

Reliabilism and the Value of Knowledge

Draft 29-06-06

Comments welcome!

Alvin I. Goldman and Erik J. Olsson

Abstract: Most contemporary epistemologists agree that knowledge is more valuable than mere true belief, and many argue that process reliabilism is unable to account for this fact. We begin with a partial challenge to the first claim. There is a weak sense of ‘knowing,’ we suggest, in which knowing simply consists in believing truly. In this weak sense, knowing isn’t more valuable than believing truly. Of course, there is also a stronger sense of ‘knowledge’ that involves more than true belief and is more valuable. We explore two ways in which process reliabilism can account for this added value. The first approach contends that, under one legitimate conception of instrumental value, the instrumental value of a reliable process token isn’t wholly derived from the value of the true belief token it causes, and hence its value isn’t “swamped” by the value of that token. The second approach takes a different angle. If knowing consists of true belief formed by a reliable process, a knower has the ability to acquire more true beliefs on similar occasions by taking advantage of the same reliable mechanism. Thus reliabilist knowing, conceived as a state-of-affairs, has greater value than merely believing truly in that it makes future true belief more likely.

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 seems to obtain, or to be 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. So, what is the evidence of 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 on a given dimension, at least for a given person and a given 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.

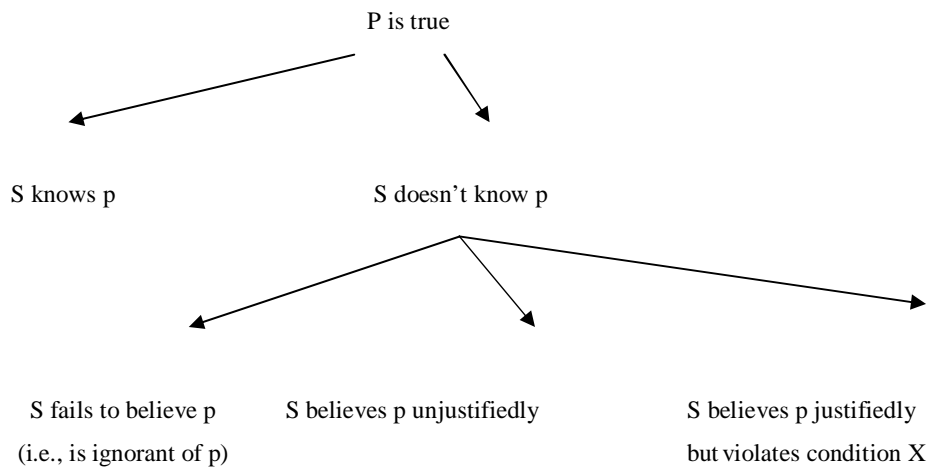
To illustrate these suggestions, consider a case discussed by John Hawthorne (2002). If I ask you how many people in the room know that Vienna is the capital of Austria, you will tally up the number of people in the room who possess the information that Vienna is the capital of Austria. Everyone in the room who 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 were given the information by somebody they knew wasn’t trustworthy, they would still be counted as cognizant of the fact, i.e., 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, i.e., *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. 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 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.

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 position of Crispin Sartwell (1992), who holds 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.^{1 2} 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 will be concerned with knowledge in its strong sense.

2. The value of reliabilist knowledge

What is it that makes us think that knowledge, in the strong sense, is 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 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 value problem can be used to test the adequacy of a given account of knowledge. Suppose a given analysis of knowledge cannot account for the added value of knowledge. That would be a pretty strong argument against the adequacy of that analysis. Recently a

¹ That there are both weak and strong senses of ‘know’ is advocated in Goldman, 1999. For a critical discussion, see Le Morvan (2005).

² It should be noted that even though knowledge, in its weak sense, is no more valuable than mere true belief for the trivial reason that such knowledge equals true belief, it can still be more valuable than its absence, ignorance. For, if p is true, it is generally better to believe p than not to believe it, especially if the truth of p is something which the agent cares about. Moreover, if p is believed, it is better that it be true than false. Thus, the problem of accounting for the greater value of knowledge over its conceptual parts – what Pritchard (forthcoming) calls the secondary value problem (as opposed to the primary problem of showing why knowledge is more valuable than mere true belief) – has a relatively simple positive solution for knowledge in its weak sense. Stephen Stich (1990), pp. 122-3, has challenged the claim that true belief is generally more valuable than false belief when it comes to promoting practical ends (1990), pp. 122-3. For a discussion, see Goldman (1999), pp. 73-5.

number of authors have argued that process reliabilism³ does not pass the value test. 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. The purpose of the anti-Gettier clause is to exclude various counterexamples similar to those produced by Edmund Gettier in his 1963 paper. 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” (p. 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” (p. 58).

