1

Reliabilism and the Value of Knowledge

Alvin I. Goldman and Erik J. Olsson

1. WEAK KNOWLEDGE IS MERE TRUE BELIEF

It is a widely accepted doctrine in epistemology that knowledge has greater value than mere true belief. But although epistemologists regularly pay homage to this doctrine, evidence for it is shaky. Is it based on evidence that ordinary people on the street make evaluative comparisons of knowledge and true belief, and consistently rate the former ahead of the latter? Do they reveal such a preference by some sort of persistent choice behavior? Neither of these scenarios is observed. Rather, epistemologists come to this conclusion because they have some sort of conception or theory of what knowledge is, and they find reasons why people *should* rate knowledge, so understood, ahead of mere true belief. But what if these epistemological theories are wrong? Then the assumption that knowledge is more valuable than true belief might be in trouble. We don't wish to take a firm position against the thesis that knowledge is more valuable than true belief. But we begin this paper by arguing that there is *one* sense of 'know' under which the thesis cannot be right. In particular, there seems to be a sense of 'know' in which it means, simply, 'believe truly.' If this is correct, then knowledge—in this weak sense of the term—cannot be more valuable than true belief. What evidence is there for a weak sense of 'knowledge' in which it is equivalent to 'true belief'?

Knowledge seems to contrast with ignorance. Not only do knowledge and ignorance contrast with one another but they seem to exhaust the alternatives, at least for a specified person and fact. Given a true proposition p, Diane either knows p or is ignorant of it. The same point can be expressed using rough synonyms of 'know.' Diane is either *aware of* (the fact that) p or is ignorant of it. She is either *cognizant of* p or ignorant of it. She either *possesses the information* that p or she is uninformed (ignorant) of it.

To illustrate these suggestions, consider a case discussed by John Hawthorne (2002). If I ask you how many people in the room know that Vienna is the capital of Austria, you will tally up the number of people in the room who possess the information that Vienna is the capital of Austria. Everyone in the room who

Alvin I. Goldman and Erik J. Olsson

possesses the information counts as knowing the fact; everybody else in the room is ignorant of it. It doesn't really matter, in this context, where someone apprised of the information got it. Even if they received the information from somebody they knew wasn't trustworthy, they would still be counted as cognizant of the fact, that is, as knowing it rather than as being unaware of it.

The point can be expressed by the following principle:

(COMPL) $\sim (K_sp) = IGN_sp$

(COMPL) applies only where p is true, or factive. Given the truth of p, it says that ignorance and knowledge are complements of one another, that is, S is ignorant of p if and only if S doesn't know that p. How could this principle hold, however, if knowledge consisted in something more than true belief? Suppose, for example, that knowledge is justified true belief plus an anti-Gettier condition X. Then, assuming the truth of p, S's failure to know p wouldn't imply his being ignorant of p. Instead of being ignorant of p, he might believe p unjustifiedly, or might believe it justifiedly but without fulfilling condition X. So, when p is true, failure to know p in a strong sense of knowing (e.g. JTB + X) would not imply ignorance. The correctness of (COMPL) implies that, at least in one sense, knowing is nothing more than having true belief.

We can illustrate the foregoing argument diagrammatically (Fig. 1.1). If knowledge is something like JTB + X, then the terrain is exhaustively captured by the set of possibilities displayed in the diagram. The complement of knowing is not knowing, but not knowing p (where p is true) can occur in any of three different ways: (1) by being ignorant of p (not believing it), (2) by believing p unjustifiedly, or (3) by believing p justifiedly but violating condition X. Under this concept of knowledge, no inference is licensed from not knowing p to being ignorant of p. We contend, however, that there is a sense of 'knowing' in which this inference is licensed. People commonly make this inference. The
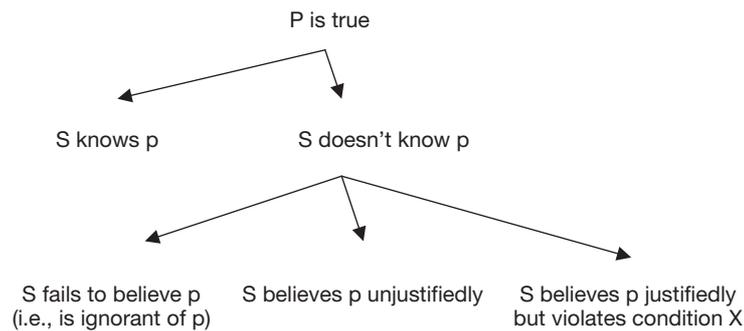

Figure 1.1.

only concept of knowledge compatible with this inference is the one in which knowledge = true belief.

Someone might challenge this conclusion by challenging the claim that ignorance of p (where p is true) is failure to believe p. The challenger might claim that there are three ways to be ignorant of p: by failing to believe it, by believing it unjustifiedly, or by violating condition X. If this were right, (COMPL) wouldn't imply that knowledge is mere true belief. But this claim about the meaning of 'ignorance' is plainly wrong. It is highly inaccurate, inappropriate, and/or misleading to characterize somebody who unjustifiedly believes (the fact that) p as being *ignorant* of p. Similarly, it is highly inaccurate, inappropriate, and/or misleading to characterize somebody who justifiedly believes p but fails to satisfy condition X as being ignorant of p. Thus, the exhaustiveness of the dichotomy between knowledge and ignorance is best explained by the thesis that knowing p (in one sense of 'knowing') is simply believing p where p is true. It does not consist in anything beyond true belief.[1]

If there is a weak sense of 'knowledge' in which it is equivalent to true belief, then the unqualified thesis that knowledge is more valuable than true belief goes by the board. If a state of knowing, in this sense of 'know,' is nothing more than a state of true belief, then neither knowing nor truly believing can be more valuable than the 'other.' However, we do not maintain that weak knowledge is the *only* kind of knowledge, or the only sense of 'know.' In this respect we depart from the radical positions of Isaac Levi (1980) and Crispin Sartwell (1991, 1992), who both hold that 'know' *uniquely* means 'believe truly.' We cheerfully grant that there is a stronger sense of 'know,' which epistemologists have long pursued and which involves more than the two conditions of belief and truth. For this stronger sense of 'know,' the thesis that knowledge is more valuable than true belief is not so easily disputed. In the rest of the paper we shall be concerned with knowledge in its strong sense.

2. THE VALUE OF RELIABILIST KNOWLEDGE

Why is knowledge, in the strong sense, more valuable than mere true belief? The question was first raised in Plato's dialogue *Meno* where it was pointed out that a mere true belief seems instrumentally just as valuable as knowledge. What

---

[1] That there are both weak and strong senses of 'know' is advocated in Goldman (1999). According to Stephen Hetherington (2001), knowledge is a concept that admits of degrees and mere true belief is a minimal kind of knowledge. Similarly, Keith DeRose (2002) says that he is 'tempted' by the contextualist thesis that, in very low-standard contexts, nothing more than true belief is required for knowledge. The identification of knowledge with true belief has been defended in the German philosophical literature by Wolfgang Lenzen (1980), Franz von Kutchera (1982), and, more recently, Ansgar Beckermann (2001). For a critical discussion of Goldman's defense of the weak sense of knowledge, see Le Morvan (2005).

