

## Utilitarian epistemology

Steve Petersen

Received: 5 April 2010 / Accepted: 2 February 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** Standard epistemology takes it for granted that there is a special kind of value: epistemic value. This claim does not seem to sit well with act utilitarianism, however, since it holds that only welfare is of real value. I first develop a particularly utilitarian sense of “epistemic value”, according to which it is closely analogous to the nature of financial value. I then demonstrate the promise this approach has for two current puzzles in the intersection of epistemology and value theory: first, the problem of why knowledge is better than mere true belief, and second, the relation between epistemic justification and responsibility.

**Keywords** Epistemology · Utilitarianism · Epistemic value · Meno problem · Value of knowledge · Justification · Responsibility · Epistemic justification · Epistemic responsibility

Standard epistemology has it that there is a particularly epistemic type of value. A belief might be disastrous in many other ways; it might bring about great misery, and be horribly ugly to contemplate, and even cause you to forget many other important things. If that belief is an instance of knowledge, though, then most would say it is valuable in at least one important way: epistemically.<sup>1</sup> The nature of this epistemic value is mysterious, however, and a variety of notable epistemologists will cheerfully confess (when pressed) that they have no good idea what it is. It is fine to rely here and there on pretheoretical notions, of course—but given the weight this particular notion has borne in recent epistemology, I think it is time to start theorizing.

---

<sup>1</sup> Knowledge is what is of typically epistemic value in *standard* epistemology; Kvanvig (2003) and others have argued otherwise, but they still agree that there is particularly epistemic value in something, such as understanding.

---

S. Petersen (✉)  
Department of Philosophy, Niagara University, PO Box 2043, Lewiston, NY 14109, USA  
e-mail: steve@stevepetersen.net

That such special epistemic value exists is, at any rate, a normative claim, and it had better be consonant with one's overall theory of value. If one is a Kantian about value, it is easy enough to imagine how to make room for an epistemic dimension; presumably epistemic value is (roughly) to do with forming correct intentions with respect to believing. It is also easy enough to see how an Aristotelian might start to accommodate epistemic value; probably, epistemic value inheres in stably virtuous epistemic habits.<sup>2</sup> An explicit account of either such picture of epistemic value would probably find quick proponents. But it is not at all clear how utilitarianism might cohere with a particularly epistemic value. According to what we might call "classical" act-utilitarianism, the *only* thing of value is welfare, and individual acts are only good instrumentally, insofar as they contribute to it. If some belief brings about more misery than happiness, the utilitarian must say it is simply of disvalue, however true, justified, *etc.* it may be. Can a philosopher be both a utilitarian and a normative epistemologist?

I argue that the answer is yes. In fact, I have come to think a utilitarian account of epistemic value provides a powerful approach to puzzles that arise on the other two pictures. I sketch such an account and its potential implications here.

## 1 Utilitarian epistemology

In a certain fussy sense, the classical utilitarian must of course say that there is no epistemic value; only states of welfare are valuable or not. Knowledge is not good in-itself, and neither is understanding, nor epistemic justification, nor true belief. All are good only insofar as they *tend* to enhance welfare. On this picture, then, the normal objects of epistemic evaluation are good only in the same way that money is good. Money, all agree, has neither intrinsic nor final value. Still, financiers study it and its means of acquisition a great deal, because it is typically of such instrumental value. Epistemology, on the utilitarian picture, is very like the study of finance.

Think again of the piece of knowledge that brings about great net misery. The utilitarian is committed to saying that this belief is of disvalue, and thus it appears that the utilitarian cannot account for epistemic evaluation. But consider an analogous circumstance: someone's financial earnings bring about great net misery. This possibility does not show that a utilitarian cannot study finance. We might even say those earnings were nonetheless of "financial value", but this does not contradict anything in the utilitarian dogma, for it is clear to us in this case that to be of "financial value" means something close to "makes money". The utilitarian epistemologist hears "epistemic value" similarly; we can say that knowledge (and more generally, perhaps, true belief) is of epistemic value without thereby committing to the existence of a value independent of welfare.

