

## Reflection On: *On Reflection*

Declan Smithies

The role of reflection in human rationality has been a central theme in the history of philosophy from Plato and Aristotle through Descartes and Kant to contemporary figures as diverse as Bonjour and Sosa, Shoemaker and McDowell, Frankfurt and Korsgaard. In his book, *On Reflection*, Hilary Kornblith criticizes what he regards as a chronic tendency in philosophy towards inflating the significance of reflection in ways that manifest a combination of philosophical naiveté and scientific ignorance about how reflection actually works. He writes: “Philosophers have typically assigned a great deal more value to reflection than it really deserves, and, more than this, they have done so because their view of what reflection is and what it is capable of achieving is terribly inaccurate” (2012: 1).

The book has a broad scope. It encompasses no less than four different areas of philosophy: epistemology, philosophy of mind, philosophy of action, and metaethics. The role of reflection is examined in connection with knowledge and justification (Ch. 1), belief and reasoning (Ch. 2), agency and free will (Ch. 3), and the nature of normativity (Ch. 4). The final chapter (Ch. 5) turns to cognitive science where Kornblith surveys dual-process theories of reasoning and argues that much of this work lends credence to the skeptical conclusion of the book.

*On Reflection* has much to recommend it. The book is written in an accessible and engaging style; it draws illuminating parallels between various areas of philosophy; it brings philosophy into contact with contemporary work in cognitive science; and it poses a powerful challenge to traditional views about the importance of reflection in understanding human nature.

In my opinion, Kornblith’s book provides a valuable corrective to views that are overly sanguine about the powers of reflection. At the same time, I am not yet persuaded that we should accept its negative conclusion about the philosophical significance of reflection. This is a substantive disagreement, but not one that is

easily resolved, and so I want to begin by acknowledging the contribution that Kornblith's book makes in articulating such a clear and forceful challenge. My goal in these comments will be to take some preliminary steps towards meeting this challenge in the hope of advancing the conversation further.

I will begin – as Kornblith does – with the debate in epistemology between internalist and externalist theories of knowledge, since it provides a good entry point into more general debates about the role of reflection in human nature. In Chapter 1, Kornblith criticizes theories of knowledge that give a central role to reflection, including those of Bonjour and Sosa. My main aim in these comments is to explain how my own theory (in Smithies 2015) provides the resources to answer the main challenges that Kornblith raises in his discussion of Bonjour and Sosa. A secondary aim is to sketch a more general account of the philosophical significance of reflection that withstands Kornblith's critique.

## **1. Justification and Reflection**

Bonjour holds that a belief is knowledge only if it is justified, where a justified belief is one that is held on the basis of reflection. He motivates this view with his example of Norman, the clairvoyant:

Norman, under certain conditions which usually obtain, is a completely reliable clairvoyant with respect to certain kinds of subject matter. He possesses no evidence or reasons of any kind for or against the general possibility of such a cognitive power or for or against the thesis that he possesses it. One day Norman comes to believe that the President is in New York City, though he has no evidence either for or against this belief. In fact the belief is true and results from his clairvoyant power under circumstances in which it is completely reliable. (1985: 41)

Bonjour argues Norman's belief is unjustified. The argument is that Norman neglects his epistemic duty to reflect on the credentials of his belief with the result that his belief is held in a way that is irresponsible and so unjustified. He writes:

Norman's acceptance of the belief about the President's whereabouts is epistemically irrational and irresponsible, and thereby unjustified, whether or not he believes himself to have clairvoyant power, so long as he has no justification for such a belief. Part of one's epistemic duty is to reflect critically upon one's beliefs, and such critical reflection precludes believing things to which one has, to one's knowledge, no reliable means of epistemic access. (1985: 42)

We can make Bonjour's reasoning explicit as follows. The first premise states a necessary condition for justified belief, according to which a belief is justified only if it is held on the basis of reflection, while the second premise adds that Norman does not satisfy this condition:

- (1) A belief is justified only if it is held on the basis of reflection.
- (2) Norman's belief is not held on the basis of reflection.
- (3) Therefore, Norman's belief is unjustified.

The key idea then is that a justified belief is one that withstands reflective scrutiny.