22   *Alvin I. Goldman and Erik J. Olsson*

matters for someone who wants to get to Larissa is to have a *true belief* about
its location. Satisfying the stronger requirement of *knowing* where Larissa is does
not seem to make you any more likely to get there. Still we do believe that
knowledge is somehow better.

The extra-value-of-knowledge (EVOK) problem can be used to test the
adequacy of accounts of knowledge. Suppose an analysis of knowledge is
incompatible with knowledge having an added value. That would be a pretty
strong argument against the adequacy of that analysis. Recently a number
of authors have argued that process reliabilism does not pass the value test.[2]
According to process reliabilism, a subject S knows that p if and only if (1) p
is true, (2) S believes p to be true, (3) S's belief that p was produced through
a reliable process, and (4) a suitable anti-Gettier clause is satisfied. Ward Jones
(1997) raises the value objection as follows:

In short, given the reliabilist's framework, there is no reason why we should care what the
method was which brought about a true belief, as long as it is true. We value the better
method, because we value truth, but that does not tell us why we value the true beliefs
brought about by that method over true beliefs brought about by other less reliable ones.
(1997: 426)

Richard Swinburne (1999) makes essentially the same point:

Now clearly it is a good thing that our beliefs satisfy the reliabilist requirement, for
the fact that they do means that . . . they will probably be true. But, if a given belief of
mine is true, I cannot see that it is any more worth having for satisfying the reliabilist
requirement. So long as the belief is true, the fact that the process which produced it
usually produces true belief does not seem to make that belief any more worth having.
(1999: 58)

Similar arguments are presented by Linda Zagzebski (1996; 2000; 2003),
Wayne Riggs (2002*a*; 2002*b*), Jonathan Kvanvig (2003), and Ernest Sosa (2003).
As Kvanvig points out, the common element of these criticisms of reliabilism is
the identification of the 'swamping effect' that the value of truth seems to have
on the value of reliably acquired belief. Once truth is in place, its value appears
to swamp the value of reliability, thus making the combination of truth and
reliability no more valuable than truth itself. Accordingly, the argument is often
referred to as the 'swamping argument' against reliabilism.

In response to the swamping argument one could point out that few reliabilists
have claimed that knowledge amounts to nothing but true belief reliably pro-
duced. As we have already noted, reliabilists about knowledge usually insist on an
anti-Gettier clause. Adding such a clause opens up the possibility that satisfaction
of that clause is what gives reliabilist knowledge its additional value over mere
true belief. Even if the value of true belief reliably formed does not exceed the
value of true belief *simpliciter*, the value of a true belief reliably formed *in a way*

[2] For statements of process reliabilism, see Goldman (1979; 1986).

that satisfies the anti-Gettier clause could conceivably exceed the value of a mere true belief. Nevertheless, the idea of knowledge depending on the existence of a reliable connection is the central one behind reliabilism, and it would be unfortunate for the theory if that very component failed to produce an added value. In the following we will, with one exception, be concerned with a simple reliabilist theory according to which knowledge requires the satisfaction of only (1)–(3).

The standard conclusion of the swamping argument is that reliabilism must be rejected. This raises the question of whether other accounts of knowledge can solve the value problem. Swinburne and Kvanvig argue that certain internalist theories fare better in this regard. Kvanvig also thinks—as do Sosa, Riggs, and Zagzebski—that virtue epistemology holds special promise when it comes to accounting for the added value of knowledge. The basic idea here is that S knows that p only if S acquires her belief in p by exercising some epistemic virtue and, furthermore, that a person who knows can therefore be credited for his or her true belief in a way in which a person who has a mere true belief cannot.[3] But is it really true that process reliabilism is incompatible with knowledge having an added value? We will, in the course of the paper, explore how process reliabilism can be defended against the challenge posed by the swamping argument.

3. THE SWAMPING ARGUMENT

The standard swamping argument, as endorsed by Jones, Swinburne, and others, runs simply as follows:

(S1)   Knowledge equals reliably produced true belief (simple reliabilism).

(S2)   If a given belief is true, its value will not be raised by the fact that it was reliably produced.

(S3)   Hence: knowledge is no more valuable than unreliably produced true belief.

Since (S3) is a highly counterintuitive conclusion and the argument appears valid, one of the premises must be false. The most common reaction is to reject (S1), that knowledge equals reliably acquired true belief.

Let us take a closer look at the swamping argument. While some theorists, for example, Swinburne, seem to think that this short argument is good as it is, others have tried to present some form of argument for why (S2), the characteristic swamping premise, should be considered true. In Linda Zagzebski's view,

[T]he reliability of the source of a belief cannot explain the difference in value between knowledge and true belief. One reason it cannot do so is that reliability per se has no value

---

[3] Ward Jones (1997) takes the less common position of arguing that the value problem, while problematic, does not show reliabilism to be false.

or disvalue . . . The good of the product makes the reliability of the source that produced it good, but the reliability of the source does not then give the product an additional boost of value . . . If the espresso tastes good, it makes no difference if it comes from an unreliable machine . . . If the belief is true, it makes no difference if it comes from an unreliable belief-producing source.   (2003: 13)

What Zagzebski is saying is that the value of a good espresso is not raised by the fact that it was produced by a reliable espresso machine *if taste is all that matters*; and, likewise, the value of a true belief is not raised by the fact that it was produced through a reliable process *if truth is all that matters*. On this view, (S2) depends for its justification on the following additional premise:

> (Veritism)   All that matters in inquiry is the acquisition of true belief.

Hence, the swamping problem can be seen as arising from combining reliabilism with veritism.[4] Once a true belief is in place, it does not matter whether it was reliably produced, provided attaining true belief is all we strive for in inquiry. Veritism has been advocated within a reliabilist framework by one of the authors of this paper whose theory is also one of the prime targets of swamping theorists.[5]

The standard swamping argument should be distinguished from the swamping argument offered by Kvanvig (2003). Kvanvig too argues that 'ordinary reliabilist theories of knowledge cannot explain the value of knowledge over true belief' (p. 44). His argument runs essentially as follows:

> (K1)   Knowledge equals true belief produced by a process that normally produces true belief (simple reliabilism).
>
> (K2)   Being produced by a process that normally produces true belief just means being likely to be true.
>
> (K3)   The value of having a true belief that is likely to be true is no greater than the value of having a true belief *simpliciter*.
>
> (K4)   Hence: the value of knowledge, reliabilistically construed, is no greater than the value of true belief *simpliciter*.

It is noteworthy that (S2) plays no role in Kvanvig's argument. It is not assumed that if a given belief is true, its value is not enhanced by the fact that it was reliably produced. The crucial premise in Kvanvig's reasoning is rather (K2), which says that being reliably produced just means being likely to be true.[6] The fact that Kvanvig's version of the argument is essentially different from the standard version seems to have gone unnoticed in the literature.

---

[4] For a clear statement of this point, see Percival (2003: 33). As for the role of veritism, see also Jones (1997: 424) and, following him, Riggs (2002*b*: 82).

[5] See e.g. Goldman (2002: 53). Goldman's theory is the explicit target of Jones's (1997: 438) and Riggs's (2002*b*: 80) swamping reasoning. Zagzebski (2003: 14) mentions Alston, Plantinga, Sosa, and Goldman as advocating epistemological theories that are vulnerable to swamping problems.