This analogy might not seem to recover any normative aspect to epistemology, since finance appears to be just about maximizing expected monetary return, and thus a purely descriptive enterprise of determining instrumental means. Similarly, if epis-

<sup>2</sup> I am thinking the likes of e.g. [Chisholm \(1977\)](#) or [Ginet \(1975\)](#) for the former, and [Goldman \(2000\)](#) or [Sosa \(1997\)](#) for the latter.

temology is just the study of how to get true belief (or knowledge, or some other epistemic good), we seem to be back to a Quinean absorption of epistemology into psychology. But consider this example, on the financial side: you are offered a contract in which you can stake everything you own for a 1% chance at a 1,000-fold return. Looking strictly at expected monetary value this is a good bet, but for most of us it is a very bad bet in terms of the contract's expected contribution to welfare. Here I think the right *financial* advice is to abstain. If so, the study of contributions to welfare—which we utilitarians take to be a significant part of normative ethics—is bound up with what look to be purely financial decisions.

Perhaps some will insist that recommending against such a contract would be to step outside the role of strictly financial advisor. Similarly, some epistemologists insist that to worry about the goodness of beliefs not ultimately contributing to their truth is to step outside epistemic bounds. As Ernest Sosa puts his *epistemic truth monism*,

Truth may or may not be intrinsically valuable absolutely, who knows? Our worry requires only that we consider truth the *epistemically fundamental* value, the ultimate explainer of other distinctively epistemic values.<sup>3</sup>

Thus on Sosa's view, a recommendation that does not aim at truth is not *epistemic*—even when it is a belief-related recommendation for the sake of welfare.

Still, the finance example above shows there is an interesting realm where finance and utility meet, even if one refuses to include it in the domain of “finance”. Similarly there is an interesting realm where belief-formation and utility meet, and that is how I propose a utilitarian epistemologist understand normative epistemology. This realm might include examining what kinds of knowledge are most important (for example, deeply explanatory knowledge vs. utter trivia), or how other epistemic states like wisdom and understanding contribute to welfare, and so on. *Perhaps* these concerns aren't strictly epistemic—it sounds like conceptual legislation to me, to say so—but at any rate they are close enough to be within a professional epistemologist's purview.<sup>4</sup> Thus my own position in effect reverses Sosa's; I might instead say

Truth may or may not be the value fundamental to epistemology, who knows? Our point requires only that welfare alone is *the* fundamental value, the ultimate explainer of all other values.

In other words, determining the exact boundary of epistemology is not so urgent when, like finance, it is understood as the study of an important *instrumental* good, rather than an ultimate (and fairly mysterious) good in its own right.

Perhaps then the best characterization of “financial value” is something like “money-related kind of thing that has high expected return on utility.” Similarly, I am suggesting the best utilitarian characterization of “epistemic value” is something like “belief-related kind of thing that has high expected return on utility.” Note that this does not make the account rule-utilitarian; the right thing to do is still the utility-

<sup>3</sup> Sosa (2007, p. 72), Pritchard (2011) suggests we understand saying some good is “epistemically fundamental” when “its value is not instrumental value relative to further [epistemic] goods.”

<sup>4</sup> Sosa calls such matters not part of the “theory of knowledge” but rather part of “intellectual ethics” (Sosa 2007, p. 89).

maximizing thing, even if it does not get financial (epistemic) value. Note also—as an important point for later—that this talk of “tendencies” and “expected” value is essential to doing utilitarian normative ethics generally. When we look at a *type* of thing, rather than a very specific thing in a very specific circumstance, we cannot talk about its actual utility; presumably anything with causal powers could in principle lead to low welfare or high. To do normative ethics requires making some generalizations, though, and so utilitarians resort to averages and expected utility. Thus when a classical act utilitarian says a charitable type of act is good and a murderous type of act bad, they mean that charitable acts have a high *expected* utility (absent further information) and murderous acts have a low one. The point is, for the utilitarian to generalize about the value of anything (other than welfare) requires a certain degree of uncertainty about particulars.

One special case of such utility-based generalizations is the financier, who can say that buying shares in a diversified mutual fund is a good kind of financial act, and playing high-stakes roulette at the local casino is a bad type of financial act—even though (a) the mutual fund might tank and the roulette might pay off, *and* even though (b) money from a profitable mutual fund might finance dastardly schemes, while money lost at the casino might ultimately save starving children. According to my proposal, epistemology is another special case of such utility-based generalizations. The utilitarian epistemologist can say that deduction is a good type of cognitive act, and wishful thinking a bad one—even though (a) deduction might get a false belief and wishful thinking a true one, *and* even though (b) a deduced true belief might bring misery and a wishfully thought false belief might alleviate it.