A different account of the connection between justification and reflection is proposed by Alston (1989) and developed further in work of mine (Smithies 2015). Alston's claim is not that reflection is needed in order to make one's beliefs justified, but rather that reflection is the activity with reference to which the standards for justification are defined. He writes:

It would be absurd to suggest that in order to be...justified, a belief must actually have been put to the test and emerged victorious.... [T]he idea is rather that what it is for a belief to be justified is that the belief and its ground be such that it is in a position to pass such a test; that the subject has what it takes to respond successfully to such a challenge. A justified belief is one that *could* survive a critical reflection. (1989: 225)

To a first approximation, Alston's proposal is that justification is the epistemic property in virtue of which a belief has what it takes to survive reflective scrutiny. Bonjour claims that a justified belief must *actually* withstand the test of reflective scrutiny, whereas Alston claims that a justified belief is one that has the *potential* to withstand such a test. On this view, a justified belief is *stable under reflection* in the sense that if it were subjected to reflective scrutiny, then it would survive; and moreover, it would survive on the basis on which it is actually held.<sup>1</sup>

We can use this theory of justification to give an alternative explanation of why Norman's clairvoyant beliefs are unjustified. Norman's beliefs are not stable under reflection because, as Bonjour points out in the passage quoted above, "reflection precludes believing things to which one has, to one's knowledge, no reliable means of epistemic access." So, we can argue, as before, that Norman does not satisfy a necessary condition for justified belief:

- (1) A belief is justified only if it is stable under reflection.
- (2) Norman's belief is not stable under reflection.
- (3) Therefore, Norman's belief is unjustified.

In what follows, my aim is to defend this account of the connection between justification and reflection against various objections raised in Kornblith's book. I divide these into four categories: (i) the over-intellectualization problem; (ii) the regress problem; (iii) the empirical problem; and (iv) the value problem.

## **2. The Over-Intellectualization Problem**

The first problem that Kornblith raises is *the over-intellectualization problem*. He writes: "Most of our beliefs are formed without the benefit of critical reflection, and

---

<sup>1</sup> This clause is needed to rule out unjustified beliefs that would survive reflection if their basis were to change during the process. Additional qualifications are needed to avoid the conditional fallacy objection, but see Smithies (2015) for details.

so, on Bonjour's view, most of our beliefs are not in fact justified, and we thus have precious little knowledge" (2012: 11).

Bonjour's account has skeptical implications. Traditionally, the skeptic argues that knowledge is impossible, whereas Bonjour maintains that knowledge is possible in principle, even if this possibility is rarely actualized in practice. Still, it is quite implausible to suppose that beliefs held unreflectively on the basis of perception, memory, testimony, induction, and deduction can never be justified by the evidence on which they are held and so can never constitute knowledge.

One response is to weaken the conditions for justification: for instance, it might be said that one need not actually reflect in order to have justified beliefs so long as one has the capacity to do so. But this condition is not *sufficient* for justification, since we may suppose that Norman has the capacity for reflection but negligently fails to exercise the capacity: his belief should not thereby count as justified. And the more important point for current purposes is that this condition is not *necessary* for justification, at least not given the plausible assumption that children and animals can have justified beliefs while lacking any capacity for reflection at all.

In Chapter 2, Kornblith defends and motivates this assumption. Many philosophers (including Shoemaker, McDowell, Brandom, and Williams) have argued that having justified beliefs, and even having beliefs at all, requires being responsive to reasons, which in turn requires the capacity for reflection. Kornblith blocks this argument by distinguishing two senses of responsiveness to reasons: he argues that justified belief requires sensitivity to reasons, but does not require conceptualizing reasons as reasons, and so does not require any capacity for reflection. Moreover, Kornblith argues on empirical grounds for the existence of cognition without metacognition, citing the broken wing display in piping plovers, termite fishing in chimpanzees, and linguistic abilities in three-year-old children who fail the false belief task. In these cases, he argues, information is not only represented, but also integrated in ways that manifest the kind of sensitivity to

reasons that is required for justified belief. I am in full agreement with Kornblith on these points.<sup>2</sup>

A different response to the over-intellectualization problem is to maintain that justified belief requires the capacity for reflection, while denying that knowledge requires justified belief so construed. Sosa articulates a version of this response in terms of his distinction between *animal knowledge* and *reflective knowledge*:

