Can Virtue Reliabilism Explain the Value of Knowledge?

Berit Brogaard

Canadian Journal of Philosophy, Volume 36, Number 3, September 2006, pp. 335-354 (Article)

Published by Canadian Journal of Philosophy
DOI: *https://doi.org/10.1353/cjp.2006.0015*

CANADIAN JOURNAL OF PHILOSOPHY
Volume 36, Number 3, September 2006, pp. 335-354

Can Virtue Reliabilism Explain the Value of Knowledge?

BERIT BROGAARD
University of Missouri-St. Louis
St. Louis, MO 63121
USA

I Introduction

A fundamental intuition about knowledge is that it is more valuable than mere true belief. This intuition is pervasive. We have an almost universal desire to know and nearly no desire to believe the truth accidentally. However, it turns out to be extremely difficult to explain why knowledge is more valuable. Linda Zagzebski and others have called this the 'value problem.'[1] They argue that the value problem is particularly difficult to unravel for generic reliabilism. According to generic reliabilism, knowledge is true belief produced by reliable belief-forming processes or faculties. But, the critics argue, 'the reliability of the source of a belief cannot explain the [value difference] between knowledge and true belief.'[2] For reliably formed beliefs allegedly are valuable only insofar as they tend to be true. So if a belief is already true, then the fact that it is

1 See e.g. M.R. Depaul, 'Is Truth Our Epistemic End?' (Pacific Division APA, 1989); L. Zagzebski, *Virtues of the Mind* (Cambridge: Cambridge University Press 1996), 300-2; 'The Search for the Source of Epistemic Good,' *Metaphilosophy* **34** (2003): 12-28; W. Jones, 'Why Do We Value Knowledge?' *American Philosophical Quarterly* **34** (1997): 423-40; J. Kvanvig, 'Why Should Inquiring Minds Want to Know?' *The Monist* **81** (1998): 426-51; *The Value of Knowledge and the Pursuit of Understanding* (Cambridge: Cambridge University Press 2003); and W.D. Riggs, 'Reliability and the Value of Knowledge,' *Philosophy and Phenomenological Research* **65** (2002): 79-96.

also reliably formed adds no further value to the belief. In general, the good of the product makes the reliability of its source good but the reliability of the source does not add value to the product. The critics, furthermore, believe that even if generic reliabilism could find a reason that a reliable source is independently valuable, this would not solve the value problem, because the value of a cause does not transfer to its effect. If we want to guarantee that a belief-producing source confers value on its outcome, we must shift our focus from the belief alone to the overall state of knowing *p*. For a source can confer value on its effect only if cause and product are internally connected.

Virtue reliabilism is a common response to these difficulties.[3] Virtue reliabilism says that knowledge is true belief produced by one of the agent's intellectual virtues, i.e., one of her enduring and reliable cognitive abilities. The focus shifts from evaluating the belief itself to the state the agent is in when she is responsible for believing something true. Virtue reliabilism, it is said, thus has more resources for addressing the value problem. When a true belief is virtuously produced, the truth of the belief is attributable to the agent as his or her own doing. When it is produced accidentally, its truth is attributable to lucky circumstances. Forming a true belief in a virtuous way is thus more valuable than doing so accidentally, because the agent deserves more credit in the former case. While the extra credit the knower is due does not make the known belief more valuable, it supposedly adds value to the overall state of knowing that *p*. A similar response is apparently unavailable to the generic reliabilist, because (1) she fails to remove focus from the true belief to the overall state of knowing that *p*, and (2) she makes no distinction between the true beliefs that derive from stable and reliable dispositions and those that derive from 'strange and fleeting' mechanisms.

In this paper I argue that the appearance that virtue reliabilism is better equipped to handle the value problem is illusory. More specifically, I argue (1) that to solve the value problem the generic reliabilist need not make a shift from a focus on evaluating the belief itself to a focus on the overall state of knowing that *p*, and (2) that it is far from clear that a principled distinction can be drawn between the reliable belief-forming methods that are grounded in the knower's intellectual virtues, and

---

2 Zagzebski, 'The Search for the Source of Epistemic Good,' 12-13

3 See e.g. John Greco, 'Virtues in Epistemology,' in *The Oxford Handbook of Epistemology*, P.K. Moser, ed. (Oxford/New York: Oxford University Press 2002), 311; W.D. Riggs, 'Reliability and the Value of Knowledge'; L. Zagzebski, 'The Search for the Source of Epistemic Good'; and J. Kvanvig, *The Value of Knowledge*.

those that are not. Without a principled distinction of this sort it cannot be established that the extra value of knowledge derives from the extra credit the knower is due. At the end of the paper, I argue that virtue reliabilism fails to address another side of the value problem, which is that of explaining why knowing p is sometimes more valuable than being justified in believing truly that p. I conclude by considering what it would take for a theory to explain this value difference.

II The Machine-Product Model of Belief

Zagzebski compares the reliability of the source of a belief to the reliability of an espresso maker.[4] This tasty cup of espresso is not made any better by the fact that it comes from a reliable espresso maker. The lesson of Zagzebski's espresso analogy is that a reliable belief-forming mechanism does not automatically confer value on true belief. One reason Zagzebski gives in support of this claim is that a reliable belief-forming mechanism purportedly is valuable only insofar as it is truth-conducive. If truth is all that matters — if we want reliably formed belief only because such beliefs tend to be true — then mere true belief ought be treasured to the same extent as reliably formed true belief.

