



# LUND UNIVERSITY

## In Defense of the Conditional Probability Solution to the Swamping Problem

Olsson, Erik J

*Published in:*  
Grazer Philosophische Studien

2009

[Link to publication](#)

*Citation for published version (APA):*

Olsson, E. J. (2009). In Defense of the Conditional Probability Solution to the Swamping Problem. *Grazer Philosophische Studien*, 79, 93-114.

*Total number of authors:*

1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00



# **In Defense of the Conditional Probability Solution to the Swamping Problem**

Erik J. Olsson  
Lund University

Olsson, E. J. (2009). In Defense of the Conditional Probability Solution to the Swamping Problem. *Grazer Philosophische Studien*, 79, 93-114.

## **1. Introduction**

It is commonly agreed that knowledge is more valuable than mere true belief. Many also believe that process reliabilism is incompatible with this item of common sense.<sup>1</sup> The reason is the so-called swamping problem: if a belief is already true, adding that it was reliably produced doesn't seem to make it more valuable. The value in reliable production seems to lie merely in its being indicative of the truth of the belief thus produced. But if that belief is already assumed true, no further value is conferred by assuming, in addition, that the acquisition process was reliable.

In Goldman and Olsson (2009), we pointed out that a reliabilist can defend her theory against the swamping objection by referring to the distinct value pertaining to reliabilist knowledge in virtue of the fact that such knowledge makes *future* true belief more likely.<sup>2</sup> In other words, the fact that a person knows that p, in the reliabilist sense of "knows", is indicative of that person's acquiring further true beliefs. This we called the *conditional*

---

<sup>1</sup> Classical formulations of process reliabilism can be found in Goldman (1979) and (1986).

<sup>2</sup> I use "probability", throughout this article, in its objective sense. I frequently refer to one thing q making another thing more p likely. By this I mean that the objective probability of p is greater given q than given non-q. I use "q is indicative of p" and "q improves the prospect of p" etc. as mere linguistic variations of "q makes p more likely".

*probability (CP) solution*. This value is attained normally and not in every single case. We went on to formulate a second and independent solution to the swamping problem, which may be called the *type-instrumentalism and value autonomization (TIVA) solution*. Among other things, this solution would explain why some people tend to think that knowledge is *always* more valuable than mere true belief.

In this paper, I will take the opportunity to discuss some objections that have been raised against the CP solution. But first I will give a more detailed account of the swamping problem and explain, in greater depth, how the CP solutions bears on it.

## **2. The CP solution to the swamping problem**

The swamping argument, as endorsed by Kvanvig (2003), Swinburne (1999), Zagzebski (2003) and others, may be presented schematically 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.

The characteristic swamping premise, (S2), derives support from an appeal to a principle of veritism:

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

If S's belief is already assumed true and all that matters in inquiry is the acquisition of true belief, then learning that S's belief was reliably produced does not make the belief more valuable, just as (S2) says.<sup>3</sup>

The standard reaction to the swamping argument is to reject (S1), that knowledge equals reliably acquired true belief. The CP solution is different in this regard. Rather than questioning the argument's premises, it purports to shed doubt on its validity. What it proposes is that reliabilist knowledge can be more valuable than mere true belief even if the belief in question is made no more valuable through becoming known. This is so because a *state* of knowledge can be more valuable than a *state* of mere true belief. This works because a state of knowledge is also a state of reliable acquisition and as such it is valuable not only as an indicator of the truth of the belief thus acquired but also as indicative of the production of further true beliefs (of a similar kind), namely, true beliefs resulting from reapplications of the reliable method in question. This is a reason why knowledge is more valuable than mere true belief even if the truth of both the premises employed by the swamping argument is granted.

How can it be that the probability of future true belief is greater conditional on knowledge as opposed to conditional on mere true belief? This conditional probability claim is true in virtue of certain empirical conditions characterizing our world. They are the conditions of *non-uniqueness*, *cross-temporal access*, *learning* and *generality*. By non-uniqueness, the same kind of problem will tend to arise over and over again. Once you encounter a problem of a certain type, you are likely to encounter a problem of the same type at some later point. Finding the way from point A to point B is that kind of problem. By cross-temporal access, a method that was used once will tend to be available also when the same type of problem arises in the future. This is true, for instance, of the method of

---

<sup>3</sup> For discussions of veritism, see Goldman (1999) and (2002).

consulting a map or your car's GPS navigation equipment. By the learning assumption, if a particular method solves a problem once, and you have no reason to believe that it did so unsuccessfully, then you will tend to use the same method again, if available. If, for example, you rely at point A on your GPS and arrive safely at point B, you are likely to rely on it again in the future. By generality, finally, if a method was reliable in the past, its reliability is unlikely to be discontinued. Rather, the method is likely to remain reliable in the future.