For animal knowledge one needs only belief that is apt and derives from an intellectual virtue or faculty. By contrast, reflective knowledge always requires belief that is not only apt but also has a kind of justification, since it must be belief that fits coherently within the epistemic perspective of the believer. (1991: 145)

Animal knowledge does not require justification in any sense that enjoins a capacity for reflection: it requires only true belief that is “apt” in the sense that it derives from an “intellectual virtue” or reliable disposition of the believer. Reflective knowledge, in contrast, requires not only true belief that is formed in a reliable way, but also belief that is justified by reflecting upon the fact that it is formed in a reliable way. In a slogan, reflective knowledge is “apt belief aptly noted”.<sup>3</sup>

Kornblith’s main complaint is that Sosa’s distinction, though well defined, is not worth making because it fails to carve at the epistemological joints. In that respect, it is like the distinction between knowledge acquired on weekdays as opposed to weekends. The challenge that Kornblith lays down for Sosa and others is

---

<sup>2</sup> See Boghossian’s contribution to this symposium for doubts about the first-order account of reasoning that Kornblith endorses. My own view is closer to Kornblith’s insofar as I am inclined towards accepting a first-order account of reasoning.

<sup>3</sup> Sosa (1991: 144-5; 245) defines animal knowledge in terms of a contrast between aptness and justification, but elsewhere (e.g. Sosa 2009: 238-9) he defines it in terms of a contrast between reflective and unreflective justification. As far as I can see, the difference in presentation is purely terminological.

to explain why reflective knowledge is different in kind from animal knowledge and not just more knowledge of the same kind.

As I will explain in due course, I agree with Sosa that there is an important distinction to be drawn between animal knowledge and reflective knowledge. However, I do not accept Sosa's account of the distinction. In particular, I reject the suggestion that mere reliability of the right kind is sufficient for animal knowledge. Sosa's account puts Norman's clairvoyant beliefs on the same footing as the perceptual beliefs of children and animals insofar as they satisfy the conditions for animal knowledge.<sup>4</sup> But that is the wrong result. There is no sense in which Norman's belief is knowledge and no sense in which it is justified. To avoid this result, we cannot sever the connection between animal knowledge and reflection altogether. We just need a more nuanced account of the connection.

On my view, knowledge requires justified belief, but justified belief requires no capacity for reflection. To say that a justified belief has the potential to survive reflection is not to say that the subject has the reflective capacities needed to realize this potential. A justified belief has the potential to survive reflection in virtue of the basis on which it is held and not in virtue of the subject's reflective capacities. A justified belief is one that is stable under reflection in the sense that it would survive on its actual basis if it were subjected to reflection by some *idealized counterpart* of the subject who has the very same evidence but also has the requisite capacity for reflection.

Animals and children can have justified beliefs in the absence of any capacity for reflection so long as their beliefs are stable under reflection in the sense defined.<sup>5</sup> So, unlike Bonjour, we need not deny that animals and children have knowledge. But unlike Sosa, we needn't concede that Norman has knowledge too,

---

<sup>4</sup> Sosa (1991: 240) says that "no human blessed with reason has merely animal knowledge of the sort attainable by beasts." But we can avoid this complication by stipulating that Norman is a child or an animal with a faculty of clairvoyance and adjusting the details of the case accordingly.

<sup>5</sup> Here I disagree with Alston (1989: 226, n. 45) who denies that the concept of justification applies to dogs and other unreflective animals.

since his clairvoyant beliefs are not stable under reflection in the relevant sense. In this way, we can avoid the problems for Bonjour and Sosa alike.

### **3. The Regress Problem**

The second problem that Kornblith raises is *the regress problem*.<sup>6</sup> On Bonjour's account, a first-order belief is justified only if the second-order reflections on which it is based are themselves justified. But these second-order reflections are justified only if they are held in turn on the basis of justified third-order reflections. And so the infinite regress begins. As Kornblith remarks: "No amount of reflective scrutiny is enough, for, whenever one stops reflecting, there is always some belief playing a would-be justificatory role which has itself gone unreflected upon" (2012: 13).