[6] Kvanvig commits himself to (K2) in the course of his chocolate analogy at 2003: 478.

Reliabilism and the Value of Knowledge   25

Unfortunately for Kvanvig, however, premise (K2) is false. While it is plausible that being produced by a process that normally produces true beliefs implies being likely to be true, the implication does not go in the other direction. Being likely to be true does not imply the existence of a reliable process that produced the belief in question. John may have acquired his belief that he will contract lung cancer from reading tea leaves, an unreliable process, and yet if John is a heavy smoker, his belief may well be likely to be true.

Later in his book, Kvanvig seems to distance himself from (K2). There he equates reliability not simply with (objective) likelihood of truth but with 'objective likelihood derived from the process or methods employed' (p. 49), the suggestion being that reliability is but a 'special kind of objective likelihood' (ibid.). This, he goes on to say, does not save reliabilism from the swamping problem because 'once it is assumed that truth is present, this special kind of objective likelihood has no power to increase the value of the composite beyond that involved in true belief itself' (ibid.). However, Kvanvig's new proposal is not easy to make sense of. While reliability is a feature of a process—roughly speaking, the feature of leading to beliefs that are mostly true—objective likelihood is rather a property of a belief or proposition. Hence, in saying that reliability is but a special kind of objective likelihood Kvanvig seems to commit a category mistake.[7]

Kvanvig's formulation of the swamping problem is also afflicted by another difficulty. Actually, this problem may be shared by other writers' formulations as well, but it is particularly clear in Kvanvig's case. His formulation focuses on the error of allowing a property of an item whose value is *parasitic* on the value of another property of the item to add value to that item. Here is how the argument goes:

If we have a piece of art that is beautiful, its aesthetic value is not enhanced by having as well the property of being likely to be beautiful. For being likely to be beautiful is a valuable property because of its relationship to being beautiful itself. Once beauty is assumed to be present, the property of being likely to be beautiful ceases to contribute any more value to the item in question. Likelihood of beauty has a value parasitic on beauty itself and hence has a value that is swamped by the presence of the latter.   (2003: 45)

Similarly,

. . . [W]hen the value of one property is parasitic on the value of another property in the way that the likelihood of X is parasitic on X itself, the value of the first is swamped by the presence of the second. So even if likelihood of truth is a valuable property for a belief to have, adding that property to a belief already assumed to be true adds no value to the resulting composite that is not already present in true belief itself.   (2003: 45)

---

[7] Consider the following claim: (K2*) Being produced by a process that normally produces true belief entails being likely to be true (unless we have explicit reasons for thinking that the belief is false). This claim is true. Moreover, substituting it for (K2) in Kvanvig's argument makes that argument difficult to separate from the standard swamping argument.

As stated, this argument doesn't work. Here is an example that shows why. Suppose you are offered a choice between options (A) and (B).

(A)   Having one thousand dollars.

(B)   Having one thousand dollars plus having a lottery ticket with a 10 percent chance of winning another thousand dollars.

A swamper who appeals to the principles Kvanvig lays down in the previously quoted passages could argue as follows. Option (B) offers outright possession of a thousand dollars plus a certain probability of acquiring a thousand dollars. But the property of having the chance of acquiring a thousand dollars is parasitic on the property of having a thousand dollars. So the value of this property cannot add to the value of the first property. Thus, option (B) is no more valuable than option (A).

This argument is absurd, of course. Where does it go wrong? It follows Kvanvig's formulation in posing the issue in terms of *properties* and their values, where the value of one property is parasitic on the value of another. The idea can be formulated as follows:

(Property Parasitism)   If the value of property P* is parasitic on the value of property P, then the value of P and P* together does not exceed the value of P.

As the money example demonstrates, however, this cannot be a correct formulation. The point of the principle is to avoid the mistake of *double counting.* If the value of one item is wholly *derived* from the value of a second, we don't want to count the derived value in addition to the original, or *fundamental,* value. Here is another example. Suppose you own a lump of gold, which you keep in a safe-deposit box at a bank. You receive a certificate from the bank that specifies that the contents of the box belong to you, and insures those contents. This certificate, in a sense, has value. But its value is wholly derived from the value of the lump of gold. The certificate doesn't add anything to the value of the box's contents. It would be a double-counting error to suppose that having the certificate doubles the value of what you own. The lottery ticket example, however, is entirely different. The probability of getting the second thousand dollars isn't derivative from possession of the first thousand dollars, because the two quantities of money are distinct and independent.

It's a delicate problem to identify an adequate anti-double-counting principle to replace Property Parasitism. We are inclined to think that it must involve some notion of property *instances* (or property exemplifications, or states of affairs) rather than properties per se. But we shall not try to formulate a more satisfactory principle at this juncture. We simply note that Kvanvig's formulation of the swamping argument appeals to Property Parasitism. He writes:

If we want to answer Plato's question about what makes knowledge more valuable than true belief, it is insufficient to cite a further property of knowledge beyond true belief *even*

if that property is itself valuable. The parts may each have value, but when put together, the whole still may have no more value than if one of the parts were missing altogether. (2003: 48)

When formulated in terms of properties and property composites, the swamping objection hinges on the acceptability of the Property Parasitism principle. But that principle, we have seen, is inadequate. So it isn't clear that a compelling challenge has been mounted against the process reliabilist account of knowledge. Nonetheless, we shall attempt to reply to this challenge. Clearly, there is at least a *potential* problem of double counting of value, even if this problem has yet to be formulated satisfactorily.

In the next two sections, we propose two distinct solutions to the EVOK problem from the vantage point of reliabilism. The solutions are independent, but they also compatible with one another and perhaps complementary. We regard each of the solutions as well motivated, but we recognize that readers may favor one over the other. The authors themselves have different favorites (EJO favors the first solution and AIG favors the second).

Before proceeding, let us back up a bit. The central problem on the table is whether reliabilism can account for the extra value of knowledge as compared with true belief. This problem is common to *all* proffered theories of knowledge. Each must try to explain this extra value. Now, in the case of reliabilism it is generally presupposed that the only way reliabilism can solve the EVOK problem is to say that a reliable process itself has value, of one kind or another. This value can be added to that of the resulting true belief to yield a composite state of affairs (a knowledge state) with more value than the true belief alone. This presupposition about reliabilism's best prospect for solving the EVOK problem is shared by most critics of reliabilism, and perhaps by its proponents as well. If this presupposition is granted, the swamping or double-counting objection immediately comes into play, and supporters of reliabilism are obliged to answer this objection. It is possible, however, for reliabilists to *reject* this presupposition and try to solve the EVOK problem by circumventing the double-counting objection. This approach is adopted in our first solution, which doesn't try to show that the extra value of knowledge is derived from the (token) reliable process that produces the target belief. This solution doesn't solve the swamping or double-counting problem; it just sidesteps it.

4. THE CONDITIONAL PROBABILITY SOLUTION

According to our first solution, if a true belief is produced by a reliable process, the composite state of affairs has a certain property that would be missing if the same true belief weren't so produced. Moreover, this property is a valuable one to have—indeed, an epistemically valuable one. Therefore, *ceteris paribus*,

Alvin I. Goldman and Erik J. Olsson

knowing that p is more valuable than truly believing that p. What is this extra valuable property that distinguishes knowledge from true belief? It is the property of *making it likely* that one's future beliefs of a similar kind will also be true. More precisely, under reliabilism, the probability of having more true belief (of a similar kind) in the future is greater conditional on S's *knowing* that p than conditional on S's *merely truly believing* that p. Let's call this proposed solution to the EVOK problem *the conditional probability* solution. Probability should here be interpreted objectively.