Clearly, there are a lot of exceptions to these empirical regularities – problems arising only one in a lifetime, navigation equipment that is stolen or suddenly breaks down etc. – but this is of little consequence in the present context so long as those regularities hold in normal circumstances. For, if they do, that means that knowledge is normally more valuable, i.e., more valuable in the conditional probability sense, than mere true belief. In my view, this is precisely what was to be demonstrated.<sup>4</sup>

For the record, it is not my view that the CP solution captures the *whole* sense in which knowledge is more valuable than mere true belief. I also happen to believe that the fact that a person knows that p, in the reliabilist sense, as opposed to merely truly believes that p, makes that person's belief that p more stable. More precisely, the probability that a given true belief will remain in place is greater given that it was reliably as opposed to unreliably acquired, and stability, as Williamson (2000) observes, is practically valuable in cases of prolonged action. For a detailed account of the stability aspect of reliable belief acquisition, see Olsson (2007).<sup>5</sup>

---

<sup>4</sup> Some researchers take a different stance on this matter. Thus Jonathan Kvanvig thinks that knowledge is always more valuable than mere true belief (at least for knowledge of non-trivia). The issue was dealt with in some length already in Goldman and Olsson (2009). Recently, Kvanvig has published an extended critique of the CP solution focusing partly on its lack of compliance with his view that the distinct value of knowledge should be attained universally. A reply to Kvanvig can be found in Olsson (to appear). The purpose of the present article is to focus on other objections to the CP solution.

<sup>5</sup> See also Olsson (2008). It might be thought that stability is not always a good thing. As the world changes, the truth values of sentences also change. What was once true can

### 3. Kvanvig's chocolate objection

In his useful 2003 book, Kvanvig presented an alternative version of the swamping argument based on a chocolate example. In Kvanvig (2008), he returns to the example in the context of a prolonged critical discussion of the CP solution, arguing that it presents a severe obstacle to the latter. There, the example is stated in the following form:

Suppose I want chocolate. I google to find places close to me. I get two webpages: one entitled “places that sell chocolate in Waco”; the other “places likely to sell chocolate in Waco”. We may assume accuracy for both lists, and that the second list is generated from correlations: places that sell food are likely to sell chocolate, places that sell hard candy are too, etc. ... We then note [that] ... [i]f all I care about is chocolate, it would be no better to use the list of places that both sell chocolate and are likely to than to use the list of places that sell chocolate.

What the example demonstrates is that “truth plus likelihood of truth is not preferable to truth alone” (21). Kvanvig draws the lesson that “one better not identify justification with statistical likelihood of truth” (21).

So far so good. It is still unclear, though, how this is supposed to be relevant to process reliabilism, which does not equate knowledge with true

---

become false at a later point in time. For instance, the sentence “Ronald Reagan is the president of the United States” was once true, and yet now it is false. Rather than focusing on retaining true beliefs, it is better, so the objection goes, to be adapted to the environment in the sense that one, at any given time, believes what is true and is ready to stop believing what is, or has become, false. This objection hinges on the use of sentences that are temporally indeterminate, i.e. time is not an explicit part of the sentence but supplied by the context of utterance. The answer to the objection is that we can always switch to what Quine (1960, pp. 191-195) called eternal sentences in which time, place etc are explicitly mentioned, without suffering any theoretical loss. If we do, the truth values of sentences will no longer be sensitive to the time of utterance. What was once true will always be true. Retaining true beliefs is therefore to be recommended after all.

belief that is statistically or objectively likely to be true. In anticipation of this objection, Kvanvig writes:

After all, if objective probability itself succumbs to the swamping problem, why would the fact that there is an etiological relationship to a process or method responsible for that probability relieve the theory of the problem? Such a causal relationship to methods or process doesn't seem to be the kind of feature that adds value beyond the value of true belief, so there is no apparent reason here to think that ordinary process reliabilism is in a better condition with respect to the swamping problem than is the simple objective probability theory [i.e. the theory that equates knowledge with true belief that is objectively likely to be true].

Kvanvig goes on to note that both the CP solution and the solution that focuses on value autonomization “go beyond the identification of justification with objective likelihood of truth, and thus provide some hope of avoiding the swamping problem” (23). However, Kvanvig finds that this initial hope quickly vanishes on closer inspection. His lack of enthusiasm for the CP solution is reflected in the following passage:

Once we appreciate the nature of the swamping problem as a problem concerning properties of belief that are non-additive of value in the presence of true belief, it becomes hard to see how the above proposal is helpful at all. In the analogy involving chocolate, I don't even know how to begin thinking about applying this idea to new businesses of the same type, conditional on the first list (places that sell chocolate) and the third list (places that both sell chocolate and are likely to).

The problem with Kvanvig criticism is that it is question-begging in the present argumentative context. Taking the chocolate analogy for granted, as



Kvanvig does, involves focusing at the outset exclusively on the objective likelihood of the belief produced and abstracting from everything about a reliable process that doesn't amount to that process indicating the truth of that belief. According to the CP solution, by contrast, reliable acquisition is indicative not only of the truth of the belief thus produced, but also of the future acquisition of true beliefs (of the same general kind). Excluding, as it does, the CP solution from the start, Kvanvig's analogy cannot be taken as a point of departure in a neutral examination of the possible merits of that proposal.<sup>6</sup>

Does this mean that the CP solution for some reason is inapplicable to cases involving beliefs about where to buy chocolate? Not at all. The CP solution entails that, if you obtained your belief about where to buy chocolate from a reliable website, you are more likely to acquire further true beliefs about where to buy chocolate in the future, the reason being, roughly speaking, that you are likely to use the same website again which is likely to remain reliable.