So you might agree that in this sense, at least, the utilitarian can rescue a notion of epistemic value. You might be left wondering, though, about the point of such a rescue. Shouldn't the epistemologist simply study contributions to epistemic value, and leave questions about the nature of that value to the ethicists? Predictably, my answer here is “no”, for many central epistemic questions are inextricably bound with ethical ones. I suspect that is because, at its core, epistemology is the study of *good* (or *right*) thinking.<sup>5</sup> I will demonstrate the connections by outlining the implications of utilitarian value for just two current issues in epistemology: the “value problem” for knowledge, and the relation between epistemic justification and epistemic responsibility.

## 2 The value of knowledge

One obvious place of intersection between value theory and epistemology is in the comparatively recent resurrection of a problem from the *Meno*: namely, that of explaining why knowledge is (or seems!) more valuable than mere true belief. After all, as Plato pointed out, both knowledge about the location of Larissa and mere true belief about its location will get you to Larissa. This problem turns out to be surprisingly resil-

<sup>5</sup> Perhaps you protest that epistemology is really the study of knowledge. Yes, that has etymology on its side; but *if* knowledge is the one defining characteristic of good thinking, then my characterization subsumes it. If not though—if knowledge is only part of good thinking, or none at all—my characterization is ecumenical enough for these possibilities too. Surely apostates of the church of knowledge—like [Kvanvig \(2003\)](#), [Stich \(1990\)](#), and [Williams \(1991\)](#)—are still epistemologists?

ient. There have been many attempts at solutions, most notably perhaps the “credit” approach.<sup>6</sup> I will not survey these attempts, nor their problems, here—instead I will merely note that it remains an open problem, and so it is worth canvassing options for answers.

How, then, does the utilitarian epistemologist approach this problem? If the value problem is understood to be the question of why knowledge has more *final* value than mere true belief, the utilitarian’s answer is immediate: knowledge does not have more final value than mere true belief, because both have exactly zero. To the classical utilitarian under consideration only states of welfare are of value; all other value is instrumental.<sup>7</sup>

We might expect the utilitarian to fare better in explaining why knowledge has more *instrumental* value than mere true belief. Tradition is behind this understanding of the question; in fact it seems to be how Plato heard it, since he answers his question with the assertion “knowledge exceeds true belief by its binding ties.”<sup>8</sup> Knowledge, Plato seems to suggest, is true belief that has been tied down so that it cannot easily escape. It is thus more valuable to us in the same way a valuable statue in an alarmed display case is more valuable than that same statue left unguarded among thieves.

This answer is not satisfactory, though; as Timothy Williamson aptly put it, “surely [Plato] recognized that mere true beliefs can be held with dogmatic confidence, and knowledge lost through forgetting.”<sup>9</sup> That is, Plato’s answer does not explain why knowledge is more valuable than a lucky true belief held dogmatically, since both are apt to stick around. It also fails to explain why knowledge we are about to forget is more valuable than mere true belief we are about to forget, given that both are about to disappear.

<sup>6</sup> Greco (2003), Riggs (2002), Sosa (2003), Zagzebski (2003), et al.

<sup>7</sup> There are complications hereabout whether knowledge can help *constitute* welfare; I here assume it cannot. As a quick gesture in defense of this assumption, consider knowledge of utter trivia, such as of “Dennett’s lost sock center: the point defined as the center of the smallest sphere that can be inscribed around all the socks [he has] ever lost in [his] life” (Dennett 1991, p. 28). Would such knowledge plausibly increase one’s welfare? If not, what could explain why some knowledge constitutes welfare and other knowledge does not, without referring to extrinsic aspects of that knowledge?

Pritchard (2011) suggests that we would, at least, choose knowledge of any arbitrary proposition over mere true belief in the same proposition when given the chance—even under the stipulation that there is no “practical benefit to choosing one option over the other.” Given some kind of preference-based account, this would be evidence that knowledge does help constitute welfare. Consider a comparable offer, though, between two lives stipulated to go equally well overall, but in one you have more friends. Initial intuitions might tempt us to the one with more friends, but on reflection this should look like a mere framing effect, and not a proper preference.