The solution can be illustrated in connection with the espresso example. If a good cup of espresso is produced by a reliable espresso machine, and this machine remains at one's disposal, then the probability that one's next cup of espresso will be good is greater than the probability that the next cup of espresso will be good given that the first good cup was just luckily produced by an unreliable machine. If a reliable coffee machine produces good espresso for you today and remains at your disposal, it can normally produce a good espresso for you tomorrow. The reliable production of one good cup of espresso may or may not stand in the singular-causation relation to any subsequent good cup of espresso. But the reliable production of a good cup of espresso does raise or enhance the probability of a subsequent good cup of espresso. This probability enhancement is a valuable property to have.

The following example shows that reliable production of true belief is no different from reliable production of good espresso when it comes to probability enhancement. Suppose you are driving to Larissa but are at loss as to which turns to take at various crossroads. On the way to Larissa there are two forks. If you choose correctly on both occasions, you will get to Larissa on time. If not, you will be late at best. Your only assistance in forming beliefs about the right ways to turn is the on-board computerized navigation system. We consider two situations differing only in that the navigation system is reliable in Situation 1 and unreliable in Situation 2. We assume that in both cases the navigation system tells you correctly how to turn at the first crossroads. In the first scenario this is to be expected, because the system is reliable. In the second it happens by chance. Suppose the correct information at the first crossroads is 'The best route to Larissa is to the right'. Hence in both situations you believe truly that the road to Larissa is to the right (p) after receiving the information. On the simple reliabilist account of knowledge, you have knowledge that p in Situation 1 but not in Situation 2. This difference also makes Situation 1 a more valuable situation (state of affairs) than Situation 2. The reason is that the conditional probability of getting the correct information at the second crossroads is greater conditional on the navigation system being reliable than conditional on the navigation system being unreliable.

We said that the conditional probability approach bypasses the swamping, or double-counting, problem. How does this transpire? As presented, the conditional probability approach is silent about the value that attaches to the reliable process

per se (as opposed to the value that attaches to the state of affairs of knowing). It is equally silent on the legitimacy of *adding* any value that attaches to that process to the value of the true belief in order to obtain a new value that exceeds that of the true belief. It simply doesn't address these issues. Instead, it looks directly at the composite state consisting of knowing (by means of a reliable process or method) and compares its value to the composite state consisting of truly believing (without arriving at that belief by means of a reliable process). The solution contends that, other things being equal, the former composite state has a valuable property that the latter composite state lacks. Moreover, we might remark, the value of this property is not already contained in the value of the true belief that helps constitute the knowledge state. Thus, there is no way for a critic of reliabilism to re-introduce the swamping problem for the conditional probability approach.

Obviously, the extent to which a knowledge state enhances the conditional probability of future true beliefs depends on a number of empirical regularities. One is that people seldom face unique problems. Once you encounter a problem of a certain type, you are likely to encounter a problem of the same type at some later point. Problems that arise just once in a lifetime are relatively rare. In our navigation example, the question of what is the best turn for driving to Larissa occurs more than once. Another observation is that if a particular method successfully solves a problem once, this method is usually available to you the next time around. In our example, you use the navigation system to solve the problem of what road to take at the first crossroads. This method is also available to you when the same question is raised at the second crossroads. A further empirical fact is that, if you have used a given method before and the result has been unobjectionable, you are likely to use it again on a similar occasion, if it is available. Having invoked the navigation system once without any apparent problems, you have reason to believe that it should work again. Hence, you decide to rely on it also at the second crossroads. Finally, if a given method is reliable in one situation, it is likely to be reliable in other similar situations as well. Let us refer to these four empirical regularities as *non-uniqueness, cross-temporal access, learning*, and *generality*, respectively.

To see what roles these regularities play, suppose S knows that p. By the reliabilist definition of knowledge, there is a reliable method M that was invoked by S so as to produce S's belief that p. By non-uniqueness, it is likely that the same type of problem will arise again for S in the future. By cross-temporal access, the method M is likely to be available to S when this happens. By the learning assumption, S is likely to make use of M again on that occasion. By generality, M is likely to be reliable for solving that similar future problem as well. Since M is reliable, this new application of M is likely to result in a new true belief. Thus the fact that S has knowledge on a given occasion makes it to some extent likely that S will acquire further true beliefs in the future. The degree to which S's knowledge has this value depends on how likely it is that this will happen. This,

in turn, depends on the degree to which the assumptions of non-uniqueness, cross-temporal access, learning, and generality are satisfied in a given case.

Clearly, no corresponding conclusion is forthcoming for unreliably produced true belief. While non-uniqueness and cross-temporal access are usually satisfied quite independently of whether or not the method used is reliable, there is no reason to believe that an unreliable method that yields a correct belief on its first occasion of use will also yield a correct belief on the second occasion. This blocks the step from the availability of the method on the second occasion to the likely production of true belief on that occasion.[8]

On the conditional probability solution, knowledge has its extra value provided that a number of empirical conditions are satisfied. They are the conditions of non-uniqueness, cross-temporal access, learning, and generality. These conditions are normally satisfied, but we would be hard-pressed to claim that those conditions always hold. When they fail to hold, knowledge fails to have an extra value in the present sense. Is this a flaw in the account? Should it follow from an account of the extra value of knowledge that knowledge *always* has this extra value? This is a matter of controversy. Several authors express satisfaction with weaker conclusions. Swinburne (1999: 64) concludes that knowledge has an added value by arguing that it has this value 'almost always.' Williamson (2000: 79) maintains that knowledge is more valuable provided that the cognitive faculties of the knower are in good order, a condition that may occasionally fail to hold. Finally, Percival (2003: 38) thinks that what needs to be shown is that knowledge has added value 'by and large.' Clearly, these authors think that the value problem can be solved without there being a need to show the greater value of knowledge in every single case. For these authors, the conditional probability solution should be deemed satisfactory in this respect. However, other authors disagree. Riggs conceives of the 'value principle' as '[k]nowledge is *always* more valuable than (mere) true belief' (2002*b*: 79, emphasis added). Kvanvig similarly insists that what is to be certified is the '*unqualified* value of knowledge over true belief' (2003: 57, emphasis added). These authors would consider the conditional probability solution incapable of accounting for the whole of the value of knowledge.

---

[8] Bits and pieces of what we have called the conditional probability solution can be found in different places in the literature. It is noted in passing by Armstrong (1973) in his reply to an objection raised by Deutscher. There Armstrong also acknowledges the importance of generality. Williamson (2000: 100–2) presents an account similar to our conditional probability account. However, he focuses on the special case of beliefs with temporally related contents. The approach has rarely been invoked in connection with the swamping problem for reliabilism. For instance, there is no discussion in Kvanvig (2003), which is otherwise well informed by the existing literature. Jones (1997) does attend to the present strategy but fails to see its full import. In Jones's view, the proposal, which he misrepresents as a purely social account of the value of knowledge, can 'only explain the value we place on knowledge in other people' (1997: 430) and is 'of no help in explaining why we value our own knowledge' (ibid.). Our Larissa example shows Jones to be in error. The fact that you have reliabilist knowledge of which way to turn at the first crossroads is better *for you*, the agent, for it makes it more likely that you will acquire further true beliefs in the future.