#### **4. Do we value future true belief?**

The purpose of this section is to discuss two objections that were raised by Isaac Levi in response to the CP solution.<sup>7</sup> In Levi's view, the value problem, as it was posed by Plato, is an artificial one, the reason being that there does not seem to be any context where there is a genuine choice to be made between knowledge and mere true belief. This is especially true from a first-person perspective, for it doesn't appear that a person could be faced with a decision whether to know or merely truly believe that *p*. If so, why should we care which is more valuable? How can a solution to the value

---

<sup>6</sup> For a discussion of the chocolate example, see also Goldman and Olsson (in press) and Olsson (to appear).

<sup>7</sup> Personal communication. What follows is an excerpt of a longer email correspondence.

problem, whatever it turns out to be, matter in our inquiries and deliberations?

As it stands, this is a general objection to the fruitfulness of trying to account for the extra value of knowledge which, as such, is not directed specifically at the CP solution. I will respond to it anyway. I am willing to grant that, from where X is standing, there is no genuine choice between knowing that p or merely believing truly that p. This may be due to the fact that a person cannot coherently grant both that p is true and that her belief has no reliable basis. From her perspective, all her (true) beliefs are cases of reliabilist knowledge.<sup>8</sup> Rather, my first move will be to question the validity of the inference from the lack of first-person concern to the lack of real import.

My approach to the value problem is thoroughly externalist. The fact that a creature has knowledge at a particular instance improves the prospects for that creature to acquire more of the same in the future. This fact need not be something that the creature in question appreciates or is even intellectually capable of appreciating. But it is a fact nonetheless and, one might add, a valuable fact. The ultimate value in having true beliefs is plausibly practical: having true representations – now and in the future – allows a creature to navigate successfully in its environment which is of obvious survival value, whether the creature in question is aware of it or not.

But let us grant, for the sake of the argument, that what is valuable to us must be something that we are, or can be, consciously concerned with in our own inquiries and deliberations. Even so, what the objection leaves out is that, while the difference between X's knowing reliably that p and X's merely believing truly that p may be elusive, or even non-existent, from X's own point of view, there is clearly a difference from another person's, Y's, perspective. Y can note that X's true belief that p was not reliably acquired.

---

<sup>8</sup> As I argue in Olsson (2004), F. P. Ramsey, the first modern reliabilist, seems to have held this view.

Y can then reason, in accordance with the CP solution, to the conclusion that X would be better off had he based his belief on a reliable method. For then X would have been more likely to obtain further true beliefs in response to similar inquiries in the future.

Y may now decide to inform X about the unreliability of the process leading up to X's belief and inform X about a reliable method to arrive at the same belief. So, while it may be difficult, if not impossible, for X to appreciate the difference between knowledge and mere true belief and hence difficult, if not impossible, for X to make that difference matter in his conduct, it is quite possible for another person Y not only to appreciate the difference but to let that difference influence his conduct.

Levi, however, thinks that this reply will not work. The reason, he submits, is that we attribute no value to obtaining future true beliefs. Indeed, in his 1980 and 1991 books, Levi rejected, on independent grounds, what he calls "messianic realism", i.e., the view that we care now about avoiding error or gaining truth not only at the next stage of inquiry but also further down the line.<sup>9</sup> This is not the place for a detailed scrutiny of Levi's complex argument against this view. I will confine myself to making an observation that I believe sheds *prima facie* doubt on the proposal that it would be unreasonable, in all cases, to attach value to the prospect of arriving at true beliefs in the future. At least, it shows that that proposal is in serious need of clarification.

Suppose X is embarking on a journey to Larissa. On the way there are three junctions. In order to get to Larissa, X must take the correct road at each junction. X knows this, but this, alas, is all he knows. In particular, X doesn't know in advance what road to take at each junction. X now

---

<sup>9</sup> Cf. Levi (1991), p. 161: "I suggest we distinguish between being concerned to avoid error *at the next change in belief state* and being concerned to avoid error *not only at the next stage but at  $n (>2)$  or all stages down the line* ... I shall call those who advocate avoidance of error as a desideratum of inquiry in the first sense *secular realists* and those who favor one of the other two sense *messianic realists*."

confronts the first junction, asking himself whether he should choose “left” or “right”. This is his first inquiry. X knows, at this point, that he will face two more inquiries of a similar kind at the two remaining junctions. If Levi is right, however, the inquirer attributes no value to obtaining further true beliefs in response to those remaining inquiries. But this seems implausible. X, by assumption, wants to go to Larissa. But in order to do so he must make the right choice at each of three junctions. Surely, then, X cares about each and every one of these choices now. *Pace* Levi, X desires now to make the right choice at each junction.

Let us add some more details to the example in an attempt to make it even more compelling. X now decides to ask a local resident, Z, about the way. Z tells X to take the road to the right, which indeed is the correct choice, and volunteers to accompany X to Larissa. Believing Z to be a reliable local guide, X agrees. Let us compare the following two scenarios: First scenario: Unbeknownst to X, Z is unreliable and his recommendation a mere lucky guess. Hence, X’s true belief, arrived at via Z’s testimony, that the road to the right leads to Larissa is not a case of reliabilist knowledge. Second scenario: Z’s recommendation is not only true but was also reliably acquired.