Still, some in my experience find my assumption implausible given cases like the blissfully ignorant person, who is subjectively happy because unaware that she is in fact faring poorly. It may seem that the missing piece for improving her welfare is the knowledge of her wretched state. ‘Improving’ here is ambiguous, though. The claim is probably correct on an instrumental reading of the word, since the knowledge of her poor state will usually help her to fix it. I think it is not correct on a constitutive reading, though, where the mere knowledge alone improves her state. If we stipulate that knowledge of the actual state of affairs can have absolutely no effect on improving the situation, then I think a classical utilitarian should say—not too implausibly—that her welfare is actually worse off with the knowledge (due simply to the subjectively worse state).

<sup>8</sup> Plato (*ca.* 380 BCE), 98a.

<sup>9</sup> Williamson (2000, p. 78).

**Table 1** The financial analogy

Epistemology	Finance
True belief	Profit
Knowledge	Earnings
Mere true belief	Windfall

Williamson amends this answer to one Platonic in spirit; in summary, Williamson says knowledge is more (instrumentally) valuable because the justification that comes with knowledge makes it more probable that the knower will keep the true belief when new doubts arise. For example, suppose the mere true belief is a Gettiered one based on false premises.<sup>10</sup> Knowledge is more valuable than that, on this story, because the Gettiered belief is more vulnerable to having its justification undermined. John Kvanvig argues that Williamson's answer is also unsatisfactory, however, for in many ways knowledge is significantly more fragile than the corresponding mere true belief. You merely need a defeater to lose the former—perhaps just thinking of skeptical scenarios will do!—while something must change your mind to lose the latter. *Perhaps* the world is such that you are on balance more likely to keep knowledge than mere true belief, but this would be at best a contingent thesis, and so could hardly capture the intuition that it is in the nature of knowledge to be more valuable than mere true belief.<sup>11</sup>

I think Kvanvig's point here generalizes; any account that explains how knowledge is better than mere true belief by adverting to instrumental value is going to be hostage to contingencies. Any piece of knowledge could, in the right circumstances, be far less valuable than the corresponding mere true belief, and even generalizations about how knowledge is *usually* more valuable will depend on the local causal tendencies. It looks like instrumental approaches to explaining knowledge's value will at best result in empirical hypotheses. Such answers do not seem, on reflection, to be in the spirit of the question.

To see the problem through the lens of utilitarian epistemic value, let us develop the financial analogy a bit further (Table 1). Holding true beliefs as analogous to money, the gaining of a true belief is analogous to profit. Leaning on the common idea that knowledge is true belief that was not "lucky",<sup>12</sup> we can then say knowledge is like *earnings*—that is, money gained through sound investments. Mere true belief, on the other hand, is like a financial *windfall*—money gained by luck.

To ask "why is knowledge of more instrumental value than mere true belief?" is, on this picture, like asking "why are earned profits of more instrumental value than monetary windfalls?" The answer to the financial version of this question is clearly that the earnings are *not* more valuable. By analogy, then, neither is knowledge. The epistemic utilitarian embraces this conclusion and denies the intuition that knowledge is better

<sup>10</sup> Obligatory citation: Gettier (1963).

<sup>11</sup> Kvanvig (2003), Chap. 1.

<sup>12</sup> Duncan Pritchard takes this to be a consensus and a "platitude" (Pritchard 2005, p. 1) citing cases like that in Steup (2006); at any rate, whether platitudinous or not, it is at least a good start.

than mere true belief, even on the instrumental version of the value question. The reason, illustrated by the analogy, is fairly simple: like anything but welfare, epistemic states are at best of instrumental value, and (as we noted earlier) *generalizations about instrumental value only make sense under uncertainty*. Generally charity is more valuable than murder, but to the classical utilitarian (and to the classical utilitarian alone) it is not sensible to ask “why is a charitable act more valuable than a murder that results in the same amount of utility?” To assume there is an answer here begs the question against the utilitarian. The same goes, one step down the instrumental chain, for the question “why are earnings more valuable than windfalls?” Under uncertainty, investments with high expected monetary value are in an important, instrumental sense more valuable than those with poor expected monetary value, but this question builds in the assumption that both result in the same monetary value (given of course that all else is equal). Finally, the same goes for knowledge and lucky-but-true belief; in the description of the case, both have gotten the relevant epistemic (instrumental) good. To stipulate that, despite the odds, luck-sensitive belief formation nonetheless resulted in a true belief is just like stipulating that the murder under consideration ended up net benefitting people, or that the stupid casino bet ended up paying off.

