The Intrinsic Value of Self-Knowledge

Eric Schwitzgebel

February 6, 2015

For presentation at the April 2, 2015, Pacific APA Author-Meets-Critics Session on Quassim Cassam's *Self-Knowledge for Humans*

Note to reader: If you're not familiar with Cassam's book, you can skip straight to Section 3. Section 3 to conclusion can be read (almost entirely) as a freestanding defense of the intrinsic value of self-knowledge.

1. Introduction.

So as I'm reading through Quassim's book, I'm in despair. What am I going to disagree with? This pretty much all seems right – okay, I have a few quibbles about whether Quassim is critiquing the most interesting version of transparency theory, and what does he mean, really, by "inference"? – but none of that seems worth hanging a commentary on, for me.

And then I hit the final chapter, titled "The Value of Self-Knowledge". What a relief! I totally disagree! So that's what I'm going to pick at.

Quassim argues that self-knowledge doesn't have intrinsic value. I think self-knowledge does have intrinsic value. I think if we lack self-knowledge, that stinks, even if we're happy.

According to Quassim, self-knowledge only has instrumental value. Substantial self-knowledge of your character traits, your values, your attitudes, can sometimes help you get what you want, can sometimes bring you more happiness or well-being, but other than the fact that it might contingently help you attain those other ends, it has no value of its own. This instrumentalist view of self-knowledge helps take some of the sting out of Quassim's pessimism about self-knowledge. Look, if we don't have very good self-knowledge, no biggie! It doesn't matter so much, anyway, as long as we're happy. And we are happy! Aren't we? Wait....

I see three ways to argue that something has intrinsic value. One I'll call the Argument from Addition and Subtraction. One I'll call the Argument from Nearby Cases. And one I'll call the Argument from Identity.

2. Quassim's Argument.

First, let me summarize Quassim's argument as I understand it. It appears to be either a burdenof-proof argument or an inference to the best explanation. He says it's not clear why selfknowledge would be valuable independent of its instrumental contribution to your happiness or well-being. I'm not sure quite what Quassim means by "well-being", but it seems evident from his argumentative structure that he can't regard it as partly *constitutive* of well-being that one have self-knowledge. Maybe by "well-being" Quassim means just hedonic pleasure and one's subjective sense of how well things are going; or maybe he can also include goods like friendship and artistic and intellectual accomplishment. I'll return to this issue later.

In any case, Quassim considers and rejects three ways of trying to establish that self-knowledge is intrinsically valuable. If he intends a burden-of-proof argument, his conclusion would be on the grounds that the burden of proof is not met. If he intends a best-explanation argument, his conclusion would be on the grounds that the best explanation of previous philosophers' failure to provide a compelling argument for the intrinsic value of self-knowledge is that self-knowledge is not in fact intrinsically valuable.

The three arguments he considers are these. One: that substantial self-knowledge is essential for rationality, and rationality is intrinsically valuable. Two: that substantial self-knowledge is essential for authenticity, and authenticity is intrinsically valuable. Three: that substantial self-knowledge is essential for living a unified life, and living a unified life is intrinsically valuable. I'm inclined to grant Quassim his arguments on all three points: You *can*, he rightly says, be rational, be authentically true to yourself, and live a unified life without substantial self-knowledge. Substantial self-knowledge is not essential for any of these things. So all three arguments start from a similar type of false premise.

But there are other arguments for the intrinsic value of self-knowledge. Here are three.

3. Argument 1: The Argument from Addition and Subtraction.

Here's what I want you to do: Imaginatively *subtract* our self-knowledge from the world, or at least large portions of our self-knowledge, while keeping everything else as constant as possible, especially our happiness or subjective sense of well-being. Now ask yourself, is something valuable missing?

Now imaginatively *add* lots more self-knowledge to the world, while keeping everything else as constant as possible, especially our happiness or subjective sense of well-being. Now ask yourself, has something valuable been gained?

Okay, I see two big problems with this method of philosophical discovery. Both problems are real, but they can be *partly* addressed, I think.

