The Value and Expected Value of Knowledge

JULIEN DUTANT University of Geneva

ABSTRACT: Meno's Thesis—the idea that knowing something is better than merely having a true belief about it—is incompatible with the joint claims that (a) believing the truth is the sole source of the value of knowledge and (b) true belief and knowledge are equally successful in believing the truth. Recent answers to that so-called "swamping" problem reject either (a) or (b). This paper rejects Meno's Thesis instead, as relying on a confusion between expected value and value proper. The proposed solution relies on an externalist view of rationality, which is presented.

RÉSUMÉ: La thèse du Ménon—l'idée que savoir quelque chose est meilleur que d'avoir une croyance vraie à ce sujet—est incompatible avec la conjonction des affirmations que (a) croire le vrai est la seule source de valeur de la connaissance et (b) la croyance vraie et la connaissance ont des niveaux égaux de réussite en termes de croire le vrai. Des réponses récentes à ce «problème de la submersion» rejettent (a) ou (b). Cet article rejette au contraire la thèse du Ménon, comme reposant sur une confusion entre valeur espérée et valeur proprement dite. La solution proposée repose sur une théorie externaliste de la rationalité, qui est présentée.

Knowledge is a good thing. That fact appears readily explainable. Knowing something entails having the truth. Having the truth is a good thing. So knowing is a good thing.

This explanation faces a difficulty, as pointed out by Plato in the *Meno* (97b–d). One may have true belief without having knowledge. One might have guessed rightly, for instance. But having a true belief is no less having the truth than knowing. So, knowing would not be better than merely having a true belief.

Dialogue 51 (2012), 141–162.

© Canadian Philosophical Association/Association canadienne de philosophie 2012 doi:10.1017/S0012217312000285

Yet knowledge *does* appear better than true belief. Knowing that a storm is a coming, for instance, appears better than merely guessing that a storm is coming. The simple explanation appears to fail.

We have the makings of a paradox: knowledge is good only because it is good to have the truth; merely believing the truth is no less having the truth than knowing; yet it is better to know than to have a true belief. Call this *Meno's Paradox*, and call the thesis that knowledge is better than true belief *Meno's Thesis*. ¹

A variant of the paradox—the so-called "Swamping Problem"—has sparked a debate on the value of knowledge in the recent literature. Most of the contributors accept a qualified version of Meno's Thesis, according to which knowledge is at least most often better than true belief. In this paper, I put forward a radically opposed solution which denies that knowledge is better than mere true belief.

The core idea of the proposal is that knowledge has a higher expected value than a mere true belief. More precisely, as far as the value of believing the truth is concerned, the expected value of believing what one believes is higher when one knows than when one has a true belief. The result relies on a certain externalist picture of rationality according to which rational expectations are ones based on knowledge and not simply belief. (This raises the worry that I am assuming what is to be explained, a worry that is defused in due course.) The result explains why one ought to prefer believing when one knows to believing when one merely believes. In that sense, knowledge is preferable to true belief. That, I think, is all there is to the intuition behind Meno's Thesis. Hence, the proposal suggests that Meno's Thesis relies on a confusion between expected value and value proper. It thus puts the burden of proof on those who think that knowledge is not only preferable to true belief, but genuinely better.

The paper is structured as follows. The first section states Meno's Paradox and the Swamping Problem. Section two reviews the existing solutions. They are "liberal" in the sense of upholding Meno's Thesis and in being generous in their ascriptions of value to knowledge. Section three puts forward the "austere" solution I favour which rejects Meno's Thesis entirely.

1. Meno's Paradox on the Value of Knowledge

1.1. Meno's Thesis

Let us start with the idea that knowledge is a good thing. It can be stated as follows:

VK For any fact p, knowing that p is a good thing.