We do have the stubborn intuition that knowledge is more valuable than mere true belief, though, and the utilitarian needs to accommodate this intuition in some way. Fortunately, such a task is familiar ground for the utilitarian; she already has to explain away intuitions about the theoretical possibility of justified torture and such. She can simply borrow the typical response strategy already available for such cases. It runs roughly like this: because we are finite creatures who cannot be arbitrarily sensitive to the infinite vagaries of each situation, we rely on heuristics to guide us—rules of thumb about what generally is of use.<sup>13</sup> These become ingrained in us so deeply that we are tempted by deontological intuitions to the effect that an act can be wrong even if it makes the world better.

Alvin Goldman makes a comparable response to mine in his half of [Goldman and Olsson \(2009\)](#). Like me, he suggests that we may be confusing the instrumental value of a reliable process with greater final value of the product, and like me he alludes to money as an analogous case. The difference, though, is that Goldman suggests that knowledge actually does thereby get more value, through a process of “value autonomization”. This is a psychological thesis about how we come to *attribute* independent value to otherwise instrumentally valuable things, combined with the view that such intuitive attribution is a guide to what is *actually* of value. In the cases at hand this latter claim is implausible, though; we surely do not think such intuitions are a guide to what is actually of final value when it comes to money, for example. And though I am sympathetic to letting intuitions be a defeasible starting place for theorizing, they generally no longer have the same evidential weight when we have an undermining story about their source—just as stubborn intuitions about which line is longer in the Müller-Lyer illusion do not remain persuasive after learning the nature of the illusion.

Unlike Goldman, pure utilitarians can simply let go of any attempt to make sense of knowledge as of more final value, and use the typical utilitarian story—one to which

---

<sup>13</sup> See Harman (1977, p. 155).

they are already committed—to do so. Knowledge is not more valuable than mere true belief, just as earnings are not more valuable than windfalls. In both cases the former *seem* more valuable because it makes sense to seek them, while seeking the latter is at best unwise and at worst incoherent.

### 3 Justification and epistemic responsibility

The theory of value behind an epistemology also has implications for the nature of epistemic justification. Two central characteristics of epistemic justification are that (a) it is a positive evaluation of a belief, and that (b) this positive evaluation can obtain even when the resulting belief is false. One common and natural way to capture both these aspects is to say that justified beliefs are those that have been *responsibly* formed. Justification is then a positive evaluation because it represents epistemic responsibility on the part of the thinker, and yet we know that despite such responsibility, an uncooperative environment can still cause a thinker to fail to get knowledge.

The typical understanding of “responsibility” here involves deliberative cognition of some kind, but this construal has at least two serious problems. First, such responsibility seems to many unnecessary for justification, since many are willing to make justification-like positive evaluations of beliefs even when they clearly were not under the thinker’s deliberate control. It is hard to say the thinker is “responsible” in this deliberative sense for beliefs that result from straightforward perceptual mechanisms, for example, but such beliefs (even when false) seem importantly well-formed in a way that at least looks like typical epistemic justification. In effect, taking justification as this kind of epistemic responsibility seems to beg the question against justification externalists, who (roughly speaking) deny that justification supervenes on deliberative cognition. Second, such a notion of epistemic responsibility seems to require a robust doxastic voluntarism. If there is no such thing, as many believe, then deliberative epistemic responsibility (and thus justification, here) is not even possible.

These problems should sound like special cases of more general problems in the intersection of action theory and normative ethics. Is responsible choice in intention what is essential to positive evaluation? Is there such a thing as a choice for which we can be (truly) responsible in the first place? The Kantian deontologist must answer both of these positively, but the utilitarian need not. This is especially good for the epistemic utilitarian, since the conceptual coherence of free will faces tougher challenges than usual at the doxastic level.<sup>14</sup> Traditional libertarian views will ascribe some separate form of causation to agents from the natural causation to which we are accustomed, and as applied in epistemology this mysterious causation would foreclose on any scientific account of how we form beliefs. It is far too early in the science side of the game for such foreclosures, and indeed given how far the science already has come, such foreclosures would look entirely arbitrary.<sup>15</sup> It is possible a compatibilist

<sup>14</sup> *The locus classicus* on doxastic voluntarism as an objection to deontic epistemology is Alston (1988). For a collection on such issues, see Steup (2001).