It may be objected that the overall value of a reliable source derives from the value of all the true beliefs it produces and so will be greater than the value of any true belief it produces on some single occasion.[5] It would follow that a reliable source of truth could explain a difference in value between knowledge and true belief. The objection is amiss, however. To say that a mechanism is truth-conducive is not to say that it actually produces true beliefs but only that it is likely to do so. But if a reliable source is valuable only insofar as it is likely to produce true beliefs, then a reliably produced belief is valuable only insofar as it is likely to be true, and an entity that already has *F* cannot acquire additional value by acquiring the property of *being likely to be F*.

Thus it might be suggested that we look for a value in the source of true belief that is independent of reliability or truth-conduciveness. According to Zagzebski, however, identifying a value in the cause of true belief that is independent of truth-conduciveness is not sufficient to solve the value problem.[6] This is because, she says, a source may be valuable

---

4 'The Search for the Source of Epistemic Good,' 13.

5 See Philip Percival, 'The Pursuit of Epistemic Good,' *Metaphilosophy* **34** (2003): 38.

6 'The Search for the Source of the Epistemic Good,' 14-15.

even if its effect is not. In other words, the value of a source never transfers to its product automatically. So even if we were to show that a reliable belief-forming mechanism had, say, intrinsic value (suppose, for example, that the mechanism is an intrinsically valuable intellectual virtue) this would be insufficient to explain the extra value of knowledge. The problem is not, of course, that an intrinsically valuable mechanism is not a good thing but that the value of the mechanism does not accrue to the known belief itself.

Zagzebski thinks the problem facing generic reliabilism is that it adheres to what she calls the 'machine-product model of belief.'[7] On this model, a known belief is the external product of a good source. But if the product is external to the source, Zagzebski says, then the value of the source does not transfer to the product. So if knowledge is true belief that is the output of a valuable source, then it has no more value than true belief. Zagzebski believes that if knowledge is to have more value than true belief, its source must have an internal connection of the same sort as that between an act and its motive. The state consisting of a virtuous motive and right action has more value than right action alone. Zagzebski therefore suggests that we understand knowledge, not as a state consisting of a known belief, but as a whole consisting of the true belief and its source. If knowledge is such a whole, and the source of the belief has independent value, then knowledge is more valuable than true belief. Virtue reliabilism identifies knowledge with a broader state that comprises not only the known belief but also the intellectual virtues of the agent. So, according to Zagzebski, the extra value of knowledge derives in part from the independent (i.e., intrinsic) value of intellectual virtue.

A similar suggestion has been made by Wayne Riggs.[8] According to Riggs, a person who is causally efficacious in bringing about some valuable product is due some amount of credit for having done so. So, it is natural to think that at least part of the extra value of knowledge derives from the knower's achievement in acquiring it. But, Riggs argues, the value of the knower's achievement contributes to the value of knowledge only if knowledge is regarded as the *state* the agent is in when she is causally responsible for her belief. Since virtue reliabilism shifts our focus from the belief itself to the state the agent is in when she is causally responsible for believing something true, virtue reliabilism can explain (at least some of) the value of knowledge.

---

7 *Ibid.*

8 W.D. Riggs, 'Reliability and the Value of Knowledge,' 95

III Value and External Sources

According to the virtue reliabilists, there are thus two reasons that virtue reliabilism is better equipped than generic reliabilism to address the value problem. One is that generic reliabilism, unlike virtue reliabilism, adheres to the machine-product model of belief. But, the virtue reliabilists reason, we cannot explain what makes knowledge more valuable than true belief if we use the machine-product model of belief, because a valuable source does not automatically confer value on its product. The other is that generic reliabilism, in contrast to virtue reliabilism, takes the source of knowledge to be valuable only because it is reliable. But, the virtue reliabilists insist, truth plus a reliable source of truth cannot explain the value of knowledge.

I will now argue that these two arguments in favor of virtue reliabilism are less than fully convincing. First, while it might be true that a valuable source does not automatically confer value on its product, it is not true that an external source cannot ever confer value on its product. Second, while it might be true that generic reliablism cannot explain the value of knowledge because it holds that that the source of knowledge is valuable only insofar as it is reliable, it is unclear how virtue reliabilism differs from generic reliabilism in this respect. I will discuss the first problem in this section, and the second in the three subsequent sections.

Riggs and Zagzebski both hold that we cannot explain the value of knowledge if we adhere to what Zagzebski calls the 'machine-product model of belief.' So they suggest that we remove focus from the belief alone to the overall state of knowing that *p*. Riggs adds further reason for making such a shift in focus. He doubts that there is any such thing as an item of knowledge, because he cannot think of any intrinsic property of a known belief that is not also a property of an accidentally true belief.[9]

I think the latter reason is amiss. Whatever can be said in favor of making a shift in focus from known belief to the overall state of knowing that *p*, the shift is not needed in order to be able to distinguish true belief from knowledge. If Riggs were right, then by analogy there would be no such thing as a true belief, because there is no intrinsic property of a true belief that is not also a property of a false belief. But on the standard conception of identity, two entities can be intrinsically alike and yet numerically distinct. Moreover, plenty of things exist partly in virtue of their extrinsic properties. A bachelor would not be what he is without

---

9 W.D. Riggs, 'Reliability and the Value of Knowledge,' 95.

being single. Or consider a paperweight, an insurance policy, or a door lock. Neither would be what it is without its extrinsic properties.