For similar arguments, see Linda Zagzebski (1996, 2000, 2003), Wayne Riggs (2002), Jonathan Kvanvig (2003), and Ernest Sosa (2003). As Jonathan 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.

Against the swamping argument one could hold that few reliabilists have claimed that knowledge amounts to nothing but true belief reliably produced; as we have already noticed, reliabilists have usually also insisted on an anti-Gettier clause. Adding such a clause opens up for the possibility that it is the satisfaction of that clause that gives reliabilist knowledge its additional value over mere true belief. Even if the value of true

³ For an early statement of process reliabilism, see Goldman (1979); see also Goldman (1976) and (1986).

belief reliably formed does not exceed the value of true belief *simpliciter*, the value of a true belief reliably formed *in a way 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 mainly be concerned with a simple reliabilist theory according to which knowledge requires the satisfaction of only (1) – (3).

The conclusion of the swamping argument is generally taken to be 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.⁴ 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.

⁴ Ward Jones (1997) takes the less common position of arguing that the value problem, while problematic, does not show reliabilism to be false.

Since (S3) is a highly counterintuitive conclusion and the argument appears valid, one of the premises must be false.⁵ 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, e.g. 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,

“the 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 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, p. 13)

What Zagzebski is saying, or implying, 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, that 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.⁶ 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.⁷

⁵ The validity of the standard swamping argument will be questioned below.

⁶ For a clear statement of this point, see Percival (2003), p. 33. As for the role of veritism, see also Jones (1997), p. 424, and, following him, Riggs (2002), p. 82.

⁷ See for instance Goldman (2002), p. 53. Goldman's theory is the explicit target of Jones's and Riggs's swamping reasoning. See footnotes 4 and 5 on p. 438 in Jones (1997) and p. 80 in Riggs (2002). Zagzebski mentions Alston, Plantinga, Sosa and Goldman as advocating epistemological theories that are vulnerable to swamping problems. See footnote 5 on p. 14 in Zagzebski (2003).

The standard swamping argument should be distinguished from the swamping argument given 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 as given 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 expresses that being reliably produced just means being likely to be true.⁸ The fact that Kvanvig’s version of the argument is essentially different from the standard version seems to have gone unnoticed in the literature.

Unfortunately for Kvanvig, however, his (K2) premise is false. While it is plausible to assume 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

⁸ Kvanvig commits himself to (K2) in the course of his chocolate analogy on pp. 47-48 in his 2003 book.

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 is committing what seems to be a category mistake.

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)

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% chance of winning one thousand dollars (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).

Of course, this argument is absurd. 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 it 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.

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. We call this central problem the *extra-value-of-knowledge* (EVOK) problem. Now, in the case of reliabilism it is generally presupposed that the only way reliabilism can solve the EVOK problem is to say that the use of 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) that has 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 also possible, however, for reliabilists to *reject* this presupposition and try to solve the EVOK problem by circumventing the double-counting objection.

In the next two sections we offer two possible solutions to the EVOK problem. Our first solution, presented in section 4, accepts the indicated presupposition and tries to show that the swamping objection can be surmounted. Our second candidate solution, presented in section 5, rejects the indicated presupposition. It doesn’t try to show that the extra value of knowledge is derived from the instrumental value of the (token) reliable process that produces the target belief. This solution doesn’t have to solve the swamping or double-counting problem; it just sidesteps that problem. We shall present these two solutions, explore their relative advantages and disadvantages, but then leave it to the reader to decide which is more attractive all things considered. Either one, in our opinion, is a pretty compelling solution, and it isn’t essential for us to decide which is superior.⁹

4. Two conceptions of instrumental value and a solution to the swamping problem

⁹ One of us (AIG) favors the first solution and the other (EJO) favors the second. .

We turn now to our first solution to the EVOK problem. It tries to answer the EVOK problem by appeal to the reliable belief-forming process that is (allegedly) present in each case of knowledge. Each such reliable process has epistemic value above and beyond the possession of true belief (which also has epistemic value). Proponents of the swamping argument claim that the value of the reliable process will always be *derivative* from the value of the true belief it produces. This derivative value isn't really *extra* value, because it is wholly traceable to the true belief it causes. It is like the value of the certificate in our safe-deposit box example. Its value is wholly traceable to the lump of gold in the box. It would be double-counting to suppose that the reliable belief-forming process adds epistemic value to that of the true belief. This can be explained in terms of what we'll label the "Simple Derivation Thesis," which resembles Property Parasitism but improves upon it by citing *states* rather than *properties* as the bearers of value.

(Simple Derivation Thesis) If the value of one state is entirely derivative from the value of a second state, the value of the composite state consisting of the two components does not exceed the value of the second component.

