaltun E (2016) Practical knowledge of what happens: a reading of §45. *Klēsis Rev Philos* 35:52–73

### Data availability

Data sharing is not applicable to this article as no datasets were generated or analyzed.

### Additional information

**Competing interests:** The author declares no competing financial interests.

**Reprints and permission** information is available online at <http://www.nature.com/reprints>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017