The thesis has some initial plausibility—as opposed to parallel claims about belief or desire, for instance—but it faces counterexamples. Some knowledge is useless. Knowing whether the XIIth century Catalan king, Ramon Berenguer

IV, was faithful to his wife Petrolina of Aragon would not help me with any past, present or future, or practical or theoretical, aim of mine. Some knowledge is uninteresting. Medieval gossip is intrinsically interesting to many, but the same cannot be said about whether one's local directory contains an even number of digits. Some knowledge is harmful, in practical or theoretical respects. Some unlucky witnesses get killed. A misleading piece of information may put an inquiry on a wrong track.

Three mutually compatible amendments of VK can be introduced to deal with the counterexamples. First, we may restrict the claim to interesting facts:

VK-I For any interesting fact p, knowing that p is a good thing.²

If to say that a fact is interesting is just to say that it is good to know it, the claim is tautologous. VK-I plausibly assumes that we have an independent notion of interesting facts. Second, we may make a generic claim instead of a universal one:

VK-G In general, knowing something is a good thing.³

Generality allows for exceptions. What it requires is not entirely clear: it may neither be necessary nor sufficient that most instances of the claim hold. Perhaps normal or typical instances should hold. We can leave the question aside here.⁴ Third, we can construe VK as a claim about *ceteris paribus* value as opposed to value "all things considered".

VK-CP For any p, knowing p is a good thing, everything else being equal.⁵

While knowing something may be a good thing, other things being equal, if it gets one in trouble, it may be a bad thing overall.

One should carefully distinguish *ceteris paribus* value from *prima facie* value. The former is a component of a thing's total value. The second is merely an appearance of value. If a flat is located in a nice area, that makes it valuable, everything else being equal. But if the rent is so high that it would ruin you, it is altogether a bad flat for you. Still its location remains a positive component of its overall value. By contrast, the fact that the flat you rent for a winter holiday is equipped with central heating is *prima facie* a good thing. But upon realizing that the central heating does not work and will not be repaired during your stay, the appearance of value vanishes. It is not as if the fact that it is equipped with central heating still confers it some *ceteris paribus* value that is simply offset by its not working. Rather, the flat appeared to have some value that it does not in fact have. *Prima facie* value is not value at all.

Whether knowledge of uninteresting facts is valuable is unclear. If it is not, then the restriction to interesting facts is needed. Here, we sidestep the issue by

focusing on interesting facts. The generic and *ceteris paribus* amendments are each able to deal with the other counterexamples.

Meno's Thesis is the view that knowledge is better than true belief. Its variants parallel those of the idea that knowledge is a good thing.

1.2. Meno's Paradox

We want the truth. Typically, someone who raises a question does not want to hear a beautiful or coherent story. They want to get the true answer.

But what does "having" or "getting" the truth mean? I write "she loves me" and "she loves me not" on two pieces of paper. Necessarily, I have a piece of paper that states the truth. But I do not thereby "have the truth" in the required sense. Having the truth in the desired sense could be either "believing the truth (and only the truth)" or "knowing". Meno's Paradox arises with the first reading. We assume that believing the truth is good and try to explain on that basis why knowing is.

There are worries with the set-up. Is believing the truth a good thing? That is not obvious. Believing the truth generally favours success in action, but that is neither universal nor necessary. Success in action itself is at best valuable *ceteris paribus* or generally. One may think that believing the truth is fundamentally good, but if one is entitled to that claim, why not just as well claim that knowing is fundamentally good? If it is, we need not explain its value in terms of the value of true belief.⁶

These worries may be put aside dialectically. Meno's Paradox challenges the simple explanation of the value of knowledge. Proponents argue that from the assumption that true belief is good, we cannot derive that knowledge is better. They may deny Meno's Thesis or claim instead that knowledge is irreducibly valuable, or valuable on some other ground. Opponents argue that we can, so that knowledge is better than true belief if the latter is any good. The debate proceeds on the assumption that believing the truth is a good thing. We follow suit here.