<sup>15</sup> The naturalistic libertarian picture in Balaguer (2004), on the other hand, has problems explaining how we can be responsible for genuinely random occurrences in the brain.



version of doxastic voluntarism would be sufficient for Kantian epistemic responsibility, but compatibilism, too, has worse than usual problems at the cognitive level. If we explain epistemic agency in terms of causal processes, our inclination to epistemic blame is even more likely to balk than in the ordinary action cases. A compatibilist approach to epistemic agency will require a close look at the causal mechanisms at the sub-propositional level that underwrite the formation of belief. These are not likely to be susceptible to, say, a Frankfurt-type account of freedom, since it will make little sense in most cases to speak of the agents' desires (of any order) with respect to such fine details of cognition.<sup>16</sup>

Again, at the level of normative ethics the utilitarian has a typical strategy for handling these problems—one that can be carried straightforwardly into the epistemic realm. Since utilitarian value does not depend on the formation of free intentions, the utilitarian can consistently deny the existence of any robust free will, including doxastic freedom. The utilitarian also must deny the existence of what we might call “thick” moral responsibility—the moral responsibility of the retributivist, who claims that it can be just to punish even when net consequences of such punishment are bad.<sup>17</sup> To the utilitarian, punishment and reward are only justified, as anything, by its impact on welfare. The utilitarian thus has what we might call a “thin” notion of moral responsibility. Roughly speaking, for a utilitarian to say someone is *responsible* for an outcome is just to say that the person houses the most appropriate place to apply change in order to bring about more future utility with respect to that outcome. When Squeaky is responsible for murder, this means that efforts to redress current and future such harms are best concentrated on her. Squeaky is not morally responsible in the thick sense for the act; she does not deserve punishment for its own sake. Nonetheless it can be just to subject her to isolation and rehabilitation for the sake of future consequences. To the extent we discover that Charlie *caused* Squeaky to murder, that is the extent to which we discover Charlie is the more effective place to apply such change, and thus in the utilitarian sense we attribute him more of the moral responsibility for the crime.<sup>18</sup>

Usually, people are the most appropriate places to look for assigning utilitarian responsibility. There are two reasons for this. First, people are such inherently complex creatures that they often appear as black boxes when it comes to their own causes. Usually, we cannot trace the causal chain any further back—and even if we could, we would probably find a very scattered lot of small causes. (Note that when the cause external to the person is both transparent and focused, as in brainwashing cases, we then invariably place the responsibility further back onto that cause.) Thus individual people typically serve as a kind of causal bottleneck for events of moral importance.

<sup>16</sup> Frankfurt (1971).

<sup>17</sup> This is, I take it, the same thing as “desert”, or what Galen Strawson calls “true” moral responsibility—“responsibility of such a kind that, if we have it, then it *makes sense*, at least, to suppose that it could be just to punish some of us with ... hell and reward others with ... heaven” (Strawson 1994, p. 216).

<sup>18</sup> Of course this picture of moral responsibility, like all others, has its problems—most notably that it seems to justify scapegoating, should the resulting deterrence prove a net benefit. I am not here attempting to vindicate utilitarianism and its implications, though (except insofar as providing a good theory of epistemic value is in its favor). Utilitarianism is, at least, a going hypothesis worthy of consideration, and thus worthy of consideration as a picture of epistemic value.

The second reason we normally hold a person responsible is that people are literally more response-able than most other potential candidates for change. The very nature of intelligence makes its possessors extremely adaptable to new circumstances.