This thesis is appealing if we focus on the popular distinction between *fundamental* and *instrumental* value, which seems applicable to the case of a true belief caused by a reliable process. Suppose we continue to assume *veritism*, a form of epistemic value monism that holds that true belief is the only fundamental epistemic value. Process reliabilism isn't *committed* to epistemic value monism, we maintain, but let's assume it for the sake of argument. (For further discussion, see Whitcomb, 2006.) The epistemic value of a reliable process, then, cannot be fundamental; presumably it will be instrumental. This is a very natural assumption, because the reliability of a belief-forming process is a matter of the belief-forming process having a propensity to generate a high proportion of true beliefs (Goldman, 1979). So it's a conceptual truth that reliable belief-forming processes are good "instruments" to achieving true belief.

Next, it is quite tempting to think that instrumental value is entirely derivative from fundamental value. And if the Simple Derivation Thesis is correct, then a state of affairs that has a certain quantity of fundamental value does not gain in overall value by having a certain quantity of instrumental value, at least if the instrumental value is wholly derived from the fundamental value *of the same state*. If it's derived from the fundamental value of *other* states, the matter may be different, of course. But the proponent of the swamping argument insists that this is precisely the situation with a reliable process that successfully

produces a true belief. The reliable process has derivative, i.e., instrumental, value; but that instrumental (epistemic) value is wholly derived from the fundamental (epistemic) value of the true belief in the very same knowledge state. So it adds nothing to the overall (epistemic) value of that knowledge state, above and beyond the value of the true belief.

Is this correct? No. The claim rests on a particular conception of instrumental value that is admittedly widespread among philosophers but is open to challenge. In particular, we propose that there are two viable conceptions of instrumental value, and the second is the more appropriate one for process reliabilism. Under this second conception of instrumental value, the double-counting charge can be rejected for process reliabilism.

Let us distinguish *token instrumentalism* and *type instrumentalism*, which can be formulated as follows:

(Token Instrumentalism) One token event (or state of affairs) inherits value from any other token event (or state of affairs) that it causes if and only if the resulting token event has fundamental value.

In other words, according to Token Instrumentalism, one token event inherits value from a second token event if and only if the former bears the singular-causation relation to the latter and the latter has fundamental value.

(Type instrumentalism) A type or class of events C inherits value from another type or class of events E if instances of C tend to cause instances of E and instances of E (invariably or usually) have fundamental value. Derivatively, each instance or token of such a type C has instrumental value by virtue of belonging to such a type.

The standard conception of instrumental value among philosophers, we believe, is token instrumentalism. But the conception of instrumental value that seems most naturally applicable to the case of knowledge and process reliabilism is type instrumentalism. So let us focus on this kind of instrumentalism, and give an example of it.

Relief from pain – e.g., the pain of a headache – is plausibly considered a fundamental (hedonic) value. Taking aspirin is a type of event the tokens of which tend to produce pain-alleviating results. Thus, taking aspirin is a type of action that has instrumental value. Hence, any particular instance or token of such a type also has instrumental (hedonic) value under the type-instrumentalist conception, by virtue of belonging to this type. This holds of

each particular token of the type even if it fails to produce pain alleviation, or any other result with fundamental hedonic value.

Apply this theme to the case of knowledge and instrumental epistemic value. Consider a type of belief-forming process that generally produces true beliefs, say, the reasoning process of modus ponens. (We assume, with some loss of accuracy, that modus ponens deserves to be called a reasoning process, as contrasted with an argument form). If modus ponens reasoning as a type is generally reliable, this means that tokens of that type usually produce true belief results. (Strictly, this is guaranteed only when the premises that the process takes as inputs are true. But let's not worry about these details.) Since true belief is assumed to have fundamental epistemic value, the modus ponens type of reasoning should be deemed to have instrumental epistemic value. Moreover, any particular *token* of this type also has instrumental epistemic value, under the type-instrumentalist conception. Notice, however, that this doesn't depend on each particular token actually causing a true belief. A particular instance of modus ponens reasoning (one that starts with at least one false premise) might not produce a true belief. But it still has instrumental (epistemic) value under the type-instrumentalist conception.

Now suppose that a particular token reliable process does produce a true belief, and hence, under simple process reliabilism, the whole state of affairs comprises a case of knowledge. Does the epistemic value of this knowledge state of affairs exceed the epistemic value of its component true belief, in virtue of the fact that another one of its components – the reliable process – has instrumental epistemic value? Or, as the swamping argument contends, would adding that instrumental epistemic value constitute an objectionable instance of double counting? If we are right that instrumental epistemic value accrues to the reliable process token via the *type-instrumentalist* conception, it should not be considered an instance of double counting. The instrumental epistemic value of the reliable process token doesn't wholly derive from the fundamental value of the true belief token that the process causes. It ultimately derives from the fundamental value of "foreign" sources, i.e., true beliefs that are (or would be) caused by *other* tokens of the same reliable process type. Thus, the double-counting objection lodged against process reliabilism doesn't go through.¹⁰

A critic might respond that under type-instrumentalism, so-called instrumental "value" isn't a genuine form of value (*simpliciter*), but rather a form of *expected value*.¹¹ Taking aspirin is considered a good course of action when one has a headache because (i) taking

¹⁰ The embryo of a type-instrumentalist approach to the value problem can be found in Percival (2003), p. 38.