Still, it is far from clear that our pre-systematic thinking on the matter demands that knowledge always be more valuable than mere true belief. Most generalizations we subscribe to are arguably of a 'typical' rather than an 'absolute' kind. Money is a valuable thing to have, yet rich people are sometimes killed because of their wealth; so for them money was actually something bad. Birds fly, yet birds with feathers covered in oil don't; and so on. The generalizations we make in our daily lives are not universal generalizations in the sense of predicate logic but elastic generic claims that can survive a limited number of counter-instances. If so, why should our claim that knowledge is more valuable than mere true belief be any different?[9]

FN:9

5. VALUE AUTONOMIZATION

The conditional probability solution explains why reliabilist knowledge is *normally* but not *always* more valuable than mere true belief. We have also given some significant reasons for thinking this is all that needs to be shown. Still, some philosophers think that we are *always* prepared to attribute greater value to knowledge than to mere true belief. If this is a general attributional tendency of ours, does reliabilism have the resources to explain why this is so? In this section we try to offer an explanation of this sort. Notice that we have also slightly shifted the question to the question of value *attribution*. A rationale for this shift will be presented later in this section.

The crucial feature of our proposed explanation is an elucidation of the psychological mechanisms whereby reliable belief-forming processes come to be accorded 'autonomous' value. Although the ascribed value of reliable processes is initially derivative from the ascribed value of the true beliefs they cause, reliable processes ultimately acquire autonomous value—value that isn't dependent, on a case-by-case basis, on the value of resultant true beliefs. A subsidiary part of the solution is hinted at by a phrase in the preceding sentence ('on a case-by-case basis'), which indicates that the value of a token reliable process isn't a function solely of that token's own consequences. Rather, the value imputed to a token process is inherited from the value imputed to its type, a possibility that seems to be ignored by the swamping argument. This approach might be labeled 'type-instrumentalism' as opposed to 'token-instrumentalism.'

The swamping argument assumes that, according to reliabilism, the value of a token reliable process derives from the value of the token belief it produces.

---

[9] Cf. Jones (1997: 434): 'I value going to fairs because I have fun when I go to them, even though I can distinctly remember occasions when I got sick on the rides and did not have any fun at all. The fact of my having fun at fairs is responsible for the value I place on fairs, but my having fun is only a contingent property of my attending them. Knowledge is like fairs. We value them both even though we do not always get what we want from them.'

The argument doesn't highlight the type–token contrast, but it hinges on the token interpretation. The argument contends that the value accruing to a token reliable process cannot be *extra* value, because it is wholly traceable to the token belief that it causes. So the value of the process token is analogous to the value of the certificate in our safe-deposit box example, which is wholly traceable to the lump of gold. As in the certificate–gold example, it would be double-counting to add the value of the process token to the value of the true belief token.

Our present solution to the swamping problem, by contrast, involves the claim that instrumental value isn't imputed exclusively because of a singular causal relation between a token instrumental event and a token result. There is a second kind of instrumentalism-based value inheritance. When tokens of type T_1 regularly cause tokens of type T_2, which has independent value, then type T_1 tends to inherit (ascribed) value from type T_2. Furthermore, the inherited value accruing to type T_1 is also assigned or imputed to each token of T_1, whether or not such a token causes a token of T_2.

An example of value inheritance that fits the type-instrumentalist pattern is money. Money (especially paper money) is not the sort of thing that has fundamental value. But since possessing money (in non-trivial amounts) frequently produces events or states of affairs that do have fundamental value (pleasure, satisfaction, etc.), possessing money comes to be viewed as an instrumentally valuable type of state. Furthermore, each *token* of this type inherits instrumental value from the type, even if some tokens don't actually cause events or states of affairs with fundamental value. (Either the money isn't spent at all, or it's spent on things that bring no pleasure, happiness, satisfaction, etc.)

How does type-based value inheritance allow us to solve the swamping problem? Does it allow us to say that whenever a reliable process causes a true belief, the compound state of affairs consisting of the process token and the true belief has more value than the state of affairs consisting of the true belief alone? Consider an analogy. Taking aspirin regularly causes headache relief. Since headache relief is valuable, taking aspirin is instrumentally valuable. Now consider a particular token of headache relief produced by taking aspirin. Does the value associated with taking an aspirin *add anything* to the value of the headache relief? Does the compound state consisting of aspirin taking plus headache relief have more value than headache relief alone? It seems not. Contemplate a possible world in which you undergo headache relief at a certain time through no cause at all, and then contemplate another possible world exactly like the first except that the headache relief is caused by aspirin taking. Is the second possible world more valuable than the first? Given a choice, would you prefer to live your life in the second world rather than the first? Presumably not. Spontaneous headache relief is just as good as headache relief caused by aspirin taking. On analogy with this case, why should the use of a reliable process *add* value to that of a true belief? (Thanks to Dennis Whitcomb for this challenge.)

What these cases suggest is that instrumental value, even of a *type*-derived sort, does not *generally* combine with fundamental value so as to increase overall value. However, to say that it doesn't *generally* increase overall value does not entail that it *never* does so. Under what conditions might it do so?

The main possibility we suggest is that a certain type of state that *initially* has merely (type-) instrumental value eventually acquires *independent,* or *autononomous,* value status. We call such a process *value autonomization.* Consider the relationship between (morally) good actions and good motives. The primary locus of moral value, quite plausibly, is actions, for example, acts of duty fulfillment or acts of altruism. Actions, however, are not the only things regarded as morally good or valuable. We also value good motives, for example, a desire to help others in need or a desire to do one's duty. Why do we value such motives? A straightforward explanation is that such motives regularly bring about corresponding actions, actions which themselves are valuable. It is therefore plausible that there is a pattern of inheritance by which value passes from certain types of actions to corresponding motive types, which regularly produce those actions. Notice that a token motive of an appropriate type is regarded as good or valuable even if it fails to produce a good action. For example, a token motive may not generate any action, because there are conflicting considerations that yield indecision. Or it might produce an action that doesn't really assist the intended party, or isn't really what duty requires. Despite failing to produce good consequences of a standard sort, the token motive is still good or valuable, presumably because such value is inherited from the type of which it is a token. And this value it retains autonomously, even without triggering an independently good action.

Good motives or good intentions are naturally thought of as good *in themselves.* Famously, Kant held that *only* a good will, or a good intention, has fundamental moral value. We don't have to join Kant in holding that *only* a good will has such a status. But it is very plausible that good motives or intentions are *among* the things rated as *independently* good. This is confirmed by intuitive judgments to the effect that a compound state consisting of a good motive and a good action is (morally) *better* than a compound state consisting of the same good action done from a bad (or non-good) motive. Apparently, a good motive's value can be added to the value of a good action, suggesting that a good motive doesn't have merely (type-) instrumental value, in contrast with the aspirin case.