Clearly, X is better off in the second scenario than he is in the first. Because Z used a reliable method to find the way at the first junction, Z is likely to have available the same method (map, memory etc.) at the other two junctions, and he is likely to use that method there as well. Finally, since the method used is likely to remain reliable, its reemployment is likely to lead to further correct recommendations, so that X will eventually reach his destination. The same could not be said of the first scenario. Z, to be sure, is likely to reemploy the same method but, being unreliable, that method is relatively likely to yield one or more false recommendations at the remaining junctions.

We can test our intuitions as follows: Focusing once more on the first scenario, let us assume that, at the first junction, X's friend Y comes forward pointing out that, while Z will get it right this time, he will do so as a matter of luck. If Levi were right in saying that X should not care about future inquiries, X should not find Y's remark about Z at all relevant. Since Y gives the right recommendation now, Y's unreliability could only be a matter of concern in the future, and, says Levi, we should not care now about our future inquiries. But, to repeat, this sounds wrong in my ears. Of course X would find Y's remark directly relevant to his concerns. It should make X strongly hesitant to accept Z's offer to help.<sup>10</sup>

### **5. Does the CP solution succeed in identifying a genuine (surplus) value?**

Granted that we value future true belief, does the fact that reliabilist knowledge raises the probability of future true belief really make such knowledge more valuable? If so, what kind of value are we talking about? These issues are raised in a recent paper by Markus Werning.<sup>11</sup> There is much in Werning's carefully crafted paper that I agree with and find helpful. In particular, I consider his own proposals, as developed at the end of his article, insightful and intriguing, though I have yet to figure out exactly how those suggestions fit with my own thinking on the matter. However, I am not convinced by Werning's criticism of CP solution. Let me try to explain why.

First of all, I agree with Werning that the value which that solution attributes knowledge is not of an instrumental kind. In other words, Werning

---

<sup>10</sup> In our email communications, Levi replied to this argument of mine by referring to a manuscript he had recently written in response to Edward Craig's 1990 book *Knowledge and the State of Nature*. The manuscript, entitled "Knowledge as True Belief", will be published in a forthcoming book which I am editing with Sebastian Enqvist called *Science in Flux*.

<sup>11</sup> See Werning (2009).

is correct to observe that the value adding mechanism which is at work here is not subsumable under what he calls “the means-to-end relation” (nor have Goldman or I claimed it is). Nor is it a case of intrinsic or final value. If, as Werning seems to take for granted, instrumental value is the only legitimate kind of extrinsic (non-final) value, it would follow that the CP solution fails to account for the added value of knowledge, which is precisely the conclusion drawn by Werning.

However, the view that instrumental value exhausts the domain of extrinsic value is implausible or at least highly controversial. There is, among other things, also what we may call “indicator value”, the value accruing to something in virtue of its indicating something good. This is acknowledged by leading value theorist like Michael Zimmerman, who writes in his entry on intrinsic and extrinsic value in *Stanford Encyclopedia of Philosophy*:

Many philosophers write as if instrumental value is the only type of extrinsic value, but that is a mistake. Suppose, for instance, that the results of a certain medical test indicate that the patient is in good health, and suppose that this patient's having good health is intrinsically good. Then we may well want to say that the results are themselves (extrinsically) good. But notice that the results are of course not a means to good health; they are simply indicative of it. (Zimmerman, 2002)

The idea behind the CP solution is that knowledge has extra value precisely in that sense: its extra value derives from its indicating something good, namely future true beliefs which are things having final value in the veritist

framework. This, again, is not a kind of instrumental value, but it is still a kind of extrinsic value.<sup>12</sup>

There is another claim by Werning that I would like to challenge, though somewhat less confidently. At one point, he writes: “[t]he extra value of reliably produced true belief as opposed to simply true belief cannot be accounted for by the means-to-end relation”. Werning may have meant to say only that the CP solution doesn’t account for the extra value of knowledge by relying on a means-to-end relation. Thus interpreted, he is of course right. But it is still worthwhile to consider the claim he is actually making, which I tend to believe is at best only partly correct. I tend to think that the extra value of knowledge can, in some degree at least, be accounted for by the means-to-end relation. This may come as a surprise considering the fact that I have emphasized the non-instrumental character of CP solution, a solution I endorse. The explanation is that I have, as I mentioned earlier, another account of the value of reliabilist knowledge according to which such knowledge receives surplus value from making the true belief in question more stable (which is good from the point of view of acting successfully over time). Although I usually spell out what I just wrote in probabilistic terms, I am inclined to think that it is true in a sense of “making more stable” that is stronger than a mere conditional probability relation, presumably to the point of expressing a means-to-end relation of the kind Werning would like to see. The two accounts – CPS and the stability solution – are compatible, and I view them as complementary. If my conjecture about reliabilist knowledge being instrumentally valuable with respect to stability is correct, the extra value of knowledge can be accounted for by the means-to-end relation, but that wouldn’t be a complete account.