I presume the real reason that both Riggs and Zagzebski want us to remove focus from the belief alone to the overall state of knowing that *p* is that they adhere to a Moorean conception of value. On a Moorean conception of value, if two things have the same intrinsic properties, then they are equally valuable. Evidently, if two things have the same intrinsic properties, then they have the same amount of instrumental value.[10] Moreover, they have the same amount of intrinsic value. So, if non-instrumental (or final) value is intrinsic, then intrinsically indistinguishable things are equally valuable. Known belief is intrinsically indistinguishable from accidentally true belief. So, given a Moorean conception of value, known belief is no better than true belief.

Presumably the Moorean account of value is what Riggs and Zagzebski have in mind when they think we must shift our focus from the belief itself to the overall state of knowing that *p*. Unfortunately, the Moorean conception of value is questionable. As Wlodek Rabinowitcz and Toni Roennow-Rasmussen have argued, we often attribute extra value to things that are extrinsically related to something else that we value.[11] They argue that when the extra value of a thing derives from something else we value, the thing is valued for its own sake in spite of the fact that the extra value is not intrinsic or unconditional. For example, we may value Princess Diana's dress more than an exact copy simply because the former but not the latter belonged to Diana. The dress's extra value depends on its relation to Diana. Likewise, we tend to value a tropical wilderness more if it has never been visited by humans. The extra value of the tropical wilderness thus depends on its relation to humans. Since we sometimes assign different quantities of non-instrumental or final value to intrinsically indistinguishable things, final value need not supervene on the intrinsic properties of the thing in question.

If Rabinowitcz and Roennow-Rasmussen are right, then the final value of an object can derive partly from an external source. But then, pace Riggs and Zagzebski, it is not clear that the generic reliabilists' adherence to the machine-product model of belief prevents them from explaining the extra value of knowledge. For the value of a source may transfer to the product, even if the source and the product are not internally connected.

---

10 More cautiously: two things that have the same intrinsic properties have the same amount of instrumental value in the same sort of environment.

11 Wlodek Rabinowicz and Toni Roennow-Rasmussen, 'A Distinction in Value: Intrinsic and for its own sake,' *Proceedings of the Aristotelian Society* **100**, part 1 (1999): 33-49; and 'Tropic of Value,' *Philosophy and Phenomenological Research* **66** (2003): 389-403.

A related idea is suggested by Philip Percival:[12] it is quite obvious, he says, that there is no difference between the value of a tasty cup of espresso produced by a reliable espresso machine and a tasty cup produced by an unreliable machine if the value of an espresso is 'determined by its taste.' Likewise, I will add, there is no difference between the value of Princess Diana's dress and an exact copy of her dress if the value of a dress is determined by its appearance. But, Percival reasons, making a claim about belief parallel to the claim about espresso 'begs the question.' Zagzebski says that if the belief is true, then it makes no difference if it is produced by a reliable or unreliable mechanism. But Percival argues, even if it 'makes no difference' to the value of the belief 'in the point of truth,' it does not follow that it makes no difference to the value of the belief in some other respect.

IV Strange and Fleeting Processes

Thus, Zagzebski and her allies have failed to show that reliabilism is unable to account for the value we typically ascribe to knowledge. What they have shown is that the reliabilists cannot conceive of reliability as being valuable only in virtue of being truth-conducive. However, if the reliabilists can give some reason why reliability is independently valuable, then their work is done. The virtue reliabilists may respond that even if it is true that we do not have to regard knowledge as a state of the agent in order for knowledge to have extra value over mere true belief, the only reason the generic reliabilists have in fact given for the claim that reliability is valuable is that it is truth-conducive. The virtue reliabilists, on the other hand, have apparently shown that reliability (of the right sort) is independently valuable because, on their view, reliably produced true belief of the sort that counts as knowledge is to the agent's credit. Being the product of something independently valuable can add value to the product. For example, being a Picasso painting can add value to a painting. Likewise, the virtue reliabilists might say, being the outcome of an intellectual virtue can add value to true belief.

Virtue reliabilism thus fares better than generic reliabilism, even if the above criticism holds up. Or so the virtue reliabilists will most likely have us think. Whether the remainder of their criticism can be sustained, however, depends on whether there is some other significant difference between virtue reliabilism and generic reliabilism. According to the virtue reliabilists, not all reliable belief-forming mechanisms give rise to

---

12 P. Percival, 'The Pursuit of Epistemic Good,' 33

knowledge. Namely, 'strange and fleeting' ones do not.[13] If the virtue reliabilists are right, i.e., if there is a principled distinction to be made between the reliable mechanisms that are strange and fleeting and those that are stable dispositions, then the remainder of their criticism can be sustained. If no such distinction can be drawn, then virtue reliabilism collapses into a form of generic reliabilism.