¹¹ This objection was raised by Holly M. Smith.

aspirin has an *expected* outcome, viz., relief from pain, and (ii) relief from pain is valuable. In other words, our so-called valuing attitude vis-a-vis aspirin is not a pure valuational attitude but a combination of valuing and expecting (e.g., assigning a subjective probability). Expected value, however, should not be confused with genuine instrumental value. Genuine instrumental value only makes sense under the token-instrumentalist conception.

At least in psychological terms, this criticism is off base. We don't generate, on the fly, a fresh subjective probability about aspirin's effectiveness each time we take one. We learn early in life to value aspirin as a means to relieving headaches (etc.), so we don't have to re-think the probabilities afresh each time. Similarly, a salary raise is valued without having to calculate whether the extra money will satisfy more desires. We learn early in life that money is of instrumental value; so wealth becomes an entrenched value. It no longer occupies our psychic lives as a combination of two things: remote final ends plus probability judgments about their greater attainability via greater wealth. Still, its valuational basis is instrumental. Similarly, epistemic valuations for reliable belief-forming processes can be well entrenched even if they are instrumentally derived from true-belief valuations.¹² This accords with type-instrumentalism.

Other critics might complain that we are now confusing questions of objective value with questions of subjective value, or valuational activity. According to such critics, this obscures an important distinction between the conditions in which things are objectively good, on the one hand, and the ways we value things, on the other.¹³ An exploration of this issue, which lies at the heart of meta-ethics, exceeds the scope of the present paper. We content ourselves, for now, with the following comment. Without disputing our access to matters of objective value, it seems that our most direct access is to our own valuational states. It is these valuational states, or intuitions (intuitings) of value, that presumably provide epistemologists with their primary evidence that knowledge has value, and, in particular, more value than true belief. These valuing, or value assignments, comprise the primary "data" that need to be explained. Just such an explanation is on offer in this section, an explanation that invokes both the process-reliabilist account of knowledge and the type-instrumentalist conception of instrumental value.

¹² We don't wish to imply that things have instrumental value for an individual only if its being valued was the result of means-ends calculation at some juncture in the valuer's ontogeny. Some values may be acquired interpersonally (by some sort of "social learning"), and their status as instrumental values may arise from the nature of the community's collective valuational processes, rather than the processes of each individual.

¹³ See, for example, Korsgaard (1983).

5. The conditional probability solution

Our first solution said that a token reliable process has a kind of instrumental value that can legitimately be *added* to the value of its resulting true belief to yield a value for their composite state (of knowledge) that exceeds the value of the true belief alone. Our second solution makes no such claim. It merely says that 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. This property, moreover, is a valuable one to have – indeed, an epistemically valuable one. Therefore, *ceteris paribus*, 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 objectively interpreted.

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 it 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 espressos 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 was to be expected, because the system is reliable. In the second it happened 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. Hence, you are more likely to arrive at Larissa on time if you know in Situation 1 than if you merely have true belief in Situation 1.

We have 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 role 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 yielded a correct belief on its first occasion of use should 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.¹⁴

¹⁴ 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's book from 2000 features an account similar to our conditional probability account (see pp. 100-102). However, Williamson 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's 2003 book, which is otherwise well informed by the existing literature. Ward 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" (p. 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

6. Advantages and disadvantages of the two solutions

As we pointed out we will not take a definite stand as to which of our two proposals is better. However we will discuss three dimensions along which they may be compared. The first dimension concerns how often knowledge comes out as having extra value. We will refer to this as the *frequency* dimension. The second dimension has to do with the degree to which the extra value of knowledge is modally stable, in a sense to be explained. This will be called the *modal* dimension. The third dimension relates to the extent to which these solutions adhere to the principles of naturalistic epistemology. We will refer to this as the *naturalistic* dimension.