---

<sup>12</sup> I take the opportunity to express my gratitude to my colleagues Wlodek Rabinowicz and Toni Roennow-Rasmussen for originally drawing my attention, in personal communication, to the distinction between instrumental and indicator value, and for suggesting that the kind of extra value that CPS attributes knowledge is of the latter sort.

## 6. Does the CP solution solve the general value problem?

According to several authors, among them Kvanvig (2003) and Pritchard (2007), there are in fact two value problems to consider. One is the traditional Platonic value problem of showing that, despite some indications to the contrary, knowledge is more valuable than mere true belief. We may refer to this as the *special value problem*. There is also what we may call the *general value problem*, namely that of establishing that knowledge is more valuable than ignorance.<sup>13</sup> To see that these two problems are indeed different, consider a full-blown reliabilist account of knowledge according to which knowledge amounts to reliably acquired true belief that satisfies an anti-Gettier clause. Solving the special value problem for this account by showing that knowledge, thus construed, is more valuable than mere true belief does not automatically mean solving the general value problem (for that account). For even if knowledge, thus understood, is more valuable than mere true belief, it could still be the case that such knowledge turns out to be no more valuable than reliably acquired true belief that does not satisfy the anti-Gettier clause.

In the rest of the section I will focus, as I did in the formulation of the swamping argument, on simple reliabilism, i.e. reliabilism without an anti-Gettier clause. Several commentators have argued that, while the CP solution may solve the special value problem for simple reliabilism, it fails to solve the general value problem for that account. According to one critic, an opponent to simple reliabilism may concede that “true belief plus the existence of a reliable method” is better than true belief only, and that this solves the special value problem for simple reliabilism, but still deny that this is very helpful for the purposes of defending simple reliabilism against

---

<sup>13</sup> This distinction corresponds to Pritchard’s (2007) distinction between the *primary* and *secondary* value problems.



a more general swamping objection. For that more general purpose, the advocate must also claim that “true belief plus existence of a reliable method together with the fact that the belief was acquired through that method” is better than just “true belief and the existence of a reliable method”. But the CP solution does not seem to give any support for the evaluative significance of the part about how the belief was acquired.<sup>14</sup>

More precisely, the claim is that the probability of future true belief of the same kind as *p* will be more or less the same whether we conditionalize on “true belief acquired through a reliable method” or merely on “true belief plus the existence of a reliable method”. This would mean that, when it comes to being indicative of future true belief, there is no difference between simple reliabilist knowledge and what is less than such knowledge, i.e., ignorance.

But I believe that there is a probabilistic difference in the two cases. The reason why it seems reasonable to believe that conditionalizing on “true belief acquired through a reliable method” significantly raises the probability of more true beliefs of the same kind is, to repeat, that certain empirical conditions are plausibly satisfied, at least in our world. According to one of them, cross-temporal access, a method that was used once is often available when similar problems arise in the future. According to another, the learning assumption, a method that was once employed in an unproblematic fashion will tend to be reemployed on similar occasions. Crucial here is obviously the actual use or employment of the method in question (again, taking “use” in an externalist sense, see the next section). The mere fact that the method “exists”, whatever that means more precisely, seems insufficient for the purpose of raising the probability of future true belief with reference to these empirical regularities.

---

<sup>14</sup> This objection was legitimately raised by Erik Carlson in response to an earlier version of the CP solution (personal communication).

There is another objection in the same category that I would also like to say a few words about. As a preliminary, note that to acquire a *false* belief in a reliable manner makes good sense on the common interpretation of reliability as not necessarily amounting to complete reliability. By saying that a method was reliable we usually mean to say merely that it is sufficiently reliable for the purposes at hand, i.e. that it is reliable at least to degree  $d$  where  $d$  is a contextually determined threshold parameter. For instance, vision at close range counts as a reliable process in most contexts even if it happens occasionally to give rise to false beliefs. Now reliably acquired true belief, according to the objection, does not seem to be any more valuable than reliably acquired false belief on the CP solution. For the probability that  $S$  acquires more true beliefs in the future does not seem to be affected by whether we conditionalize on “ $S$  reliably acquired a *true* belief” or on “ $S$  reliably acquired a *false* belief”. In both cases  $S$  used a reliable method. By the cross-temporal and learning assumptions, there is a significant likelihood that  $S$  will reuse the method, thus probably acquiring more true beliefs. Whether the belief that was produced on the first occasion is true or false is apparently immaterial to the probabilistic argument upon which the CP solution is based. Therefore, if the value of a state depends merely on the extent to which that state is indicative of future true belief, reliably acquired true belief is no more valuable than reliably acquired false belief, and hence no more valuable than reliably acquired belief.<sup>15</sup>

Compelling as it sounds, there are two difficulties pertaining to this objection to the CP solution. First, one may question whether the probability of future true belief is actually unaffected by the truth value of the belief produced on the first occasion. To see why this is so, it is important to consider the exact formulation of the learning assumption that underlies the CP approach. According to that assumption, when spelled out in full, a

---

<sup>15</sup> This objection was raised by a member of the audience at the Goldman workshop in Dusseldorf in May 2008.

method will tend to be reemployed when the occasion presents itself provided that the previous employment was unproblematic. A problematic employment would be one resulting in some sort of conflict or surprise. In this respect there seems to be a difference between true and false belief. The fact that the belief produced on the first occasion was false makes it more likely that the agent's future expectations will be frustrated, making the learning assumptions inapplicable. To illustrate, if, on its first occasion of use, a generally reliable GPS device wrongly recommends taking a left turn in order to get to Larissa, there is some probability that this will lead to a disappointing result (e.g. a dead end street) which will lower the chance of a reemployment of the same method of navigation in the future (relative to what would have happened had the recommendation been correct).