John Greco has provided the following example to show how virtue reliabilism and generic reliabilism come apart:

> René thinks he can beat the roulette tables with a system he has devised. Reasoning according to the Gambler's Fallacy, he believes that numbers which have not come up for long strings are more likely to come up next. However, unlike Descartes' demon victim, our René has a demon helper. Acting as a kind of epistemic guardian angel, every time Rene forms a belief that a number will come up next, the demon arranges reality so as to make the belief come out as true. Given the ever present interventions of the helpful demon, René's belief forming process is highly reliable. But this is because the world is made to conform to René's beliefs, rather than because René's beliefs conform to the world.[14]

Though René reasons fallaciously, his beliefs about which numbers will come up next are reliably produced, owing to the demon's intervention. Forming beliefs via the gambler's fallacy is not in general a reliable way of forming beliefs, but forming beliefs via the gambler's fallacy and a demon helper who arranges reality in the right way is. So, the source of René's beliefs is reliable in general. And it is reliable in the particular instances as well. For it is a widely acknowledged that a belief is reliable in the particular instance if it could not easily have been wrong,[15] and in the nearby possible worlds in which René forms his belief via the gambler's fallacy and a demon helper, his belief is true. Thus, if the generic reliabilist requires for knowledge true belief produced by a mechanism that is reliable in general and in the particular instance as well, then the generic reliabilist must admit that René knows which numbers will come up next. Virtue reliabilism is apparently not committed to this consequence. For according to the virtue reliabilist, all in-

---

13 J. Greco, 'Agent Reliabilism,' *Philosophical Perspectives* **13** (1999): 273

14 J. Greco, 'Agent Reliabilism,' 286

15 See e.g. S. Luper, 'The Epistemic Predicament: Knowledge, Nozickian Tracking, and Skepticism,' *Australasian Journal of Philosophy* **62** (1984): 26-60; S. Luper, 'The Causal Indicator Analysis of Knowledge,' *Philosophy and Phenomenological Research* **47** (1987): 563-87; R.M. Sainsbury, 'Easy Possibilities,' *Philosophy and Phenomenological Research* **57** (1997): 907, and E. Sosa, 'How to Defeat Opposition to Moore,' *Philosophical Perspectives* **13** (1999): 141-52.

stances of knowledge result from a stable and reliable disposition that makes up the agent's character. But the source of René's belief is supposedly not sufficiently grounded in 'a stable and reliable disposition that makes up his character.' So, René does not know which numbers will come up next.

The examples are easily multiplied. Consider a variation of Alvin Goldman's barn facsimile case. In the original rendering, S is driving in the country and stops in front of a barn. Unbeknownst to S, S is looking at one of the few real barns in an area spawned with facsimiles. The facsimiles are so realistic that if S had stopped in front of any of them, S would have been tricked into thinking she was looking at a real barn. The standard intuition is that S does not know that she is looking at a real barn, because she could easily have had the same belief while looking at a facsimile. In the variation a guardian angel would blur S's vision if she were to look at a fake barn. Again, it is only natural to think that S does not have the knowledge that she is looking at a barn. However, since the relevant belief-forming mechanism is the faculty of vision extended by a protective device (viz., the guardian angel), S's belief is produced by a mechanism that is reliable in general and in the specific instance as well. The mechanism is reliable in general because most beliefs formed by the same mechanism would be true. In fact, we may suppose that if a person is looking at a barn façade, then the angel will blur her vision, and she will not form the belief that she is looking at a barn. The mechanism is reliable in the particular instance as well because S's belief could not easily have been wrong. Had she looked at a barn façade, then the angel would have blurred her vision, and she would not have formed the belief that she is looking at a barn. So, the generic reliabilist is apparently committed to the somewhat counterintuitive consequence that S knows she is looking at a barn. The virtue reliabilist is not so committed. Of course, in the actual world the agent's belief is produced by her cognitive abilities. But her cognitive abilities do not protect her against false beliefs. Without the epistemic helper her belief that she is looking at a barn could easily have been wrong. So her true belief, it seems, is not sufficiently determined by her own cognitive abilities and powers.

Virtue reliabilism thus appears to be significantly different from generic reliabilism. Since the generic reliabilist admits that any reliable mechanism can be a source of knowledge, a mechanism can be a source of knowledge even when its success owes to an external manipulator rather than to the agent's own abilities. So, she must grant our two agents knowledge. The virtue reliabilist, in contrast, need not grant our two agents knowledge, for the source of their beliefs is not sufficiently grounded in their own cognitive abilities, and for that reason it is ruled out by virtue reliabilism as being a source of knowledge. With this

distinction in place, the virtue-theoretical response to the value problem is straightforward. More value accrues to true belief that derives from the agent's own cognitive abilities, because achieving true belief of this sort is to the agent's credit.