Notice that it doesn't help to weaken the conditions for justified belief so that the capacity for reflection is sufficient whether or not it is exercised. Even the most reflective humans have limited capacities for reflection: in particular, there are limits on the length and complexity of the reflective processes that we are capable of undertaking. But my proposal avoids this objection. Recall that a justified belief is stable upon reflection in virtue of the basis on which it is held and not in virtue of one's reflective capacities. So even if there are limits on the length and complexity of the reflective processes that we are capable of undertaking, there are no such limits on the length and complexity of the reflective processes that our justified beliefs have the potential to withstand. A justified belief is one that has the potential to withstand reflection of any finite length and complexity when conducted by an idealized counterpart with the very same evidence together with the requisite capacity for reflection. But there are no finite limits on the length and complexity of reflective processes that our idealized counterparts can perform.

The point remains that for every finite process of reflection, "there is always some belief playing a would-be justificatory role which has itself gone unreflected

---

<sup>6</sup> Compare Bergmann's (2006: Ch. 1) dilemma for internalism: he argues that internalism is either false because it entails a vicious infinite regress or unmotivated because it does not vindicate intuitions about Norman the clairvoyant. As I explain in Smithies (2015), my proposal avoids both horns of Bergmann's dilemma.



upon". But this point is not damning for me in the way it is for Bonjour. The problem for Bonjour is that one's higher-order reflections cannot contribute towards the justification of first-order beliefs unless they are justified by reflection themselves. The skeptical conclusion is therefore guaranteed for any finite subject who cannot engage in an infinite process of reflection. Unlike Bonjour, though, I don't accept that a belief is justified only if it is reflected upon. I acknowledge that a belief can play a justificatory role without being reflected upon, so long as it has the potential to withstand reflection by an appropriately idealized counterpart of the subject. But this point can be sustained so long as there are no finite limits on the length and complexity of reflective processes that our idealized counterparts can perform.

#### **4. The Empirical Problem**

The third problem that Kornblith raises is *the empirical problem*. Reflection is sometimes thought to make us more reliable by weeding out logical fallacies, hasty generalizations, baseless prejudice, and wishful thinking. Indeed, this is one clear rationale for thinking that reflection is necessary for justified belief. However, the empirical evidence suggests that reflection does not always make us more reliable and indeed often makes us *less* reliable. Thus, Kornblith concludes: "What commonsense tells us is a way of screening our beliefs in order to make them more accurate turns out, instead, in many cases, to be a route to little more than self-congratulation" (2012: 3).

Here are the main points of Kornblith's review of the empirical literature. There is a large class of cases in which our beliefs are influenced by seemingly irrelevant factors that we are unaware of and whose influence is undetectable by means of introspection. At the same time, we have a tendency to engage in confabulation that yields a post-hoc rationalization for these beliefs, while remaining ignorant of the fact that this is what we're doing. To mention just one classic study from a much larger literature: Nisbett and Wilson (1977) found that subjects displayed a strong right-hand-side bias when choosing between qualitatively identical pairs of socks, but were unaware of this bias, while also tending to rationalize their choices by citing non-existent differences in texture,

color, and so on. Kornblith surveys many other experiments in this vein and draws the following conclusion: “Asking subjects to introspect more carefully, or think longer and harder about the sources of their beliefs, is entirely useless in many of these cases” (2012: 23).

I have several points to make in response to this empirical problem. The first point is that reflection can make us more reliable when it is done well. In Nisbett and Wilson’s study, for instance, reflection alone cannot yield knowledge of the influence of right-hand-side bias, but it can yield knowledge that there’s no reason to believe that the socks on the right are qualitatively different from those on the left. Reflection, when it is done well, has the power to override confabulation and to yield knowledge that the choice was made for no good reason. Of course, what the empirical findings show is that reflection is not always done well – we do sometimes engage in confabulation and perhaps much more often than we tend to realize.<sup>7</sup>

The second point, however, is that the empirical evidence about the power of reflection is not universally negative. Reflection can and does increase our reliability in reasoning about a range of distinct topics. Here are three examples from a recent review by Baumeister, Masicampo and Vohs (2011).<sup>8</sup> First, *logical reasoning*: Gagne and Smith (1962) found that performance on the Tower of Hanoi problem was improved in subjects who were required to verbalize their reasons for each move. Second, *moral reasoning*: Small et al. (2007) counteracted the identifiable victim effect by instructing subjects to engage in deliberation about their decisions to donate money to charity. And third, *emotion regulation*: Pennebaker and Chung (2007) found that asking subjects to reflect on traumatic personal experiences caused improvements in health that resulted from analyzing the trauma.