Let us start with the frequency dimension. On the conditional probability solution, knowledge will have 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. The condition of non-uniqueness, for instance, says that it is likely that you will encounter problems similar to the present one in the future. But suppose the present problem facing you happens to be one of those rare once-in-a-life-time problems. Since you will never face a similar problem again, the probability that you will obtain more true beliefs in the same general way is no greater conditional on your reliabilist knowledge than conditional on a mere true belief. Similarly, if the condition of cross-temporal access fails to hold, you won't have access to the reliable method the next time the same type of problem crops up, and so, again, no added value arises in the conditional probability sense. The same holds if you, for some reason, decide not to use the reliable process the next time around, thus violating the condition of learning; or if the reliable method you used is reliable only on very specific occasions so that the condition of generality is not satisfied. The conditions of non-uniqueness, cross-temporal access, learning and generality are normally satisfied, which means that knowledge normally has an added value. However, we would be hard-pressed to claim that those conditions always hold. When they fail to hold, knowledge also fails to have an extra value in the present sense.

What about the type-instrumentalist solution? How often does knowledge have the extra value postulated on this approach? On type-instrumentalism, a reliable process token inherits its instrumental value from its type. The value of the type is derived from the value of all true beliefs that are – or would be – caused by tokens of the same reliable process

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.

type. Our formulation leaves it open whether what is relevant is actual or potential causation. Which option we choose has some bearing on the issue we are now focusing on. If the value of a reliable process type depends on true beliefs that are *actually* caused by (actual) tokens of that type, then, conceivably, knowledge may fail to have any added value. For there may fail to be any actual instantiations of the reliable process -- at least any instantiations beyond the target case -- and hence there may fail to be true beliefs caused by such tokens. If, however, the value of the process type depends on all true beliefs *potentially* caused, the situation is different. For then the type will always accrue added value from other true beliefs, quite independently of whether there are any actual (additional) instantiations of that type. Regardless of which option is chosen, however, type instrumental value is clearly less dependent on empirical circumstances than is the value that emerges from facts of conditional probability, and this is the main point we want to make here.

Should it follow from an account of the extra value of knowledge that knowledge has this extra value always? This is a matter of some controversy. Several authors have expressed complete satisfaction with establishing weaker conclusions. Swinburne (1999) concludes that knowledge has an added value by arguing that it has this value “almost always” (p. 64). Williamson (2000) maintains that knowledge is more valuable provided that the cognitive faculties of the knower are in good order (p. 79), a condition that may occasionally fail to hold. Finally, Percival (2003) thinks that what needs to be shown is that knowledge has added value “by and large” (p. 38). Clearly, all these authors think that the value problem can be solved without there being a need to show that the greater value of knowledge over mere true belief is realized *in every single case*. For these authors, the performance of the conditional probability solution on the frequency dimension should be deemed satisfactory. However, other authors disagree. As Riggs conceives of the “value principle”, “[k]nowledge is *always* more valuable than (mere) true belief” (2002, p. 79, our emphasis). Kvanvig (2003) similarly insists that what is to be certified is the “*unqualified* value of knowledge over true belief” (p. 57, our italics). These authors would consider the conditional probability solution incapable of accounting for the whole of the value of knowledge. Type-instrumentalism may still be a promising candidate, though.

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 that we subscribe to are arguably of a “generic” rather than an “absolute” kind. Surely you agree that money is a valuable thing to have, and yet you must also agree that that rich people are sometimes killed because of their wealth, and so for them money was actually something

bad. Surely you agree that birds fly, and yet you must also agree that 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 general claim that knowledge is more valuable than mere true belief be any different?¹⁵

Suppose that it follows from a given account that knowledge has extra value normally/always/sometimes etc in the actual world. Then we may ask how stable this conclusion is over possible worlds other than the actual world. Hence, we may compare different accounts of the value of knowledge on a *modal* dimension as well. According to the conditional probability solution, knowledge is normally more valuable than mere true belief in the actual world. This result depends on certain empirical regularities – non-uniqueness, cross-temporal access, learning and generality – that plausibly hold in the actual world. However, once we take other possible worlds under consideration, there is no guarantee that these conditions will continue to obtain. We can imagine, for instance, worlds where agents always face problems that never recur in those agents' lifetimes; or worlds where all reliable methods are such that they work only under extremely specific conditions so that the generality condition is typically violated. Thus, there are possible worlds where knowledge does not normally have an added value in the conditional probability sense.

The conditional probability solution scores badly on the modal dimension. Does the type-instrumentalist approach fare any better? It seems that it does, provided that value is inherited from all *potential* true beliefs obtained from using the reliable process rather than merely actual ones. On that reading of the type-instrumentalist solution, knowledge, as we have seen, has its extra value *always* in the actual world. For the reliable process token inherits its value from all *potential*, i.e. possible, true beliefs caused by tokens of the same general kind, and there is no shortage of such potential true beliefs from the perspective of the actual world. But if a given such true belief is possible from the perspective of the actual world it is also possible from the perspective of any other possible world (assuming the standard S5 axioms for modal logic). Hence, there are potential true beliefs caused by tokens of the same general kind *in all possible worlds*. Hence, knowledge has its extra value always *as a matter of necessity*. The alternative rendering of the type-instrumentalist

¹⁵ Cf. Jones (1997), p. 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."

solution in terms of actual true beliefs caused by reliable process tokens is less stable over other possible worlds. For even if there are such actual true beliefs caused in a particular way in the actual world, there may well be other possible worlds where this does not hold. In such worlds knowledge may lack any extra value.