V Virtue Reliabilism on the Cheap

Unfortunately, the examples just considered do not favor virtue reliabilism over generic reliabilism. In Greco's case of the gambler, René has very unreliable faculties when unprotected by a helpful demon, but a demon intervenes and systematically adjusts his beliefs to the facts. The virtue reliabilists claim to be able to explain the alleged fact that René does not know which numbers will come up next. But their explanation is at best incomplete. For they are already committed to the claim that a device that systematically adjusts the agent's belief to the facts can be a source of knowledge. Consider David Lewis's example of the prosthetic eye. According to Lewis:

> A prosthetic eye consists of a miniature television camera mounted in, or on, the front of the head; a computer; and an array of electrodes in the brain. The computer receives input from the camera and sends signals to the electrodes in such a way as to produce visual experience that matches the scene before the eyes. When prosthetic eyes are perfected, the blind will see.[16]

In the case of a prosthetic eye, the scene before the eyes causes matching visual experience by peculiar, non-standard causal processes yet a prosthetic eye can be a means for genuine seeing. Lewis adds:

> some prosthetic eyes are more convincing than others as means for genuine seeing ... it seems better if the computer is surgically implanted rather than carried in a knapsack, but better if it is carried in a knapsack rather than stationary and linked by radio to the camera and electrodes.[17]

Some prosthetic eyes are more believable than others 'as means for genuine seeing.' But once we accept that a prosthetic eye that is incorporated into the subject can help her see, why rule out that prosthetic eyes that are external to the subject can be means for genuine seeing? Lewis agrees:

---

16 D. Lewis, 'Veridical Hallucinations and Prosthetic Vision,' in *Perceptual Knowledge*, J. Dancy, ed. (Oxford: Oxford University Press 1988), 85

17 *Ibid*.

> Why should that matter, once we grant that the standard process is not required? I see no real limits on how a prosthetic eye might work. Even the least convincing cases of prosthetic vision are quite convincing enough.[18]

Whether the means for seeing is natural or artificial, internal or external, incorporated into your head or located safely at home has no bearing on whether you can see. Prosthetic eyes are means for genuine seeing. Given the naturalistic tendencies of virtue reliabilism, it is not clear what could prompt an eschewal of this contention. The virtue reliabilists admit that virtues can be acquired. What's more, they tend to hold that the acquisition and use of the skills need not be under our control.[19] Acquired skills of perception, including those that make use of advanced technology, can yield knowledge. As Greco puts it,

> innate vision can give rise to knowledge if it is reliably accurate. But so can acquired skills of perception and acquired methods of inquiry, including those involving highly specialized training or even advanced technology. So long as such habits are both stable and successful, they make up the kind of character that gives rise to knowledge.[20]

Thus, the virtue reliabilists hold that a belief is sufficiently causally determined by the agent's abilities and powers when the belief is produced by innate abilities or acquired methods of inquiry that are both stable and successful. A prosthetic eye is just another example of such an acquired skill or ability, and so should not be dismissed as a method that can give rise to knowledge.

But supernatural devices of the above sort are close kin to prosthetic eyes. Both sorts of device assist the agents in achieving correct beliefs about their surroundings. In both cases the innate cognitive faculties are unreliable when they are not coupled with the device in question. But they are stable and successful when they are indeed coupled with the device in question. The fact that the belief-forming mechanism is external to the agent does (and should not) matter. What matters is that her belief is right in a neighborhood of worlds not too distant from the actual world. It would be otherwise if all she had were a perfectly veridical belief system. Having the latter is consistent with her beliefs not being the products of reliable belief-forming processes. But in this case the

---

18 *Ibid*.

19 Zagzebski is an exception. She argues that the acquisition and use of our intellectual virtues are always under our voluntary control.

20 J. Greco, 'Agent Reliabilism,' 287

agent does not merely have true beliefs. In close counterfactual situations, she has a perfectly veridical belief system as well.

It is even more evident that the belief-forming device used in the modified barn case can be a source of knowledge. Once it is admitted that acquired methods of inquiry, including those involving advanced technology, can give rise to knowledge, it must be admitted that a belief-forming device coupled to another device that can detect the first's limitations can give rise to knowledge. Suppose, for example, that a device, M1, is reliable within the values 4-7 of a given parameter. Suppose M1 is coupled to M2, which reliably detects the values of the parameter in question, and allows M1 to operate only if they fall within the range for which M1 produces the right output. The coupled mechanism, M1 + M2, is reliable even for values outside the success range for M1, for if a value outside the range had been detected, there would have been no output.[21] Given that the virtue reliabilists admit that acquired methods of inquiry, including those involving advanced technology, can be sources of knowledge, it would be difficult for them to deny that the coupled mechanism, M1 + M2, can be a source of knowledge. But if the virtue reliabilists admit M1 + M2 can give rise to knowledge, then they must also admit that the coupled device consisting of the agent's cognitive faculties and the epistemic helper can give rise to knowledge. For the latter device is not significantly different from the coupled device consisting of M1 and M2. The agent's cognitive faculties are coupled with a supernatural device, which reliably detects the range for which the faculties produce the right output and allows them to operate only if they are successful. The cognitive faculties coupled with the supernatural device are thus reliable for values that fall outside the range for which the faculties produce the right output. So, the allegedly strange mechanism is not all that strange after all. Since it is both stable and successful in the same way that the coupled mechanism, M1 + M2, is stable and successful, it makes up the kind of mechanism that gives rise to knowledge.