Second, and more obviously, the value of a state is not merely a function of how good the state is as an indicator of future true belief. Some states have final value. A state of true belief is a case in point since such a state, according to vertism, has a value that is not derived from other valuable states. This means that the epistemic value of a state depends on (at least) two factors: (1) the final epistemic value of that state, and (2) the value that state has as an indicator of future states of final value. Even if we grant that reliably acquired true belief and reliably acquired false belief are on a par with respect to their capacity for indicating future true belief (an assumption that we just found wanting), there is still a substantial difference in the first value aspect. A state of reliably acquired true belief has a final value that a state of reliably acquired false belief lacks. It could even be argued that a state of reliably acquired false belief has negative veritistic value. In any case, the total value accruing to a state of reliably acquired true belief is greater than the total value accruing to a corresponding state of reliably acquired false belief.

## **7. Is the CP solution compatible with externalism?**

According to externalist epistemology, of which process reliabilism is but one expression, knowledge does not depend on higher order reflective capacities but can be present also in animals and young children. In order for a subject to know something, it is sufficient that it have a belief or representation of the situation that is true and, in the case of reliabilism, was acquired in a de facto reliable manner. It is not a requirement that the subject has higher level insight into these matters. It is, for instance, not required that the subject mentally represent the belief acquisition process, let alone represent it as reliable.

Consider now the following explanation from Goldman and Olsson (2009) of the learning assumption in connection with the problem of navigating from point A to point B: “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.” This talk about using a given method, having reasons to believe that it should work again and deciding, in that case, to rely on it sounds uncomfortably internalistic and raises the question whether the CP solution is really compatible with externalism.

Why? The CP solution is certainly compatible with an externalist *analysis of knowledge*. If providing an externalist account of knowledge is taken to exhaust the externalist epistemological project, then clearly there is not much to complain about. But externalist epistemology may be taken in a narrower sense to include as well certain preconceptions about the value of knowledge. It is consonant with externalism, one could hold, to insist not only that lower level creatures can know things but also that they are in a position to enjoy the distinct benefits of their knowledge. If, by contrast, the empirical assumptions underlying the CP solution turn out to be essentially

internalist assumptions, the resulting surplus value would be attainable only for introspectively relatively advanced knowers.<sup>16</sup>

In reply, I would like to point out, first, that the conditions of non-uniqueness, cross-temporal access, learning and generality were introduced in Goldman and Olsson (2009) in connection with the higher-level cognitive task of human navigation using a GPS device. In that context, it was natural – though admittedly slightly misleading – to formulate them in relatively high-level terms.

Second, and more to the point, in the paragraph immediately following the GPS example we went on to explain more carefully the role played by the empirical assumptions:

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.

---

<sup>16</sup> Christoph Jäger raised this interesting objection in his talk at the 2008 Dusseldorf workshop on Goldman.

Here, the empirical assumptions underlying the CP solution are formulated in terms of (objective) likelihoods and without recourse to non-eliminable internalistic vocabulary. For instance, there is no talk here of either reasons or decisions. Admittedly, the learning assumption is still spelled out in terms of the likelihood that the subject S will *make use* of the method M. This reference to the use of a method, however, is not intended to imply any conscious reflection on S's part. It is the same use of "use" as in "The cat uses its teeth to clean its claws". I take it, therefore, that the CP solution is compatible, after all, not only with the broader but also with the narrower conception of externalism.<sup>17</sup>

## 8. Is the central CP claim really true?

In this penultimate section I will discuss, at some length, a clever and thought-provoking objection raised by Joachim Horvath. The objection is useful to study more carefully because of the light it sheds on the CP solution and its implications.<sup>18</sup>

Horvath states the central CP claim succinctly as follows:

$$(CPC) \quad P(F/SKp) > P(F/SBTp)$$

---

<sup>17</sup> This is not quite the whole story. As I mentioned, I also believe that the fact that a person knows, rather than merely believes, that p makes that person's belief that p more stable. Stability, moreover, is practically valuable when acting over time. However, as explained in Olsson (2007), the extra value due to stability hinges on the satisfaction of a condition of *track-keeping*. The individual has to keep some sort of record of where she got a given belief from. The condition of track-keeping is a (modest) internalist condition. The upshot is that an individual, in order to enjoy the *full* benefits of knowing, needs to be equipped with a, perhaps rudimentary, capacity to reflect upon her processes of belief acquisition. Whether "reflect" is too strong a word here is something that I choose to leave open for further investigation.

<sup>18</sup> See Horvath (2009). In his manuscript, Horvath has already taken some of our previous personal communications into account. Rather than reiterating those parts I will confine myself to what I take to be the most central remaining issue.