The modal dimension of the value problem has not received much attention in the literature. An exception is Kvanvig, who not only believes that knowledge has its added value always, but also, it seems, that this is so necessarily. As he puts it, “[i]t is simply false that knowledge loses its value in worlds where the environment is less cooperative and where pragmatics play a more significant role in belief formation” (2003, p. 17). Contrarywise, Ward Jones (1997) argues that to the extent that knowledge has an added value, it “will come from contingent characteristics of knowledge” (p. 433). We tend to side with Jones in this case: as far as pre-systematic intuition is concerned, it suffices to establish that knowledge is more valuable in the actual world.

Let us now turn to the naturalistic dimension. Many philosophers are drawn to reliabilism because they think it can provide the basis for a completely naturalistic account of knowledge, i.e., an account that makes knowledge part of the natural world, as opposed to being irreducibly evaluative or metaphysical. These philosophers may wonder which solution to the value problem is most in line with a naturalistic outlook. The matter is obviously a rather delicate one. For if values are foreign to the naturalistic outlook, how could one even begin to address the value problem from such a point of view? Still, many naturalists hold that evaluative terms like “good reasons”, “justification” and so on can be given naturalistic definitions, e.g., in terms of reliability or objective probability; and if this is correct, there seems to be no difficulty in principle in providing a naturalistic account of further evaluative concepts, including the value of knowledge.

From a naturalistic perspective, the conditional probability solution seems to have a clear advantage over the type instrumentalist approach. The former seeks to explain the value of knowledge by referring to the objective (conditional) probability of obtaining further true beliefs. Knowledge has an added value in this sense, it is contended, given that certain empirical conditions – such as non-uniqueness and generality – are satisfied which generally they are. All components of this view are naturalistically available. The type-instrumentalist solution, by contrast, relies on concepts that are not entirely naturalistic. According to this solution, a type or class of events C “inherits” value from another type or class of events E if instances of C tend to cause instances of E and instances of E have fundamental value. But the extent to which such value is inherited doesn’t seem to be the kind of thing that we could, even in principle, detect in an experimental setup or by other

empirical means. On type instrumentalism, furthermore, each instance or token of such a type C has instrumental value “derivatively” by virtue of belonging to such a type. Again, whether or not this is so doesn’t seem to be an entirely empirical matter.

On the other hand, if one adopts a *psychologistic* brand of naturalism, one that tries to explain valuing activities in psychological terms, the type-instrumentalist approach may be naturalistically quite respectable. Viewed from a psychological perspective, it doesn’t seem mysterious that people should make value assignments to types based on value assignments they make to observed token events. And it doesn’t seem psychologically mysterious that they should link or import their valuations of types to valuations of further tokens.

7. 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 approaches that either display a sharp contrast with ours or that feature 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 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 (by our lights) exaggerates 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, p.19). This strikes us as imposing an unduly heavy burden on many cases of knowledge, especially unreflective knowledge by animals and young children, among others. 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?¹⁶ 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, p. 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 Diana. Having once belonged to Diana is an extrinsic (or external) rather than intrinsic (or internal) property of the dress.

Let us turn now to moderate VE. Some themes in moderate VE are fairly congenial to at least one of the two approaches 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 is one element of moderate VE that one of our approaches readily finds attractive.

¹⁶ 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).

However, it is not so clear that stability cleanly separates generic reliabilism from virtue 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.¹⁷ On the other hand, Greco hasn't persuaded us that the stability, or non-fleetingness, of a cognitive source is a strictly necessary condition of knowing. We can easily generate cases in which knowledge occurs through a fleetingly possessed method or skill, i.e., cases in which a cognitive skill or method is newly acquired and successfully applied to form 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 an emphasis on agency, attributability, and credit-worthiness. According to Riggs (2002), 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 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 the archer who hits his target by skill (i.e., 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.

¹⁷ In discussing Greco's requirement of a stable disposition, Berit Brogaard (forthcoming) 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.

However, these points seem to be pretty congenial to generic reliabilism, and not really 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 idea. In particular, reliability theories have proposed either sensitivity, safety, or the absence of relevant alternatives as a form 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 does own a Ford, but only 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 no-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 by no means the special preserve of virtue reliabilism. Virtue reliabilists introduce special *language* to describe these cases, but the nuts and bolts of explaining the conditions has usually proceeded in a generic reliabilist fashion, using concepts like belief, truth, and various possible-world permutations.