It should not be surprising, then, that people are also usually the most natural place for assigning *epistemic* responsibility. That is, though sources of risk in belief are ubiquitous and varied, often the most salient way to control for belief risk is at the level of our own belief formation. It is so important that it just might for a long time seem to be the only salient place. This serves as the utilitarian explanation for the source of internalist intuitions—they are from judgements about epistemic risk-management. And just as it was revolutionary in criminal justice to suggest that sometimes it is most useful to blame a crime on the hypnotist, or the abusive parents, or the corrupting peers, or the hormonal imbalance, or the economy, or the government, or some other factor “external” to the criminal, so too was it revolutionary in epistemology to suggest that sometimes it is most useful to place epistemic blame or praise on factors external to the deliberative powers of the thinker, as externalism suggests.

The utilitarian can thus take on the picture of epistemic justification as epistemic responsibility, with the understanding that for a thinker to be epistemically responsible is just to have controlled for risk in belief formation (with respect to epistemic goals) as best as might be expected—in other words, so that no effort to change the thinker for better future (epistemic) results is appropriate.<sup>19</sup> When we see what went wrong in the Gettier cases, for example, we do not feel the thinker needs to amend her epistemic ways. She was epistemically responsible, and thus epistemically justified. This utilitarian version of responsibility still captures the two key principles of justification, since (a) the positive evaluation of epistemic justification comes from the high expected utility of good risk management techniques, and (b) many risks cannot be controlled for, so that a belief responsible in this sense can still be false.

Here is another advantage of this account: unlike views committed to doxastic voluntarism, the utilitarian picture of responsibility does not pixelate as we zoom in to fine-grained cognitive details. In other words, the utilitarian version of justification is naturalistic in at least the sense that it plays nicely with cognitive science. Epistemic agents are more like entire investment banks than individual financiers; there is not one little epistemic homunculus sitting in the brain making all the epistemic investments, but instead an intricate organizational chart of departments and subdivisions, each node and sub-node of which performing specific duties aimed ultimately toward the goal of forming good beliefs. And when a major investment goes south at a financial institution, it is wise to dole out the responsibility at a precise level of functional organization, rather than vaguely gesture at a whole department. As we learn more about the organizational chart of the human brain, we can similarly apportion epistemic responsibility with more accuracy.

---

<sup>19</sup> This point seems to dovetail well with Riggs (2009), which construes epistemic luck as what is under the agent’s control—at least, given this utilitarian understanding of an agent’s “control”.

## 4 Conclusion

The question of knowledge's value and the nature of epistemic responsibility are just two examples of how a background value theory can inform one's epistemology. My hunch is that utilitarian epistemology has other good fruit in its branches. For example, Pritchard (2005) explores implications of the simple hypothesis that *luck* is what makes the essential difference between knowledge and mere true belief. In a later paper he leaves as an open question why luck should be so central to epistemic value:

...I think it is crucial to say more about just why having non-lucky true beliefs is so important to knowledge, and this extension of the project will inevitably lead into issues about epistemic value which are currently at the forefront of discussion.<sup>20</sup>

The utilitarian about knowledge seems to have a simple answer to this question—namely, that to rely on luck is to act on low expected utility, and this is a grave utilitarian sin. We distinguish knowledge from mere true belief for the same reason we distinguish earnings from windfalls, and we seek what excludes luck in epistemology for the same reason we seek what excludes luck in finance: something of high instrumental value is at stake.

I think the utilitarian approach also has potential for discussions of skepticism. On this account to ask whether we ever really have knowledge is like asking whether we ever really have earnings. Aren't all profits ultimately just windfalls, given the countless ways even the apparently safest investment could have gone wrong? The utilitarian is unruffled by such questions; whether we call them windfalls or earnings, mere true belief or knowledge, what is important is the expected utility. What determines the threshold of a "safe enough" epistemic investment is a matter removed from the question of how much effort is worth the marginal return on safety.<sup>21</sup> The utilitarian epistemologist is interested in the latter; like the financier, she will simply seek to act on what has the highest expected utility available, whether or not it counts as "safe enough".

If utilitarian epistemology can cash in on just some of the potential value outlined in this prospectus, then it will merit the commitment of further philosophical investment.

## References

- Alston, W. P. (1988). The deontological conception of epistemic justification. *Philosophical Perspectives*, 2, 257–299.
- Balaguer, M. (2004). A coherent, naturalistic and plausible formulation of libertarian free will. *Noûs*, 38(3), 379–406.

<sup>20</sup> Pritchard (2007, p. 293).