Conjoining the elements of type-instrumentalism and value autonomization, we have the main ingredients for an explanation of the greater (assigned) value of knowledge as compared with true belief. Some wrinkles remain to be ironed out, however. According to process reliabilism, it doesn't suffice to turn true belief into knowledge that a reliable process *precede* the true belief; it is critical that the true belief be *caused* (or causally sustained) by such a process. What we have said thus far, however, does not fully accommodate this. A compound state of affairs consisting in a reliable process followed by a true belief will be more valuable

Alvin I. Goldman and Erik J. Olsson

than the same true belief not preceded by a reliable process, and this is so even if there is no causal relation between the two, and hence no knowledge.[10]

This is not a severe problem. What we need to ensure is that true belief *caused* by a reliable process is more valuable than true belief merely *preceded* by a reliable process. This can be handled in a fairly ad hoc way, though; it isn't essential to a solution to the swamping problem. We can simply note that our valuations are sensitive to causal linkages between suitable pairs of states. For example, good actions that are *caused* by good motives get higher moral marks than good actions that are merely *preceded* by good motives. The valuation of knowledge comports with this pattern.

Let us return now to the central explanatory construct, value autonomization. Value autonomization is a psychological hypothesis, which concerns our practices of *ascribing* or *attributing* value to various states of affairs. Readers may object that the issue posed by the swamping problem is not a problem of value *attribution*. It concerns objective value, not how people *assign* valuational status. The real question is whether a state of affairs with the status of being merely instrumentally good can objectively change its status to being non-instrumentally, or fundamentally, good. That's not something on which human psychology can shed any light.

Here we beg to differ. In claiming that this or that state of affairs has fundamental or extrinsic epistemic value, what evidence do epistemologists have to rely on? Presumably, their own intuitions and those of others. But are these intuitions (intuitional states) wholly a function of the *objective* value status of various states of affairs? Are we methodologically entitled to help ourselves to that assumption about intuitions? We think not. Although we don't wish to deny categorically that intuitions track objective value facts, the epistemological enterprise needn't collapse if that assumption is mistaken. In explaining epistemic intuitions and how they relate to one another, it makes sense to utilize human psychological propensities to form linked patterns of evaluation. Just as understanding the sources of moral intuitions may profit from an understanding of human psychology, so understanding the sources of epistemic intuitions may profit from an understanding of human psychology.[11] The value autonomization hypothesis allows that some states of affairs that at one time are assigned merely instrumental value are 'promoted' to the status of independent, or fundamental, value. If this is right, it is compatible with the hypothesis that such promotion occurs in the case of reliable belief-forming processes. While

---

[10] How could a belief-forming process fail to cause a belief? Simply through some sort of malfunction.

[11] For example, recent studies in psychology and cognitive neuroscience by Joshua Greene and by Jonathan Haidt suggest that moral judgments (or intuitions) are often a product of emotion and affect, even when the respondent considers subtly different hypothetical scenarios. Other studies suggest that moral judgments are influenced by framing effects, a well-documented phenomenon in cognitive psychology. See Greene and Haidt (2002) and Doris and Stich (2006).

many such processes are originally regarded as merely instrumentally valuable to true-belief attainment, they are later upgraded to the status of independent value, thereby accommodating the legitimacy of *adding* their value to that of true-belief outcomes.[12]

Let us now step back from our two solutions and reflect on their general characteristics. Reliabilism is generally considered a form of naturalistic epistemology. What is meant by 'naturalism,' of course, varies widely from writer to writer and from topic to topic. In the context of epistemology at least two kinds of naturalism may be distinguished: *metaphysical* and *psychological* naturalism. Metaphysical naturalism holds that epistemic evaluative facts supervene on natural facts. On this understanding, our conditional probability solution to the EVOK problem is very congenial to metaphysical naturalism. This is because it seeks to explain the value of knowledge by reference to the objective (conditionality) probability of obtaining further true beliefs, given the satisfaction of the four conditions we specified in Section 4.

What does psychological naturalism consist of in epistemology? It's an approach that tries to explain our commonsense epistemic valuations in a scientific fashion, especially by appeal to psychology. The autonomization solution is a proffered explanation of our evaluative practices vis-à-vis knowledge. Although the idea of autonomization hasn't received rigorous empirical support, it's put forward here in the spirit of an empirical treatment of human valuational activities. As such it is obviously congenial to psychological naturalism.

6. OTHER APPROACHES TO THE VALUE PROBLEM

In this final section, we examine selected solutions to the value problem that other writers have offered and explain why ours are preferable. Given the size of the literature, it's impossible to discuss all the work that merits discussion. We highlight those approaches that offer either a sharp contrast with ours or some interesting overlap.

Proponents of virtue epistemology have been in the forefront of emphasizing the challenge to reliabilism posed by the value problem. However, virtue epistemology has many varieties. Keeping things simple, it is instructive to distinguish radical and moderate forms of virtue epistemology (VE). Radical VE tries to distance itself sharply from 'naturalistic' approaches to epistemology such as reliabilism, and models the study of epistemology quite closely on the study of ethics. Moderate VE, by contrast, has closer affiliations with reliabilism; it tends

---

[12] We concede that we don't have a fully detailed story about when value autonomization occurs, that is, under what conditions a state of affairs initially viewed as instrumentally valuable is subsequently upgraded to independently valuable. Notice that such a change may not occur within an individual's ontogeny, but may be more of a historical-cultural transition.

to explain epistemic virtues, to a substantial degree, in terms of truth-getting skills. Moderate VE—sometimes called 'virtue reliabilism'—does not pursue so tight a liaison between ethics and epistemology.

At the most radical end of the spectrum is Zagzebski's approach to VE, which models VE quite closely on virtue ethics. In our view, this makes for an awkward fit, especially where Zagzebski gives excessive emphasis to motivation and love in the theory of knowledge and exaggerates (by our lights) the role of the voluntary in the epistemic domain. Taking aim at the 'machine-product' concept of knowledge that she associates with reliabilism (but never explains very carefully), Zagzebski embraces a much 'loftier' agent-act conception of knowledge and epistemic value that requires credit-worthiness and even admirability. She imposes 'a motivational requirement for getting credit for the truth that involves love of truth' (2003: 19). This strikes us as an unduly heavy burden on many cases of knowledge, especially unreflective knowledge by animals and young children. Much knowledge with which we credit people and animals is of a fairly rudimentary sort, acquired by unmotivated perception or spontaneous recall that operates independently of any 'love of truth'. If a dog remembers where it buried a bone, we readily grant that the dog 'knows' where its bone is, and this piece of knowledge is valuable. But does this bit of knowledge require some sort of canine 'agency'? Must the dog perform epistemic 'acts' out of a love of truth?[13] Although our example concerns a non-human animal, we don't think there's a big difference between humans and dogs when it comes to (this kind of) memory knowledge; nor do we think that our ordinary conception of such knowledge distinguishes the two.

Zagzebski introduces a motivational theme in her treatment of knowledge, but it's not clear that she gets any mileage out of it for solving the EVOK problem. She complains that reliabilism's (alleged) commitment to the 'machine-product' model of knowledge precludes a proper account of the relation between knowledge and true belief because the value of a mere cause cannot be transferred to its effect. A reliable process, she says, must be 'external' to the true belief it causes, whereas a motive can be 'internal' to the agent on which it confers value. Thus, only in the case of a motive can value be transferred to its act. As Philip Percival notes, however, these remarks of Zagzebski's are 'little more than gestures. She gives no guidance as to *how* an "internal" connection between motive and act, or a "part—whole" relationship between act and agent, can result in the value of a motive being transferred to its effect' (2003: 34). Moreover, it is becoming well recognized that an external, or 'extrinsic,' event *can* transfer value to another event or object. A widely cited illustration is Princess Diana's dress, which has more value than an exact duplicate simply because it once belonged to

---

[13] Hilary Kornblith's (2002) treatment of knowledge as a natural phenomenon provides a good antidote to the excesses of Zagzebski's overly intellectualist picture (although Kornblith does not specifically critique VE).