Notice that the first of our proposed solutions to the EVOK problem, the type-instrumentalism solution, could be refined by reference to a suitably chosen reliabilist non-accidentality condition of the type under discussion. How might this proceed? As presented in section 4, the value of a (candidate) state of knowledge derives from the value of the token belief-forming process that is a constituent of this state. This token process inherits value from the *type* it instantiates. However, in the spirit of the preceding paragraphs, it might be argued that the *degree of value* assignable to a candidate knowledge state is not merely a function of (1) the value of the true belief that is a constituent of the state, and (2) the value associated with the instantiated process type. Such a procedure would assign too much epistemic value in the Gettier disjunction case. That isn't even a genuine case of knowledge, so the two sources of value just identified cannot alone delimit a suitable value calculation. What needs to be done, arguably, is to take into account the *degree* to which the process type used is causally responsible for the true belief outcome. If the same belief, produced by the same process or method, could easily have been false, this dilutes the causal responsibility or causal credit attributable to that process or method. Once this factor

is taken into account, a more refined procedure for assessing epistemic value would emerge.

Does this proposal, expressed in the language of “credit”, “responsibility” and “attributability”, play right into the hands of *virtue* reliabilism? Not so, we reply. Our proposal doesn’t draw on any elements foreign to generic reliabilism, or any elements foreign to a purely event-oriented metaphysics that has no truck with agency in the sense usually invoked by virtue theorists. The idea that events can have different degrees of causal influence on other events must be acknowledged in many domains. Consider voting, for example. If two corporate shareholders cast their shares in favor of a certain proposal, which wins, their respective votes need not carry equal causal weight. If one shareholder cast 500 shares for the proposal, and the other shareholder cast 10,000 shares for the proposal, the second set of votes obviously carried more weight, and deserves more causal “credit” for the outcome. Similarly, if a muscular person and a weakling each exert force on a car to push it out of a snowdrift, the causal credit due to the two acts of pushing needn’t be the same. The territory here is causal influence between events. That territory shouldn’t be ceded to *virtue* theory simply because of the language of credit or attribution.¹⁸

However, there is a serious question of how this approach can be applied to the case of influencing true belief. What does it mean to say that a belief-forming process is causally responsible for a true belief? Does it mean that the process is responsible for the *belief*, or does it mean that it is responsible for the *truth* of the belief? In the former case, it’s hard to see how one could claim, in the Gettier case, that the inference process makes only a *minimal* contribution to the production of the belief. On the contrary, it seems to make quite a major contribution. On the second reading, Smith’s reasoning admittedly has little or no causal responsibility for the truth of the disjunctive belief. But this almost always seems to be the case. Reasoning is hardly ever causally responsible for the truth-values of its propositional contents. Unless this sticky problem can be resolved, all parties to the discussion – whether virtue reliabilists or generic reliabilists – will find it hard to make clear sense of degrees of causal influence in this domain.

Finally, some writers on the value problem have explored this territory by appealing to novel types of value. We are thinking primarily here of Ernest Sosa’s (2003) recent contribution to the subject, where he introduces the novel notion of “praxical value”:

¹⁸ For a related attempt, see Goldman (1974), which formulates a theory of degrees of *power* that employs conceptual resources and strategies that might prove applicable to the causal domain generally.

“An agent A may bring about an event E. The bringing about of E by A may then itself be assessed. This event, call it E', may not have any intrinsic value beyond the intrinsic value contained already in E, but it will have instrumental value proper to the special relation involved in E's happening because of E'. Call this special sort of instrumental value *praxical value*, the sort of instrumental value in actions of bringing about something valuable. Now, for the hedonist, an event of someone's being pleased does contain some measure of intrinsic value. Supposing someone brings about that pleasure, is there also value in this further event? For our hedonist there is here no distinctive intrinsic value, no intrinsic value *beyond* that found in the pleasure brought about. But even a monistic hedonist may yet find in that action some degree of praxical value.” (2003, pp. 162-163)

Sosa then applies the idea of praxical value to knowledge:

“What precludes our conceiving of knowledge in a similar way, as a desideratum that includes an intrinsic success component, a hitting of the mark of truth, along with *how* one accomplishes that, how one succeeds in hitting the mark of truth? On this conception, knowledge is not just hitting the mark but hitting the mark somehow through means proper and skilful enough.... [N]o additional *intrinsic* or *fundamental* value could account for the value in the skilful shot over and above the mere hitting of the bull's eye. The further value might rather be just *praxical* value or the like.” (2003, p. 164)