The final point is that an account of the reflective stability of justified belief needs to be qualified in light of these empirical facts. When reflection is done poorly, justified beliefs can be abandoned and unjustified beliefs can be retained. So a

---

<sup>7</sup> The literature on implicit bias is especially relevant here. See Gendler (2014) for a recent review and discussion.

<sup>8</sup> The official topic of Baumeister’s review is the role of consciousness in reasoning, but consciousness is defined in terms of reflection.

justified belief cannot be defined as one that is stable under reflection as it is actually performed. Instead, a justified belief should be defined as one that is stable under *idealized reflection* – that is to say, reflection that is ideally rational, reasonable, or justified. On this proposal, we idealize not only the subject’s capacity to engage in reflection in the first place, but also the way in which the capacity is exercised. Ideal reflection so construed may diverge from actual reflection in its deliverances, but it should nevertheless be recognizable as an idealized version of the same activity.

In Chapter 4, Kornblith discusses a related proposal in Korsgaard’s work about the nature of reasons. She writes:

We need reasons because our impulses must be able to withstand reflective scrutiny. We have reasons if they do. The normative word ‘reason’ refers to a kind of reflective success. (1996: 93)

Korsgaard’s proposal is intended to solve a metaphysical problem about the source of normativity. Her proposal is that normativity has its source in reflection because the correct normative standards are the standards that we would endorse upon reflection. The problem is that there is no guarantee that our reflections will converge upon correct normative standards given the empirical facts about the unreliability of reflection. Therefore, Korsgaard invokes an idealization: she appeals to “a person who reasons all the way back, who never gives up until there is a completely undeniable, satisfying, unconditional answer to the question”. But the problem remains since idealizing the length of reflection without idealizing its quality does nothing to guarantee that reflection will converge upon correct normative standards. What Korsgaard needs is a more robustly normative idealization, but she cannot avail herself of this without undermining the metaphysical project of explaining the source of normativity.

My response to the empirical problem invokes a normative idealization. I claim that a justified belief is stable under reflection that is ideally rational, reasonable, or justified. This kind of normative idealization is useless for

Korsgaard's project because the idealization presupposes normative facts of the kind she seeks to explain. But my project is distinct from Korsgaard's insofar as I am not trying to give a metaphysical account of the sources of normativity and so I can appeal to normative facts in constraining the idealization.

One might suspect that the normative idealization trivializes the proposal. Certainly, there is an element of circularity involved in analyzing a justified belief as one that is stable under ideally justified reflection. So the analysis is non-reductive, but it is also non-trivial. The general form of the analysis is as follows:

S's belief B is justified if and only if S has belief B in ideal conditions C.

If ideal conditions are defined as conditions in which one's beliefs are justified, then the analysis is trivial. But on the present proposal, ideal conditions are defined as conditions in which one's higher-order reflections on one's beliefs are justified. It is a substantive commitment of the proposal that one's beliefs are justified if and only if they can be brought into alignment with these justified higher-order reflections.

Notice also that my proposal, unlike Bonjour's, makes no commitment to the claim that we have a standing obligation to engage in reflection. Kornblith raises the objection that we have no such obligation when we have empirical reason to believe that engaging in reflection is likely to make us less reliable. What this shows, I think, is that for non-ideal agents like us, sometimes the best strategy for forming justified beliefs that are stable under reflection is actually not to engage in reflection at all. But this is consistent with the proposal that reflection sets the standards for what is to count as a justified belief in the first place.

## **5. The Value Problem**

The fourth and final problem that Kornblith raises is *the value problem*. If justification is an epistemic standard that is defined in terms of stability under reflection, then why should we care about justification – that is, why should we regard it as an important dimension of epistemic value?

One response is that justification is valuable because reflection makes us more reliable and so beliefs that are stable under reflection are objectively more likely to be true.<sup>9</sup> I agree with Kornblith that this response is inadequate. The empirical problem, as we have seen, is that reflection sometimes makes us less reliable. I have argued that we can avoid this problem by invoking an idealization. But the value problem remains, since not even ideally rational reflection is guaranteed to make us more reliable. There are two points to be made here. First, ideally rational reflection does not guarantee reliability; for instance, the victim of a Cartesian evil demon can engage in ideally rational reflection without thereby achieving a high degree of reliability. And second, whatever reliability can be achieved through rational reflection can in principle be achieved without it.