Diana. Having once belonged to Diana is an extrinsic (or external) rather than intrinsic (or internal) property of the dress.[14]

Let us turn now to moderate VE. Some themes in moderate VE are fairly congenial to at least the first approach advanced here. For example, John Greco's (1999; 2000) agent reliabilism emphasizes the stability, or 'non-fleetingness,' of a cognitive skill as essential to knowing. He associates such stability with epistemic virtues and agency, claiming that this requirement goes beyond 'generic' reliabilism. The stability requirement is friendly to our conditional probability solution to the value problem. The greater the stability of a reliable source, the greater the probability that it will be used again in the future in similar cognitive tasks. So this element of moderate VE is congenial to part of our approach.

Notice, however, that stability does not cleanly separate virtue reliabilism from generic reliabilism. Although early forms of process reliabilism (e.g. Goldman 1979; 1986) placed no emphasis on stability, nothing in the spirit of generic reliabilism prevents incorporation of stability into its framework. If the VE idea hadn't made an appearance in the 1990s, nobody would have been surprised if stability had instead surfaced within the ambit of generic reliabilism.[15] On the other hand, Greco hasn't persuaded us that stability, or non-fleetingness, of a cognitive source is a strictly necessary condition for knowing. We can easily generate cases in which knowledge occurs through a fleetingly possessed method or skill, such as cases in which a cognitive skill or method is newly acquired and successfully applied to produce a true belief but then promptly lost through death, stroke, onset of Alzheimer's disease, etc.

Another facet (or family of facets) of virtue reliabilism is its emphasis on agency, attributability, and credit-worthiness. According to Riggs (2002*b*), the value assigned to an epistemic state is a function of the credit deserved by the agent. Such credit is deserved only when the state is arrived at in a sufficiently non-accidental way, a way that constitutes an 'achievement.' This theme has pervaded Ernest Sosa's writing from the early 1990s (Sosa 1991) to his most recent writings. Riggs uses the analogy of two holders of Olympic gold medals: Maude possesses one because she won it at the Olympics, whereas Martin possesses one

---

[14] The example is due to Rabinowicz and Roennow-Rasmussen (1999).

[15] In discussing Greco's requirement of a stable disposition, Berit Brogaard (2006) discusses Greco's (1999) example of a character Rene, who is reliable only through the mediation of an epistemic guardian angel. Since Rene lacks a *stable* disposition or faculty for getting the truth, says Greco, virtue reliabilism denies him knowledge, whereas generic reliabilism would have to concede him knowledge. Brogaard plausibly argues that such examples do not favor virtue reliabilism over generic reliabilism. She introduces David Lewis's (1980) example of prosthetic vision, and compares it to Rene with his guardian angel. Since virtues can be acquired, according to Greco, and needn't be under our control, virtue epistemology should credit a possessor of a prosthetic eye with knowledge derived by the use of such prosthetic vision. But if that is correct, Brogaard implies, why not say the same for Rene? In short, virtue reliabilism doesn't draw a principled distinction between sources of belief grounded in virtuous abilities and those that are not.

because he found it while taking a stroll through the woods. Clearly, Martin doesn't deserve the same degree of credit (if any) as Maude does for having a gold medal. The situation is parallel for two cases of true belief, one acquired by luck versus another acquired by cognitive skill. Only the former deserves epistemic credit. Sosa's (2003) analogy is that of an archer who hits his target by skill, that is, his virtue, versus hitting it via a lucky gust of wind that carries the arrow off its initial path into the target. In the former case, success is attributable to the archer qua agent; in the latter, it isn't so attributable. Analogously, says Sosa, what is (most) valuable in the epistemic sphere is attaining truth by one's own performance rather than by luck or accident.

However, these points seem pretty congenial to generic reliabilism; they are not the special preserve of virtue reliabilism. Even if generic reliabilism doesn't use the same *language* as VE—the language of agency, credit, attributability, etc.—it certainly seeks to exclude luck or accidentality by some permutation of the reliability theme. In particular, reliability theories propose either sensitivity, safety, or the absence of relevant alternatives as forms of non-accidentality required for knowledge. Consider Gettier's (1963) disjunction case, for example. Smith makes an entirely justified sequence of inferences from the evidence that Jones owns a Ford to a belief that Jones actually owns one, and thence to a belief that either Jones owns a Ford or Brown is in Barcelona. The inferences used are highly reliable types of inference. Still, it's only accidental that his final, disjunctive belief turns out to be true. It isn't true because Jones owns a Ford, but because Brown, by sheer coincidence, happens to be in Barcelona. Modal reliabilists try to capture what goes wrong here by invoking either a sensitivity condition, a safety condition, or a relevant alternatives condition. These sorts of conditions are standard tools in reliabilism's toolkit. Although they typically figure as a fourth condition for knowledge rather than a third condition, they are squarely within the spirit of generic reliabilism, and not the special preserve of virtue reliabilism. Virtue reliabilism introduces special *language* to describe these cases, but the nuts and bolts of explaining the conditions usually proceeds in a generic reliabilist fashion, using concepts like belief, truth, and various possible-world permutations.

Moreover, insofar as there is a difference between virtue reliabilism and generic reliabilism, the difference resides in the former trying to specify non-accidentality in terms of the agent's credit-worthy performance (or the like). This added element, however, doesn't really work as an account of non-accidentality that captures knowledge. Consider two cases: a Gettier case like the one just discussed and an exactly similar case in which everything goes smoothly—that is, all of the agent's beliefs are true. The latter case is an instance of knowledge whereas the former isn't. Can this be explained in terms of a difference in credit-worthiness of the agent's performance? No. The agent performs just as well in both cases. He infers the same conclusion from the same justified premises using the same inferential procedures. So, contrary to the claim of virtue reliabilism, we cannot

explain all differences between knowledge and merely (justified) true belief in terms of differences in credit-worthiness of performance.[16]

Finally, some virtue reliabilist theories of value seem to embrace ingredients quite similar to those advanced in the present paper. In particular, Sosa (2003) comes quite close to making such commitments. The details of his theory are not entirely clear to us, so we remain unsure just how close are its commitments to ours.[17] At a minimum our proposals might be important supplements to the ingredients he uses explicitly, so we commend them to Sosa as amplifications of his brand of virtue reliabilism (amplifications that don't really depart, as we see it, from generic reliabilism).