This appears to be the heart of Sosa's solution to the value problem (although the paper then proceeds with further elaborations). How close is it our own first solution? That depends on the metaphysical status of the special “bringing about” events that Sosa postulates, to which praxical value attaches. It is clear that Sosa regards an event E' of bringing about E as ontologically distinct from E itself. Does he also mean to regard it as an event distinct from the event or process that brings about E? For comparison, consider (1) a person's flipping a switch, (2) the light's going on (as a result of the flipping), and (3) the person's causing the light to go on (i.e., his bringing it about that the light goes on). Clearly, event (3), the bringing-about event, isn't identical with event (2). But is it identical with event (1)? This is a topic on which theorists of action have disagreed, and we're not sure where Sosa stands. If he is prepared to side with the “unifiers”, and identify (3) with (1), then his approach to the value problem may be quite similar to ours.¹⁹ At any rate, since

¹⁹ Except there is nothing in Sosa's theory that resembles the *type*-instrumentalist maneuver that distinguishes our (first) solution.

Sosa views praxical value as a species of instrumental value, he ostensibly proposes to *add* this instrumental value to the entire state of affairs, yielding a sum-total of value greater than the fundamental value of a true belief. How is this intended to circumvent the swamping challenge, however? This isn't clear. Also, could one kind of instrumental value – *non-praxical* instrumental value – attach to the causing process (analogous to event (1) in the switch-flipping case) *in addition* to the praxical instrumental value attached to the bringing about of the true belief? Could *both* kinds of instrumental value “kick in”, and add to the sum-total of value? If so, it strikes us that Sosa's account incurs a very serious risk of double counting, indeed, triple counting! We suspect there is a danger here, but we don't feel we have a clear enough grip on the nature of praxical value as Sosa conceives it to assess the seriousness of the risk.

References

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

Brogaard, B. (forthcoming), "Can Virtue Reliabilism Explain the Value of Knowledge?", *Canadian Journal of Philosophy*.

Goldman, A. I. (1974). "On the Measurement of Power," *Journal of Philosophy* 71: 231-252. Reprinted in Goldman, *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge, MA: MIT Press (1992).

Goldman, A. I. (1976), "Discrimination and Perceptual Knowledge", *The Journal of Philosophy*, 73: 771-91. Reprinted in *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge, MA: MIT Press (1992).

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

Goldman, A. I. (1986), *Epistemology and Cognition*, Cambridge, MA: Harvard University Press.

Goldman, A. I. (1999), *Knowledge in a Social World*, Oxford: Clarendon Press.

Goldman, A. I. (2002), "The Unity of the Epistemic Virtues," in *Pathways to Knowledge*, New York: Oxford University Press.

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

Greco, J. (2000). *Putting Skeptics in Their Place*. Cambridge: Cambridge University Press.

- Jones, W. E. (1997), "Why do we value knowledge?" *American Philosophical Quarterly* 34, No. 4, October: 423-439.
- Korsgaard, C. M. (1983), "Two Distinctions in Goodness," *Philosophical Review* 92: 169-196.
- Kvanvig, J. L. (1998), "Why should inquiring minds want to know? Meno problems and Epistemological Axiology," *The Monist*.
- Kvanvig, J. L. (2003), *The Value of Knowledge and the Pursuit of Understanding*, Cambridge University Press.
- Le Morvan, P. (2005), "Goldman on knowledge as true belief", *Erkenntnis* 62 (2): 145-155.
- Percival, P. (2003), "The Pursuit of Epistemic Good", *Metaphilosophy* 34, Nos. 1/2: 29-47.
- Pritchard, D. (forthcoming), "Recent work on epistemic value", unpublished manuscript.
- Riggs, W. D. (2002), "Beyond truth and falsehood: The real value of knowing that p", *Philosophical Studies* 107, 87-108.
- Riggs, W. D. (2002), "Reliability and the Value of Knowledge", *Philosophy and Phenomenological Research* 64: 79-96.
- Sosa, E. (1991). *Knowledge in Perspective, Selected Essays in Epistemology*. Cambridge: Cambridge University Press.
- Sosa, E. (2003), "The Place of Truth in Epistemology", in *Intellectual Virtue: Perspectives from Ethics and Epistemology*, Michael DePaul and Linda Zagzebski, edited. Oxford University Press (pp. 155-179).
- Stich, S. (1990), *The Fragmentation of Reason*. Cambridge, Mass.: MIT Press.
- Swinburne, R. (1999), *Providence and the Problem of Evil*. Oxford: Oxford University Press.

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

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: Cambridge University Press.

Zagzebski, L. (2000). “From Reliabilism to Virtue Epistemology,” in *Knowledge, Belief, and Character*, edited by Guy Axtell. Lanham, MD: Rowman and Littlefield.

Zagzebski, L. (2003), “The Search for the Source of Epistemic Good”, *Metaphilosophy*, 34: 12-28.