<sup>21</sup> Here I do not mean the narrow, particularly epistemological sense of "safe" (according to which, roughly, a true belief is safe if it could not easily go wrong:  $Bp \Box \rightarrow p$ ). I do not mean it as widely as "low risk", either. It means something like "high expected utility, not forgetting to account for reasonable risk aversion in diminishing marginal returns of utility on the instrumental good in question." Thus a bet of \$100 for a 1 in 10 chance at a profit of \$10,000 might be a "safe" bet, though risky, while a 1 in a million chance at a \$1 billion profit is probably not safe, though of equal expected monetary value.

- Chisholm, R. (1977). *Theory of knowledge* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Dennett, D. C. (1991). Real patterns. *The Journal of Philosophy*, 88(1), 27–51.
- Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68(1), 5–20.
- Gettier, E. (1963). Is justified true belief knowledge?. *Analysis*, 23, 121–123.
- Ginet, C. (1975). The general conditions of knowledge: Justification. In L. M. Alcoff (Ed.), *Epistemology: The big questions*, 2004 edition (pp. 79–89). Malden, MA: Blackwell Publishing.
- Goldman, A. I. (2000). Epistemic folkways and scientific epistemology. In E. Sosa & J. Kim (Eds.), *Epistemology: An anthology* (pp. 438–444). Malden, MA: Blackwell Publishing.
- Goldman, A. I., & Olsson, E. J. (2009). Reliabilism and the value of knowledge. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Epistemic value* (pp. 19–41). Oxford: Oxford University Press.
- Greco, J. (2003). Knowledge as credit for true belief. In M. DePaul, & L. Zagzebski (Eds.), *Intellectual virtue: Perspectives from ethics and epistemology* (pp. 111–134). Oxford: Oxford University Press.
- Harman, G. (1977). *The nature of morality*. Oxford: Oxford University Press.
- Kvanvig, J. L. (2003). *The value of knowledge and the pursuit of understanding*. Cambridge, MA: Cambridge university press.
- Plato (ca. 380 BCE), Meno. In G. R. Crane (Ed.), *The Perseus digital library project*. Somerville, MA: Tufts University. Retrieved April 5, 2010 from <http://www.perseus.tufts.edu/>.
- Pritchard, D. (2005). *Epistemic luck*. Oxford: Oxford University Press.
- Pritchard, D. (2007). Anti-luck epistemology. *Synthese*, 158(3), 277–297.
- Pritchard, D. (2011). What is the swamping problem?. In A. Reisner & A. Steglich-Petersen (Eds.), *Reasons for belief*. Cambridge, MA: Cambridge University Press.
- Riggs, W. D. (2002). Reliability and the value of knowledge. *Philosophy and Phenomenological Research*, 64(1), 79–96.
- Riggs, W. D. (2009). Luck, knowledge, and control. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Epistemic value* (pp. 204–221).
- Sosa, E. (1997). Reflective knowledge in the best circles. *Journal of Philosophy*, 94(8), 410–430.
- Sosa, E. (2003). The place of truth in epistemology. In M. DePaul, & L. Zagzebski (Eds.), *Intellectual virtue: Perspectives from ethics and epistemology* (pp. 155–179).
- Sosa, E. (2007). *A virtue epistemology: Apt belief and reflective knowledge* (Vol. 1). Oxford: Oxford University Press.
- Steup, M. (Ed.). (2001). *Knowledge, truth, and duty: Essays on epistemic justification, responsibility, and virtue*. Oxford: Oxford University Press.
- Steup, M. (2006). The analysis of knowledge. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford, CA: The Metaphysics Research Lab. Retrieved July 3, 2008 from <http://plato.stanford.edu/entries/knowledge-analysis/>.
- Stich, S. (1990). *The fragmentation of reason*. Cambridge, MA: MIT Press.
- Strawson, G. (1994). The impossibility of moral responsibility. In G. Watson (Ed.), *Free will* (2nd ed., pp. 212–228). Oxford: Oxford University Press.
- Williams, M. (1991). *Unnatural doubts*. Cambridge, MA: Blackwell.
- Williamson, T. (2000). *Knowledge and its limits*, 2002 edition. Oxford: Oxford University Press.
- Zagzebski, L. (2003). The search for the source of epistemic good. *Metaphilosophy*, 34(1/2), 12–28.