In short, the virtue reliabilists do not succeed in drawing a principled distinction between the sources of belief that are grounded in the agent's virtuous abilities and those that are not, because the sources that allegedly are ruled out by virtue reliabilism as being sources of knowledge

---

21 This example is due, near enough, to Mark Sainsbury. See R.M. Sainsbury, 'Easy Possibilities,' *Philosophy and Phenomenological Research* **57** (1997): 911. Sainsbury uses this example to show that features of our situation (so-called hidden snags) that may defeat knowledge by shrinking the range of our reliability need to be actual and not merely possible. I am using the example here in my own way.

because they are insufficiently grounded in the agent's abilities are not very different from a wide range of 'acquired methods of inquiry' that are not ruled out by virtue reliabilism as being sources of knowledge.

VI Achieving To Some Degree

It may be true that in most cases of knowledge the truth of the belief is attributable to the agent as her own doing, but the boundary of the class of known beliefs does not run precisely along the boundary of the class of true belief attributable to the agent. We might say that the distinction between the reliable methods that are strange and fleeting and those that are stable dispositions is one of degree, not kind. To the extent that it is one of degree and not kind, there will be a whole range of cases which varies when it comes to how causally responsible the agent is for her true belief. There might be cases in which the truth of the belief is definitely attributable to the agent's own abilities and powers, and cases in which it is definitely attributable to something other than the agent's own abilities and powers. But the virtue reliabilist cannot say that only those cases in which the truth of the belief in question is definitely attributable to the agent's own abilities count as knowledge. The virtue reliabilist must count a range of borderline cases as knowledge as well. And any attempt to draw a distinction between the cases in which the agent is sufficiently causally responsible and the cases in which the agent is not will be ad hoc. Without a clear distinction between the reliable methods that are strange and fleeting and those that are not, there is no clear distinction between generic reliabilism and virtue reliabilism. Thus, virtue reliabilism does not have the advantage over generic reliabilism that it alone can explain (some of) the extra value of knowledge. Either they both can, or neither can. If they can both explain (some of) the extra value of knowledge, then it may be in the following way: in some cases, knowledge is very valuable because the agent deserves a lot of credit. In other cases, knowledge is scarcely more valuable than true belief because the agent does not deserve much credit at all. Maybe this suggestion does have something to recommend it. It apparently accommodates the intuition that in the case of Greco's René, the knowledge he has (if any) is not very valuable and not really something we strongly desire to have. However, my point still stands: the virtue reliabilist is not so radically different from generic reliabilism that this explanation is reserved for virtue reliabilism alone.

VII The Other Side of the Value Problem

The value problem is usually taken to be the problem of explaining why knowledge is more valuable than mere true belief. But there is another side to the value problem, which has received less attention, namely, the problem of explaining why knowledge is more valuable than mere justified true belief. This may not seem to be a problem for an externalist theory, like virtue reliabilism, because externalist theories have traditionally avoided all talk of justification. However, virtue reliabilists have recently drawn our attention once again to the importance of justification. Consider the following well-known case. S lives in an evil demon world. Internally speaking, S's cognitive faculties are in as good working order as ours. Furthermore, some of her beliefs are also true. Still, she fails to know, since she is the victim of a massive deception. Though S fails to know, it is natural to think that she is justified nonetheless.[22] Virtue reliabilists tend to explain the intuition in the following way: the victim is justified because her cognitive faculties count as intellectual virtues relative to our environment.[23] They count as virtues relative to our environment because they are generally reliable relative to our environment. But S fails to have knowledge even when her beliefs are true and justified, because her cognitive faculties fail to work properly in the particular instances.

The example just considered suggests that the virtue reliabilist is faced with the problem of explaining not only the extra value of knowledge over mere true belief but also the extra value of knowledge over mere justified true belief. For the virtue reliabilist, the problem is that of explaining why knowledge is more valuable than true belief produced by a mechanism which is reliable in our normal situation for forming belief but which fails in the particular instance.

Even if there were an obvious way to make out the knowledge-as-credit thesis, the thesis will not help the virtue reliabilist explain why knowledge is more valuable than mere justified true belief. When a true belief derives from a source that fails to be reliable in the particular instance, the belief could easily have been wrong. But whether or not a true belief possesses this property has no bearing on whether the belief

---

22 This problem is due to Keith Lehrer and Stewart Cohen. See their 'Justification, Truth and Coherence,' *Synthese* **55** (1983): 191-208.

23 See Ernest Sosa, 'Intellectual Virtue in Perspective,' in *Knowledge in Perspective: Collected Essays in Epistemology* (Cambridge: Cambridge University Press 1991), 288ff.

is to the agent's credit. For, as the above example suggests, it is never solely in virtue of the agent's cognitive abilities that her true belief possesses this property. Unless the agent is placed in a suitable environment, her cognitive abilities will not produce any true beliefs.[24] We think the person who lives in the evil demon world is justified in believing what she does, because her cognitive abilities would have been in as good working order as ours, had she only been in the right environment. But then her cognitive abilities are as decent as ours, and any of her cognitive achievements are as admirable as any of ours. The fact that she didn't get to the truth on her own is not her own fault and does not prove her cognitive abilities defective. Equally, the fact that we *do* get to the truth on our own is attributable just as much to our environment as to our cognitive abilities. So if the truth of our beliefs is attributable to us, then the truth of the victim's beliefs is attributable to her as well. For she does not differ from us in terms of her cognitive abilities, but only in terms of her emplacement.