One problem is: The subtraction and addition are too vague to imagine. To really do this exercise, we need to get into details. And the details are going to be complicated, and maybe we'll get them wrong, and any detailed imagination is going to be partial and therefore possibly misleading. Right! But we can at least try and start and see where we seem to be headed. I can imagine my own case, for example. What if I suddenly knew a lot more about why I'm so attracted to philosophy? Stipulate that this new knowledge has no effect on my overall subjective sense of well being and, if possible as an idealization, that the knowledge has no effect on my future choices. I'm not depressed upon seeing my base and stupid motives uncovered. I don't gain any insight into how to further progress in my goal of being a terrific philosopher. I

just now have this self knowledge that I otherwise wouldn't have. Doesn't that seem good? Doesn't it seem like something of value has been gained, somehow, independently of those consequences?

Or subtract: Right now I think I know about myself that I'm kind of a middling extravert and a kind of a middling racial egalitarian with, probably, an ordinary middle-class-white-guy set of implicit racial biases. Subtract this knowledge. I have no idea whether I'm an introvert or an extravert, or I wrongly think I'm an introvert. Stipulate again: no practical consequences. Or suppose I have no idea where I am in implicit racial egalitarianism; maybe I falsely think I'm wholly bias free. Suppose again, no practical consequences. Isn't something important lost?

Another way to do addition/subtraction is to imagine an alien culture. On the far side of the galaxy, suppose, there's a culture that we will never interact with. What would I wish for it? Would I wish for a culture of happy beings with no self-knowledge? Or, if I imaginatively added to this culture substantial self-knowledge, to the extent possible not affecting anything else, would I be imagining something better, a better state of affairs in the universe? I think it would be better. I wish my aliens substantial self-knowledge, even if there's no gain from it as long as there isn't too much happiness lost.

We can contrast this with a case in which addition and subtraction leave us cold: seas of iron in the planet's core. Apart from its effects on planetary inhabitants, I don't care if Earth or our alien planet has more or less iron at its core. Add or subtract away. Whatever.

We could do a lot more imaginative addition and subtraction, should do a lot more, before feeling comfortable that we know where such exercises lead. I think they point toward the intrinsic value of self knowledge.

Or rather, I should say – and this brings me to the second methodological concern – they reveal the fact that *I* regard self-knowledge as something that has intrinsic value. You might differ. You might try these exercises and think: happy aliens, no self-knowledge, great! They're not missing anything important. You might think that unless there's some practical purpose that's served by knowledge of your personality traits and attitudes toward other sexes and races, you might as well do without that knowledge; you might as well not know.

I think this is just the methodological problem that's at the root of all value inquiries. I can't rationally compel you to share my values if you start far enough away in value space. I can just invite you to think about what your values are, and how they fit together, suggest that if you think about it, you'll probably find that you do already share these values with me or at least already share some commitments that should lead you toward those values.

4. Argument 2: The Argument from Nearby Cases.

Here's another way to argue for intrinsic value: Take a case of something that you and your interlocutor both presumably regard as having intrinsic value. Then point out that the target thing is a similar type of case, which doesn't merit being treated differently. For example, if you

and I both agree that beautiful paintings are intrinsically valuable, then maybe I can use that as a lever in arguing that beautiful rap songs are intrinsically valuable too, if you were having some doubts.

I'm going to assume that you agree that knowledge in general is intrinsically valuable. A world of unreflective bliss would be lacking something important that a world of bliss plus knowledge would possess. I want my alien world on the far side of the galaxy to be a world with inhabitants that know things, not just a bunch of ecstatic oysters.

Now it doesn't follow from knowledge's being intrinsically valuable in general, if you agree with that, that self-knowledge is also intrinsically valuable. Maybe self-knowledge is an exception to the general case. So here's how I'd argue instead. Knowledge of other people is intrinsically valuable. But knowledge of oneself is a similar case, so self-knowledge is intrinsically valuable.

Use the Argument from Addition and Subtraction to support the intrinsic value of knowledge of other people, specifically the motivations and values and attitudes of your friends and family. Subtracting from the world people's psychological knowledge of their friends and family would make the world a worse place. Adding more of that knowledge, other things held constant to the extent possible, makes the world a better place.