Here F stands for “future true beliefs of a similar kind”, SKp for “S knows that p”, and SBTp for “S merely truly believes that p”. As part of the stage-setting Horvath now makes the following remark:

How are we to evaluate such a contrastive claim about two objective probabilities? Suppose that S actually knows that p. Then, it will be contrary to fact that S merely truly believes that p in the same situation. So, we have to imagine how probable more future true beliefs of the same kind would be for S, if S did not know but merely truly believed that p in that situation. Thus, Goldman’s and Olsson’s contrastive probability claim (CPC) seems to have an implicit counterfactual dimension, for it can never be true of any epistemic subject S that she knows that p and merely truly believes that p in the very same situation.

Let us pause here for a moment. I agree with Horvath that there is a counterfactual dimension to (CPS). The reason, which may not coincide entirely with that given by Horvath himself, is that (CPS), as I think of it, is something of a natural law, and natural laws, as Hempel and others have observed, support counterfactual statements.<sup>19</sup> For instance, the law asserting that iron melts when heated supports the statement that this piece of iron would melt, if it were subjected to heat. Similarly, the proposed law asserting that the probability of S’s obtaining further true beliefs is higher conditional on S’s knowing that p as opposed to conditional on S’s merely believing truly that p entails the following counterfactual: if this proposition p were known, as opposed to merely truly believed, by S, the probability of S’s obtaining further true beliefs would be higher. Equivalently,

---

<sup>19</sup> See, for instance, Hempel (1965), p. 339.

(CC) If the proposition *p* were merely truly believed, as opposed to known, by *S*, the probability of further true beliefs would be lower.

Horvath now argues that, given some reasonable assumptions, this counterfactual is in fact false. This, if correct, would indeed be damaging to the CP solution, as it would entail that (CPC), the central CP claim, were also false.

Horvath's first dialectical move as part of his attempt to demonstrate the falsity of (CC) is to make the following observation:

Human subjects typically have a number of reliable belief-producing mechanisms at their disposal, like perception or memory, as well as a number of unreliable belief-producing mechanisms, like biased reasoning or testimony based on an unreliable informant. In order to evaluate (CC) according to the standard Lewisian semantics of counterfactuals, we need to consider the closest possible world where its antecedent is true, that is, the next possible world where *S* does not know that *p* but merely truly believe that *p* – or, more precisely, the next possible world where *S* has an unreliably produced true belief that *p* instead of a reliably produced true belief that *p* like in the actual world.

Horvath now introduces a further assumption to the effect that, in the actual world, *S*'s knowledge that *p* is based on *S*'s reliable faculty of visual perception. He continues:

Now, the closest possible world where *S* truly believes that *p* based on an unreliable mechanism does not seem to be a world where *S*'s visual capacities are seriously impaired, for this would be a rather drastic change compared to the actual world. More plausibly, the closest world where *S* has an unreliably produced true belief that *p* might be one, for



example, where S bases his belief that p on the testimony of an unreliable informant of his acquaintance (or on one of the other unreliable mechanisms that are actually at her disposal).

As Horvath proceeds to note, it is a plausible empirical assumption that we can acquire most of our beliefs in more than one way. But if so, then

... the closest possible world where the antecedent of (CC) is satisfied will typically be one where S disposes of the very same mechanisms of belief-production as in the actual world, but simply makes use of one of her unreliable mechanisms, like the testimony of an unreliable informant, instead of one of her reliable mechanisms, like visual perception. However, in such a world it is not the case that the probability of more future beliefs of the same kind is lower than 0.8, that is, lower than the corresponding actual probability. For, in a close possible world like that, all of S's actual belief-producing mechanisms are still present in an unmodified way.

Hence,

... given my plausible empirical assumptions about normal human subjects, the probability of more future true belief is typically not lower conditional on merely truly believing that p than conditional on knowing that p. But then, reliably produced true belief does not have the extra value of making future true belief more likely than unreliably produced true belief in actual human subjects.

As Horvath reports, I have in our previous correspondence objected to the implicit assumption that, in the closely possible world where S truly believes that p based on an unreliable mechanism is either one where S's

visual capacities are seriously impaired or one in which S makes use of another mechanism altogether. There is also the possibility that S's visual capacities are impaired but only slightly. They could be impaired only to the extent required for the reliability of those capacities to fall below a given reliability threshold. Clearly, this need not entail *serious* impairment.

The point I would like to raise now, however, is a different one. Let us return to Horvath's claim that since, in the second scenario in which S uses a different method that is in fact unreliable, all of S's actual belief-producing mechanisms are still present in an unmodified way, the probability of future true belief will be just as high as in the knowledge scenario. I doubt that this claim is true. What Horvath overlooks is that it may well be a significant fact that S in order to solve the problem at hand chooses to ask an in fact unreliable witness rather than to take a look for himself.

To see why this is so, first note that the empirical conditions that underlie the CP solution must be assumed true in the actual "knowledge" world. Moreover, since they have the status of laws, worlds in which they don't hold are relatively remote from the actual world. This means that, in the alternative world we are now considering, let us call it W, the assumptions of non-uniqueness, cross-temporal access, learning and generality still hold true.