The virtue reliabilist may attempt to salvage her position by stipulating that only beliefs produced by mechanisms that work properly in general and in the specific instances are virtuously produced. Since beliefs produced by mechanisms that fail in the specific instances are not virtuously produced, the agent deserves less credit for forming them. But this stipulation is squarely at odds with widely received views about virtue. Consider two variations of Keith Lehrer and Tom Paxson's well-known Grabit case.[25] S sees her acquaintance, Tom Grabit, steal a book from the library. But unsuspected by S, Tom's mother has said that Tom has a doppelganger who is indistinguishable from Tom and who was in the library at the time. In the first case, Tom's mother is lying: Tom has no doppelganger (or at least not one who was in the library at the time). Since Tom's mother is lying, let's suppose S's true belief that Tom stole the book could not easily have been wrong. In the second case, Tom's mother is telling the truth: Tom's doppelganger was in the library at the time. Since S could easily have been looking at Tom's doppelganger without being able to tell the difference, her belief that Tom stole the book could easily have been false. However, it would certainly be odd to say that S is more virtuous when Tom's mother is lying than when she is telling the truth. A person's belief that she is seeing her acquaintance can be virtuous even if she is unable to rule out the possibility that

---

24 A point urged vigorously by Jonathan Kvanvig. See *The Value of Knowledge*, 180ff.

25 K. Lehrer and T. Paxson, 'Knowledge: Undefeated Justified True Belief,' *Journal of Philosophy* **66** (1969), 228

she is looking at her acquaintance's doppelganger. Otherwise, very few of us would ever be virtuous. It seems that the virtue reliabilists must admit that S can believe out of intellectual virtue — that S's believing the truth can be to her credit — even if she is to some extent lucky that her belief is true. On this proposal, the truth of the belief in question cannot be to the agent's credit if the agent does not form her belief via a mechanism that is reliable in our normal situation for forming beliefs, but it can be to her credit if the mechanism fails in the particular instance. But if an agent may deserve credit for her true belief when the belief is justified but fails to be adequately grounded, then virtue reliabilism does not have the resources to explain why knowledge is more valuable than mere justified true belief.

VIII The Extra Value of Knowledge Over Mere Justified True Belief

I suggested above that the knowledge-as-credit thesis might be able to explain why we value some instances of knowledge more than some instances of true belief, even if the explanation would not be reserved for virtue reliabilism alone. On this proposal, some instances of knowledge would be more valuable than some instances of true belief because the agent's achievement is more admirable in virtue of her greater responsibility for it. Still, we have been given no reason to believe that reliabilism under any name can respond to the other side of the value problem. What is it about knowledge compared to justified true belief that causes us to hold it in such high regard? It is not that the agent who has knowledge is due a certain amount of credit than the agent whose belief is merely true and justified is not. For without a suitable environment none of our cognitive abilities are truth-conducive. So, even if it is solely in virtue of our cognitive abilities that we arrive at justified true belief, it is never entirely in virtue of our cognitive abilities that we arrive at knowledge.

Jonathan Kvanvig has argued that we tend to hold knowledge in such high regard because of the kind of cognitive handle one allegedly has on reality when one knows something.[26] We want to avoid accidentally true beliefs because they show that we lack an accurate picture of the interrelationship between things. Having only accidentally justified true

---

26 See J. Kvanvig, 'Why Should Inquiring Minds Want to Know?' and *The Value of Knowledge*, ch. 8.

beliefs bars such understanding of the explanatory connections in nature. However, Kvanvig concludes that we legitimately desire or value something that is confused with knowledge. Knowledge, he argues, does not require explanatory understanding, and explanatory understanding does not require knowledge. So what we really ought to desire or value is not knowledge but understanding. Kvanvig draws the lesson that since it is not really knowledge we want but something else, knowledge is really not all that important.

Kvanvig's proposal is appealing. But, I will now argue, the claim that knowledge does not require explanatory understanding is true only if internal justification is not required for knowledge. As we will see, if internal justification *is* required for knowledge, then knowledge does indeed require explanatory understanding. My lesson will be as follows: generic reliabilism and virtue reliabilism both fall short of explaining all of the value of knowledge, but this should not lead us to think that knowledge does not have the value usually attributed to it. Before drawing the unduly pessimistic conclusion that knowledge is not really all that important, we must consider other explanations of what it is about knowledge compared to justified true belief that causes us to hold it in such high regard. As I will argue, we tend to hold knowledge in such high regard because when we are justified in believing truly that *p*, but we fail to know that *p*, we lack an adequate understanding of the explanatory connections in the world.

Let me begin my defense of this thesis with an argument for the following claim: anyone who is (internally) justified in believing that *p* will believe (at least dispositionally) that her evidence is a reliable indicator of *p*. For example, if you believe it is raining because water is pouring down outside your window, you believe (at least dispositionally) that the water pouring down outside your window is a reliable indicator of rain. The argument is straightforward. If you are justified in believing *p*, then you believe *p* on the basis of the evidence you have for *p*. If, for example, you have evidence for *p*, and believe that *p*, but believe that the evidence indicates that *q*, then we would not say that you are justified in believing that *p*. But you believe *p on the basis of* evidence for *p* only if you believe (at least dispositionally) that the evidence is a reliable indicator of *p*, i.e. only if you believe that the evidence would not be present unless *p* were the case.