Now add the Nearby Cases step: It would be weird if psychological knowledge of your friends and family had intrinsic value but psychological knowledge of yourself did not. There doesn't seem to be any good ground for treating self-knowledge as something of lesser intrinsic value. The cases seem similar in the relevant respects. It's good to know about people intimate to you. You yourself are one of those people intimate to you. The world is a richer place because pieces of it can gaze knowledgeably upon themselves and the others around them.

Maybe I can phrase this as a challenge: Would Quassim deny that knowledge of other people is intrinsically valuable? If not, then what's the relevant difference between that case and the self-knowledge case? Of if knowledge of other people's minds is *not* intrinsically valuable – now I seem to be turning my argument into a dilemma – is there *anything* intrinsically valuable besides pleasure or hedonic well-being?

Maybe he'd say not – maybe Quassim is a hedonist. I don't really have an argument against that. I don't think most philosophers, if they really think about it, are hedonists. I think we find things besides pleasure valuable – beauty and intellect, for example. But if Quassim is a hedonist then let's bring that out onto the table as part of what's behind his argument against the intrinsic value of self-knowledge.

Back on the other horn: If Quassim is *not* a hedonist, but thinks that there's intrinsic value in the rich flourishing of human intellectual and artistic capacities – then why wouldn't self-knowledge be part of that flourishing, an intrinsically valuable part of our overall objective well-being?

5. Argument 3: The Argument from Identity.

Here's a third way to argue that something is intrinsically valuable: Argue that it is in fact identical to something that we already agree is intrinsically valuable.

What is self-knowledge? Here, maybe, I'll give a somewhat different account of the phenomenon than Quassim would prefer, so he can push back against it if he likes. On my view, to know some psychological fact about oneself is to possess a suite of dispositions or capacities with respect to your own psychology. That's too abstract to be comprehensible as-is, so let me develop a couple examples.

What is it to know that you are an extravert? It is in part the capacity to say, truly and sincerely, "Yeah, I'm an extravert". It is in part the capacity to respond appropriately to party invitations, by accepting them in anticipation of the good time you'll have. It is in part not to be surprised to find yourself smiling and laughing in the crowd. Etc.

What is it to know that you're angry? It is in part to have the capacity to say, truly and sincerely, "I'm angry!" It's in part to know not to pick up the phone if you're expecting a call from someone you need to be calm with. It's in part to feel unsurprised at the fact that you're talking loudly and making fists. It's in part the capacity to conclude that someone in the room is angry.

My thought is: Those kinds of dispositions or capacities are intrinsically valuable. I hope you'll agree with that. They are central to living a rich, meaningful life. If we subtract them away, we impoverish ourselves. Human life wouldn't be the same without this kind of self-attunement or structured responsiveness to psychological facts about ourselves, even if we might experience just as much pleasure. And self-knowledge is not just some further thing floating free of those dispositional patterns, those capacities, that could be subtracted without taking them away too. Self-knowledge is not some independent representational entity only contingently connected with those patterns. It *is* those patterns, those capacities. If those patterns and capacities are intrinsically valuable, then self-knowledge is too, by the principle that identical things have identical properties.

You might notice that this third argument creates some problems for the straightforward application of the first argument, the Argument from Addition and Subtraction. Perhaps, in trying to imagine subtracting self-knowledge from the world while holding everything else constant to the extent possible, you were imagining or trying to imagine holding constant all those dispositions I just mentioned, for example the capacity to say yes appropriately to party invitations. If my view of knowledge is correct, you can't do that. What this shows is that the Argument from Addition and Subtraction is not as straightforward as it might seem at first glance. It needs to be handled carefully. But still, that doesn't mean it's a bad argument.

6. Conclusion.

So I guess I'm more of a pessimist that Quassim. I agree with him that our self-knowledge of our morally most important traits and attitudes is poor. We tend not to know what we really value and what our moral character really is like, for example. We're amazingly ignorant about

what makes us happy. But unlike Quassim, I think this is a disaster. It's instrumentally bad because it leads to terrible decisions. And it's intrinsically bad because self-knowledge, when we have it, is one of the most intrinsically valuable things in human life.