On Kornblith's view, reflection has no distinctive epistemic value. Reflection is just one way of forming beliefs among many and it is valuable insofar as it makes us more reliable and not otherwise. Thus, he writes: "From an epistemological point of view, we should value reflection to the extent that, and only to the extent that, it contributes to our reliability. Epistemologically speaking, there is no reason to value reflectively arrived at belief in general over unreflective belief." (2012: 34)

In this passage, Kornblith seems to be assuming a version of *monism* about epistemic value on which truth is the only intrinsic epistemic good.<sup>10</sup> On this view, the epistemic value of reflection must be explained instrumentally in terms of its reliability or conduciveness towards truth. Given *pluralism* about epistemic value, in contrast, there are multiple dimensions of epistemic goodness not all of which can be explained in terms of their reliability or truth-conduciveness. I think it begs the question in the present context to assume that reliability is the only dimension of epistemic value, since many internalist theories deny that all epistemic values can be explained in terms of reliability. Even so, the challenge remains to explain why

---

<sup>9</sup> There is a trivial sense in which justified beliefs are epistemically likely to be true, since the epistemic likelihood of a proposition can be defined as the likelihood that it is true on the justifying evidence that is available.

<sup>10</sup> See DePaul (2001) for the distinction between monism and pluralism about epistemic value. See Kornblith (2002: Ch. 5) for a more extended discussion of the relationship between justification and truth.

reflection is valuable given that it is not always guaranteed to increase reliability. Are there other benefits that reflection provides that cannot be achieved in any other way?

I will argue that reflection is important because it is required for being a *person* – that is, a subject who can be held *responsible* for their beliefs and actions in the sense that it is appropriate to make demands on them by adopting *reactive attitudes*. On this account, persons are distinguished from other animals by features of their individual psychology – namely, their capacity for reflection. But the significance of this capacity emerges only given its function in the social and interpersonal context of participation in human relationships.

As I use the concept, personhood is not a biological category, but an evaluative one. Frankfurt puts the point eloquently in the following passage:

The criteria for being a person do not serve primarily to distinguish the members of our own species from the members of other species. Rather, they are designed to capture those attributes which are the subject of our most humane concern with ourselves and the source of what we regard as most important and most problematical in our lives. (1971: 6)

Even if ours is the only species whose members meet the criteria for being persons, there is no reason in principle to suppose that members of other species, such as intelligent aliens, couldn't satisfy these criteria too.

So what are the criteria for being a person? Persons are distinguished from other animals by the fact that they can be held *responsible* for their beliefs and choices. This is why we regard it as appropriate to make moral and rational demands on them by subjecting them to what Strawson (1962) called *the reactive attitudes*, such as praise and blame, gratitude and resentment, and so on. We don't regard it as appropriate to adopt these attitudes towards other animals. As Kornblith remarks: "When my neighbor's dog runs loose in my garden and destroys the flowers, it is not the dog who is responsible, but my neighbor."

Responsibility is not the same as justification: it is a more demanding status. Kornblith makes a persuasive case that many non-human animals can form beliefs and perform actions in a way that is justified by their sensitivity to reasons, but he also notes that we do not hold non-human animals responsible for the justification of their beliefs and actions. The question that remains is why it is that persons can be held responsible for the justification of their beliefs and choices in a way that other animals cannot. And here I can see no plausible alternative to the traditional Lockean answer that persons are distinguished from other animals by their capacity for reflection.

The argument so far takes the form of an inference to the best explanation: the capacity for reflection is suitably correlated with responsibility and is therefore well placed to explain it. But more needs to be said in order to elucidate the connection. Why should responsibility require any capacity for reflection? The answer is that being responsible for one's beliefs and actions is a matter of being an appropriate target of reactive attitudes whose function is to make demands on one to comply with certain normative standards. But this in turn requires that one has some understanding of the normative standards that govern one's beliefs and actions together with some capacity to bring this understanding to bear in regulating one's beliefs and actions. That is to say, it requires the capacity for reflection. So the argument can be summarized as follows:

- (1) Being responsible is being an appropriate target of reactive attitudes.
- (2) Being an appropriate target of reactive attitudes requires having a capacity for reflection – that is, for understanding normative demands and bringing this understanding to bear in regulating one's beliefs and actions.
- (3) Therefore, being responsible requires a capacity for reflection.