I say 'dispositionally' because many of our beliefs are dispositional in nature. For example, until you read this sentence you were most likely not consciously entertaining the thought that elephants are bigger than spiders. But it wouldn't be unreasonable to say that you had the belief nonetheless. Likewise, we do not typically have conscious thoughts about our evidence. If water were pouring down outside your window, you would most likely form the belief that it is raining without any

conscious thought about what explains what. However, you would most likely insist, if asked, that water would not be pouring down outside your window if it were not raining. But the latter is virtually the same as believing that the water outside your window is a reliable indicator of rain. The picture about to unfold is thus quite plausible psychologically speaking. While we do not typically have conscious thoughts about our evidence, we do quite often have dispositional beliefs about it.

Now, your evidence for *p* is a reliable indicator of *p* only if you could not easily have had the same evidence and a false belief that *p*.[27] Suppose, for example, that you believe Tom Grabit stole a book from the library, because you saw someone who looks like Tom Grabit steal a book from the library. If Tom's kleptomaniac twin brother was in the library at the time in question, then you could easily have had the same evidence but a false belief. So, your evidence is not a reliable indicator of 'Tom Grabit stole the book.' But if you could easily have had the same evidence and a false belief that *p*, then your second-order belief that your evidence is a reliable indicator of *p* is false.

It follows that if you fail to know in the Grabit case, then you fail to have a true belief about the nature of your justification. For even an internalist about justification must require that known beliefs be well grounded. Your belief that *p* is well grounded and so counts as knowledge only if you could not easily have had the same evidence and a false belief that *p*. Hence, if your second-order belief that your evidence is a reliable indicator of *p* is false, you fail to know that p. If you know that *p*, on the other hand, you truly believe that your evidence is a reliable indicator of *p*, that is, you have a true second-order belief which you would not have had, had you been justified in believing truly that *p*, but had failed to know that *p*.

How does this explain the extra value of knowledge? Well, if you are justified in believing truly that *p*, but you fail to know that *p*, then you implicitly believe that your reasons for believing *p* are good reasons — i.e., reasons that make *p* warranted. But the second-order belief about your reasons for believing *p* is false. On the other hand, if you know that *p*, then your second-order belief about your reasons for believing *p* is true. So, if you know that *p*, then you truly believe, not only that p is true, but also that you believe *p* because the evidence in your possession is a reliable indicator of *p*. Hence, if you know that *p*, then you truly believe your belief that p is warranted, and you understand why this is so. As

---

27 See e.g. S. Luper, 'The Causal Indicator Analysis of Knowledge.'

Kvanvig points out, our worst cognitive fears include the fear of being duped and the fear of missing something important.[28] When we justifiably believe but fail to know, our worst cognitive fears are realized. We are missing something important when we have evidence for our beliefs but fail to understand how the evidence makes our beliefs warranted. We are being duped when we have a false belief that our evidence for *p* is a reliable indicator of *p*. Thus, a theory of knowledge that requires that the knower be internally justified appears to be able to explain why knowing that p is more valuable than being merely justified in believing truly that *p*. We aspire to knowledge, on this account, because of the cognitive grasp we have of the explanatory connections among reason, belief and truth when we know something. The extra value of knowledge derives from the significance of believing truly that our beliefs are warranted, and understanding why this is so.

Since an internalist can explain why knowing *p* is more valuable than being merely justified in believing truly that *p*, she can explain why knowledge is more valuable than mere true belief. Knowledge is more valuable than mere true belief for much the same reason that knowing *p* is more valuable than being merely justified in believing truly that *p*; it is more valuable because of the understanding we have of the explanatory connections among reason, truth, and belief when we know something. Having only true beliefs entails a grasp of the facts your beliefs are about, but if you have only feeble reasons for believing as you do, but you believe your reasons make your belief warranted, you have a false second-order belief about the nature of your justification. If you have no reasons at all for believing as you do, you are clearly missing something important. What you are missing is an ability to explain why you believe as you do.

The lesson of the above is this: none of the existing versions of reliabilism has given us any reason why we should care about *knowledge* as opposed to accidentally true belief. A theory of knowledge that requires internal justification, in contrast, can explain what it is about knowledge compared to accidentally true belief that causes us to hold it in such high regard. A theory of this sort thus appears to have more explanatory power than does reliabilism. The question that remains is whether the claim that knowledge is more valuable than accidentally true belief should serve as an adequacy condition on a theory of knowl-

---

28 *The Value of Knowledge*, 202-3.

edge. An affirmative answer to this question may help us settle the debate between internalism and externalism once and for all.[29]

Received December 2003
Revised February 2005

---

29 I would like to thank an anonymous referee, Matt Bell, Phillip Dennis, John Gabriel, Tom Paxson, Duncan Pritchard, Joe Salerno, Barry Smith, Jim Stone, and the participants in a seminar at the University of Missouri, St. Louis, for their helpful comments and/or discussion.