Let us consider more carefully the world W in which S relies on the in fact unreliable witness. By the non-uniqueness assumption, S is likely to encounter the same kind of problem again. By cross-temporal access, the same method is likely to be available then, which in this case means asking the unreliable witness. By the learning condition, S is now likely to use the same method again. But since the method is unreliable, chances are now that this will lead to a false belief. At least this is more likely than if S had had knowledge based on visual perception (which we assume, with Horvath, to be reliable to the same degree in the scenarios under consideration).

Hence, if we compare the two scenarios – knowledge versus mere true belief – at the second shot, it is more likely that a true belief will ensue in the knowledge scenario than it is in the mere true belief scenario. I see no reason, therefore, to think that the probability of further true belief would be no higher in the knowledge scenario as compared to the true belief scenario. The CP claim still seems to me to be, in Horvath’s words, “clearly and determinately true”.

## **9. Conclusion**

According to the CP solution to the swamping problem, reliabilist knowledge has surplus value in virtue of indicating not only the truth of the belief that was acquired but also the truth of future beliefs adopted in response to similar problems. This holds to a lesser degree of a mere true belief. Since it is valuable to have true beliefs in the future, this proposal, in my view, solves the longstanding value problem first raised in Plato’s dialogue *Meno* for the attractively simple theory of process reliabilism. I am grateful for this opportunity to respond to some of the objections that have been raised against this proposal, which has forced me to think more carefully about its consequences. So far, however, I have not seen any objection to that approach that survives critical scrutiny.

Acknowledgements: I benefited greatly from the input I received from speakers and members of the audience during the Dusseldorf Goldman conference in May 2008. Apart from Alvin himself, who is always a great source of inspiration and insight, and the people I have already credited, I would like to thank, in particular, Peter Baumann, Thomas Grundmann, Christian Piller, and Gerhard Schurz.

## **References**

- Craig, E. (1990), *Knowledge and the State of Nature*, Oxford: Oxford University Press.
- Goldman, A. I. (1979), "What is Justified Belief?" in G. Pappas (ed.), *Justification and Knowledge*, Dordrecht: Reidel. Reprinted in A. Goldman, *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge, MA: MIT Press (1992).
- Goldman, A. I. (1986), *Epistemology and Cognition*, Cambridge, Mass.: Harvard University Press.
- Goldman, A. I. (1999), *Knowledge in a Social World*, Oxford: Oxford University Press.
- Goldman, A. I. (2002), "The Unity of the Epistemic Virtues", in *Pathways to Knowledge*, New York: Oxford University Press.
- Goldman, A. I., and Olsson, E. J. (in press), "Reliabilism and the Value of Knowledge", in Haddock, A., Millar, A. and, Pritchard, D. H. (eds.), *Epistemic Value*, Oxford University Press.
- Hempel, C. G. (1965), *Aspects of Scientific Explanation: and Other Essays in the Philosophy of Science*, New York: Free Press.
- Horvath, J. (2009), "Why the Conditional Probability Solution to the Swamping Problem Fails", *Grazer philosophische Studien*.
- Kvanvig, J. L. (2003), *The Value of Knowledge and the Pursuit of Understanding*, Cambridge: Cambridge University Press.
- Kvanvig, J. L. (2008), "Further Thoughts on the Swamping Problem", *Social Epistemology*.
- Levi, I. (1980), *The Enterprise of Knowledge*, Cambridge, Mass.: The MIT Press.
- Levi, I. (1991), *The Fixation of Belief and Its Undoing*, New York: Cambridge University Press.
- Levi, I. (to appear), "Knowledge as True Belief", in Olsson, E. J., and Enqvist, S. (eds.), *Science in Flux*, Springer.

- Olsson, E. J. (2004), "F. P. Ramsey on Knowledge and Fallibilism", *Dialectica* 58 (4): 549-557.
- Olsson, E. J. (2007), "Reliabilism, Stability, and the Value of Knowledge", *American Philosophical Quarterly*, vol. 44 (4): 343-355, 2007.
- Olsson, E. J. (2008), "Knowledge, Truth, and Bullshit: Reflections on Frankfurt", *Midwest Studies in Philosophy* XXXII: 94-110.
- Olsson, E. J. (to appear), "Reply to Kvanvig on the Swamping Problem", *Social Epistemology*.
- Pritchard, D. H. (2007), "Recent Work on the Value of Knowledge", *American Philosophical Quarterly*.
- Quine, W. V. O. (1960), *Word and Object*, The MIT Press, Cambridge, Mass.
- Swinburne, R. (1999), *Providence and the Problem of Evil*, Oxford: Oxford University Press.
- Werning, M. (2009), "The Evolutionary and Social Preference for Knowledge: How to Solve Meno's Problem within Reliabilism", *Grazer philosophische Studien*.
- Williamson, T. (2000), *Knowledge and Its Limits*, Oxford University Press.
- Zagzebski, L. (2003), "The Search for the Source of Epistemic Good", *Metaphilosophy* 34: 12-28.
- Zimmerman, M. (2002), "Intrinsic vs. Extrinsic Value", *Stanford Encyclopedia of Philosophy*, Summer 2009 Edition, Edward N. Zalta, URL: <http://plato.stanford.edu/entries/value-intrinsic-extrinsic/>, the paper was first published in 2002 and substantially revised in 2007.