Here it helps to reflect on the rationale for our social practice of subjecting one another to reactive attitudes in the first place. The point of adopting reactive attitudes is to make demands on one another to comply with normative standards of morality and rationality. We don't regard it as appropriate to adopt reactive

attitudes towards other animals because there is no point in doing so: they cannot understand the demands we are thereby making on them. Animals can be more or less sensitive to the normative standards of justified belief and action, but they cannot understand those normative standards and bring this understanding to bear in regulating their beliefs and actions. That is why they cannot be held responsible for the justification of their beliefs and actions.

I don't claim any originality for this line of thought: for instance, it can be found in Burge's work on the role of the first-person concept in reflection. In the following passage, Burge draws the connection between personhood and responsibility, where responsibility requires a capacity for reflection – that is, for understanding normative demands and bringing this understanding to bear in regulating one's beliefs and actions using the first-person concept:

Insofar as *full* intellectual (or any other) responsibility requires the capacity to understand the way norms govern agency and the capacity to acknowledge the responsibility, a being that lacked the first-person concept would not be fully responsible intellectually. It would not have a fully realized rational agency. Conceptualized self-consciousness seems a necessary condition for fully responsible agency. Using the first-person concept is necessary to being a fully realized person. (1998: 262)

My point here is that this line of argument can be used to answer the value problem. Reflection is valuable not because of its reliability, but because it is the *sine qua non* for being a person who can be held responsible for their beliefs and choices. Assuming that personhood is intrinsically valuable, so is reflection. To explain the value of reflection in terms of reliability alone is to overlook the evaluative significance of the distinction between persons and other animals.

Kornblith's book does not discuss Burge's work. To my mind, this is a shame because Burge proposes one of the more interesting accounts of the philosophical significance of reflection. In Chapter 3, Kornblith discusses some related themes in the work of Frankfurt, Moran, and Sosa that tie the capacity for reflection together



with notions of agency, freedom and responsibility. However, much of this discussion targets extraneous details that do not figure in the proposal articulated here. In particular, there is no commitment here to the claim that reflection involves epistemic agency, a claim that Kornblith attributes to both Sosa and Moran, and criticizes at some length. Reflection is a *sine qua non* for responsibility whether or not it involves a distinctive kind of epistemic agency.

Since Kornblith does not explicitly address this proposal, I cannot predict his response, but I can see at least two options. The more radical option is to deny that there is any important evaluative distinction between persons and other animals in respect of responsibility. The more conservative option is to maintain that there is such an evaluative distinction between persons and other animals, while denying that the capacity for reflection is what grounds the distinction.

There are hints in the book that suggest the more radical option. In Chapter 3, Kornblith considers the proposal that human cognition is distinguished from animal cognition by the role that reflection plays in cognitive self-management. On this view, reflection is presented as “the driving agent of cognitive improvement” (2012: 104). The thrust of Kornblith’s discussion is that this exaggerates the contrast between human and animal cognition by overintellectualizing the former and underestimating the sophistication of the latter. In the case of animal cognition, Kornblith argues that reflection is not needed for cognitive self-improvement, since many instances of animal learning involve not just updating beliefs in light of new information, but also more widespread changes in patterns of belief revision. Meanwhile, in the case of human cognition, Kornblith reiterates that reflection is not as effective in cognitive self-improvement as we tend to assume, since the empirical evidence suggests that it is often epiphenomenal or worse. The upshot is that human cognition and animal cognition are much more continuous on Kornblith’s view than on many of the opposing views that he is criticizing.

Now my claim is not that reflection is needed for cognitive self-improvement, but rather that reflection is needed for responsibility in a sense that involves being an appropriate target of the reactive attitudes. Much of the appeal of this claim lies in its ability to explain why we regard it as appropriate to adopt reactive attitudes

towards persons but not other animals. In this respect at least, there seems to be a striking discontinuity between human cognition and animal cognition.

One option for Kornblith here would be to argue that our practices rest on a mistake insofar as they presuppose that there is discontinuity between human and animal cognition. But this option raises difficult questions about how our practices should be revised. Should we relax them by holding adult humans to the same standards as non-human animals or should we make them more stringent by holding non-human animals to the same standards as human adults? Either way, this option is highly revisionary. The other option for Kornblith would be to argue that our practices can be explained and justified without appeal to reflection. But of course this option faces the challenge to supply such a justifying explanation.<sup>11</sup> There may be other options too. I don't know which Kornblith will take, but I expect there is much to be learned from further discussion of these options.