Sosa says that our epistemic evaluations express a preference for attaining truth (true belief) by our own virtuous performance rather than by mere happenstance. Thus, there is such a thing as 'performance value.' What is good about performance value, he says, is to be understood in a truth-connected way: '[W]hat is good about this performance value cannot be understood independently of the fundamental value of true believing, and especially of true believing that hits the mark of truth attributably to the agent' (2003: 177). This suggests that performance value is a kind of derivative value. Since Sosa also says that a performance can have performance value even when it doesn't attain its characteristic end (e.g. when it doesn't occur in a proper environment), he implies that performance value attaches in the first instance to a *type*. Here he is in the neighborhood of our type-instrumentalist account. However, Sosa also proceeds to regard performance value as a kind of *intrinsic value*. He writes, 'we seem plausibly committed to the *intrinsic* value of such intellectual deeds. . . . We want . . . to attain truth by our own performance, which seems a reflectively defensible desire for a good preferable not just extrinsically but intrinsically' (2003: 175, italics in the original). Later he speaks of a '*eudaimonist, intrinsic* value of true believing where the agent hits the mark of truth as his own attributable deed' (2003: 177, emphasis in the original). Somehow—Sosa doesn't tell us how—the performance value of a cognitive performance seems to rise from the level of extrinsic value to the level of intrinsic value. We are not sure if Sosa means this, but we feel that the account would be clearer and more persuasive if it were supplemented by the autonomization story we presented above. Autonomization would explain *why* performance value gets to be an intrinsic type of value, although it isn't initially characterized this way. Whether or not Sosa would endorse the autonomization story, we feel that his virtue account would be more

---

[16]   Dennis Whitcomb (2006) makes a similar criticism of virtue reliabilism.

[17]   The article in question (Sosa 2003) runs through a number of possible positions on the value problem and distinguishes several kinds of epistemic value. A novel type of value, 'praxical' value, is initially introduced as important to a solution to the value problem, but it seems to fade in importance toward the end of the paper. We are not sure we fully capture Sosa's position vintage 2003 here, but it's the one that seems most promising. For a new formulation (on which we won't try to comment), see (Sosa 2007: ch. 4).

40   *Alvin I. Goldman and Erik J. Olsson*

compelling if this story were added. Thus, to the extent that his account is on the right track, it does not differ sharply from the one offered here.

REFERENCES

Armstrong, D. M. (1973). *Belief, Truth and Knowledge*. Cambridge, MA: Cambridge University Press.

Beckermann, A. (2001). 'Zur inkohärenz und Irrelevanz des Wissensbegriffs—Plädoyer für eine neue Agenda in der Erkenntnistheorie,' *Zeitschrift für philosophische Forschung,* 55: 571–93.

Brogaard, B. (2006). 'Can Virtue Reliabilism Explain the Value of Knowledge?', *Canadian Journal of Philosophy,* 36 (3): 335–54.

DeRose, K. (2002). 'Assertion, Knowledge, and Context,' *Philosophical Review,* 111: 167–203.

Doris, J., and Stich, S. P. (2006). 'Moral Psychology: Empirical Approaches,' in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (spring 2008 edition), URL = <http://plato.stanford.edu/entries/moral-psych-emp/>.

Gettier, E. (1963). 'Is Justified True Belief Knowledge?', *Analysis,* 23: 121–3.

Goldman, A. I. (1979). 'What is Justified Belief?', in G. Pappas (ed.), *Justification and Knowledge*; Dordrecht: Reidel. Reprinted in A. Goldman (1992), *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge, MA: MIT Press.

—— (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.

—— (1999). *Knowledge in a Social World*. Oxford: Oxford University Press.

—— (2002). 'The Unity of the Epistemic Virtues,' *Pathways to Knowledge*. New York: Oxford University Press.

Greco, J. (1999). 'Agent Reliabilism,' in J. Tomberlin (ed.), *Philosophical Perspectives,* 13, Atascadero, CA: Ridgeview Press.

—— (2000). *Putting Skeptics in Their Place*. Cambridge, MA: Cambridge University Press.

Greene, J., and Haidt, J. (2002). 'How (and Where) Does Moral Judgment Work?', *Trends in Cognitive Sciences,* 6 (6): 517–23.

Hawthorne, J. (2002). 'Deeply Contingent A Priori Knowledge,' *Philosophy and Phenomenological Research,* 65 (2): 247–69.

Hetherington, S. (2001). *Good Knowledge, Bad Knowledge*. Oxford: Oxford University Press.

Jones, W. E. (1997). 'Why Do We Value Knowledge?', *American Philosophical Quarterly*, 34 (4): 423–39.

Kornblith, H. (2002). *Knowledge and its Place in Nature*. Oxford: Oxford University Press.

Korsgaard, C. M. (1983). 'Two Distinctions in Goodness,' *Philosophical Review,* 92: 169–96.

Kvanvig, J. L. (1998). 'Why Should Inquiring Minds Want to Know? *Meno* Problems and Epistemological Axiology,' *Monist*, 81: 426–51.

—— (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, MA: Cambridge University Press.

Le Morvan, P. (2005). 'Goldman on Knowledge as True Belief,' *Erkenntnis*, 62 (2): 145–55.

Lenzen, W. (1980). *Glauben, Wissen und Wahrscheinlichkeit*. Vienna: Springer-Verlag.

Levi, I. (1980). *The Enterprise of Knowledge*. Cambridge, MA: MIT Press.

Lewis, D. (1980) 'Veridical Hallucination and Prosthetic Vision,' *Australasian Journal of Philosophy*, 58, 239–49●).

Percival, P. (2003). 'The Pursuit of Epistemic Good,' *Metaphilosophy*, 34 (1/2): 29–47.

Pritchard, D. (2007). 'Recent Work on Epistemic Value,' *American Philosophical Quarterly*, 44: 85–110.

Rabinowicz, W., and Roennow-Rasmussen, T. (1999). 'A Distinction in Value: Intrinsic and for its Own Sake,' *Proceedings of the Aristotelian Society*, 100 (1): 33–49.

Riggs, W. D. (2002*a*). 'Beyond Truth and Falsehood: The Real Value of Knowing that p,' *Philosophical Studies*, 107: 87–108.

——(2002*b*). 'Reliability and the Value of Knowledge,' *Philosophy and Phenomenological Research*, 64: 79–96.

Sartwell, C. (1991). 'Knowledge is Merely True Belief,' *American Philosophical Quarterly*, 28: 157–65.

——(1992). 'Why Knowledge is Merely True Belief,' *Journal of Philosophy*, 89: 167–80.

Sosa, E. (1991). *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge, MA: Cambridge University Press.

——(2003). 'The Place of Truth in Epistemology,' in M. DePaul and L. Zagzebski, (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press, 155–79.

——(2007). *A Virtue Epistemology*. Oxford: Oxford University Press.

Stich, S. (1990). *The Fragmentation of Reason*. Cambridge, MA: MIT Press.

Swinburne, R. (1999). *Providence and the Problem of Evil*. Oxford: Oxford University Press.

von Kutchera, F. (1982), *Grundfragen der Erkenntnistheorie*. Berlin: De Gruyter.

Whitcomb, D. (2006). 'Knowledge, Virtue, and Truth' (manuscript in preparation, Department of Philosophy, Rutgers University).

Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

Zagzebski, L. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge, MA: Cambridge University Press.

——(2000). 'From Reliabilism to Virtue Epistemology,' in G. Axtell (ed.), *Knowledge, Belief, and Character*. Lanham, MD: Rowman and Littlefield.

——(2003). 'The Search for the Source of Epistemic Good,' *Metaphilosophy*, 34: 12–28.

Queries in Chapter 1

Q1.   Opening paranthesis is missing in this place. Please check.