## **6. Conclusions**

I want to conclude by returning to Kornblith's challenge for Sosa's distinction between animal knowledge and reflective knowledge. Why should we think this distinction is worth drawing in the sense that it carves at the epistemological joints? Here is a summary of my response. Animal knowledge is justified belief, but animals cannot be held responsible for the justification of their beliefs in a sense that involves making demands on them by adopting reactive attitudes. After all, they lack any capacity for reflection and hence for understanding these normative demands. Reflective knowledge, in contrast, is not only justified but also responsibly held on basis of reflection upon normative demands. The significance of the distinction between animal knowledge and reflective knowledge traces back to our practice of holding one another responsible by making demands on each other through the

---

<sup>11</sup> Doris (2009) takes this option: he argues that avoiding skepticism about persons requires rejecting a view that he calls "reflectivism" about persons, although he does not address the challenge of proposing an alternative conception to replace it.

adoption of reactive attitudes. A theory of knowledge that neglects this distinction therefore fails to account for one of our most basic normative commitments.<sup>12</sup>

## References

- Alston, W. 1989. *Epistemic Justification: Essays in the Theory of Knowledge*. Cornell University Press.
- Baumeister, R., Masicampo, E., and Vohs, K. 2011. Do Conscious Thoughts Cause Behavior? *Annual Review of Psychology* 62: 331-61.
- Bergmann, M. 2006. *Justification Without Awareness: A Defense of Epistemic Externalism*. Oxford University Press.
- BonJour, L. 1985. *The Structure of Empirical Knowledge*. Harvard University Press.
- Burge, T. 1998. Reason and the First Person. In *Knowing Our Own Minds*, edited by C. Wright, B. Smith and C. Macdonald. Oxford University Press.
- DePaul, M. 2001. Value Monism in Epistemology. In *Knowledge, Truth and Duty*, edited by M. Steup. Oxford University Press.
- Doris, J. 2009. Skepticism about Persons. *Philosophical Issues* 19.1: 57-91.
- Frankfurt, H. 1971. Freedom of the Will and the Concept of a Person. *Journal of Philosophy* 68.1: 5-20.
- Gagne, R. and Smith, E. 1962. A Study of the Effects of Verbalization on Problem Solving. *Journal of Experimental Psychology* 63: 12-18.
- Gendler, T. 2014. The Third Horse: On Unendorsed Association and Behavior. *Aristotelian Society Supplementary Volume* 88: 185-218.
- Kornblith, H. 2012. *On Reflection*. Oxford University Press.
- Korsgaard, C. 1996. *The Nature of Normativity*. Cambridge University Press.
- Locke, John. 1968. *An Essay Concerning Human Understanding*. London: Dent.
- Nisbett, R. and Wilson, T. 1977. Telling More than We Can Know: Verbal Reports on Mental Processes. *Psychological Review* 84: 231-59.

---

<sup>12</sup> I am grateful to Brendan Dill, Lindsay Rettler, and especially Hilary Kornblith for comments on a previous draft and to everyone who participated in the symposium on Kornblith's book at the Eastern Meeting of the APA in December 2014.

- Pennebaker, J. and Chung, K. 2007. Expressive Writing, Emotional Upheavals, and Health. In *The Handbook of Health Psychology*, edited by H. Friedman and R. Silver. Oxford University Press.
- Small, D., Loewenstein, G., and Slovic, P. 2007. Sympathy and Callousness: the Impact of Deliberative Thought on Donations to Identifiable and Statistical Victims. *Organizational Behavior and Human Decision Processes* 102.2: 143-53.
- Smithies, D. 2015. Why Justification Matters. In *Epistemic Evaluation: Point and Purpose in Epistemology*, edited by D. Henderson and J. Greco. Oxford University Press.
- Sosa, E. 1991. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge University Press.
- Sosa, E. 2009. *Reflective Knowledge: Apt Belief and Reflective Knowledge, Volume 2*. Oxford University Press.
- Strawson, P. 1962. Freedom and Resentment. *Proceedings of the British Academy* 48: 1-25.