Let us state the paradox more precisely. A "condition" is something a subject may or may not be in at a given time and world. Feeling cold, getting to Larissa, knowing p and believing the truth p are examples of conditions. We say that a condition entails another if and only if, necessarily, a subject that is in the first is in the second. We also need a graded condition of having true beliefs. We could say that one subject satisfies this condition more than another if and only if the first has more true beliefs than the second. But counting beliefs is of dubious value. For present purposes, we only need the limited and plausible assumption that if the truths believed by one subject entail those believed by another but not conversely, the first satisfies the condition more than the latter. The paradox is stated:

Evaluation Principle (EP): If conditions C and C' are good solely in virtue of how much satisfaction of graded condition C'' they entail, and if each entails as much satisfaction of C'' as the other, then C and C' are equally good.

Truth Sole Source of Value (TSSV): Knowing that *p* and believing the truth *p* are good solely in virtue of how much satisfaction of the condition of having true beliefs they entail.

Equal Theoretical Success (ES): Knowing that p and believing the truth p entail as much satisfaction of the condition of having true beliefs as each other.

Meno's Thesis (MT): Knowing that *p* and believing the truth *p* are not equally good.

The claims are logically contradictory. TSSV and ES entail that knowing p and believing the truth p satisfy the antecedent of EP, but MT contradicts its consequent. The formulation can be applied to the *ceteris paribus* and "all things considered" readings of "good." A generic version is obtained by substituting generic claims for the unrestricted ones.

The Evaluation Principle seems to me unassailable in all versions. So, one of TSSV, ES and MT has to go. Liberal solutions maintain MT and give up at least one of TSSV and ES. The austere solution I put forward rejects MT.

1.3. The Swamping Problem

In recent literature, a debate on the value of knowledge has been sparked by the so-called "swamping problem." (Jones, 1997, 426; Zagzebski, 2003; Swinburne, 1999, 58; Riggs, 2002; Kvanvig, 2003, 44–9) The problem is a variant of Meno's Paradox that involves a weaker "Truth Sole Source of Value" principle and a stronger "Equal Success" claim.

The "Truth Sole Source of Value" principle is very restrictive. It ascribes value to conditions only in virtue of their necessary consequences. An analogous principle in the practical domain would deny value to taking an aspirin even though it causes my headache to vanish because the effect of aspirin is contingent. The principle denies that something can have instrumental value, insofar as instrumental value is value in virtue of its causal consequences.

It is not obvious that there is instrumental value. We have instrumental reasons to do things that have good consequences. But that does not imply that those things are good; only that they cause something good. So it is open to claim that while the condition of "taking a pill that will cure one's headache" is a good condition, because it entails that one's headache vanishes, the condition of taking a pill itself is not good. Hence, the necessary consequence version of Meno's Paradox is defensible. We leave it aside for the moment, however, and consider whether the paradox survives the introduction of instrumental value.⁸

People who accept instrumental value should be wary of a kind of double counting. To borrow an example from Parfit (2011, 59), suppose there are two apples in a tree. It is equally good for me to have either one, but the higher one can only be reached with a ladder. Using the ladder is therefore instrumentally good. Suppose I am only allowed to take one. One cannot say that it is better to take the higher one, because it has both the value of getting an apple *and* the instrumental value of taking a ladder to get that apple. That is obviously wrong.

The instrumental value that a thing has in virtue of having some good consequence cannot be added to the value of that consequence.

Let's call consequentialism the idea that value is either intrinsic or instrumental. There are several ways to formulate it. Here is one that will do for present purposes: The satisfaction of a condition by a particular subject at a particular time is an event. We remain neutral on whether the event in question exactly consists in the subject satisfying the condition, or is a more coarsegrained entity. Some may want to distinguish direct and indirect causation, but we will ignore the distinction and say that an event causes another if and only if it causes directly or indirectly. If any event has parts, we include among its causal consequences the causal consequences of its parts, provided these consequences are not already part of it. For instance, if the event of "the spark causing the fire" includes both the event of "the spark," which caused the fire, and "the fire" itself, that event did not cause the fire. We assume that no event causes one of its own parts. We say:

Actual-Causal-Consequences Consequentialism (ACCC): The value of an event E is equal to the sum of (a) the values that its parts have fundamentally and (b) the values that its causal consequences have fundamentally. Part (a) of the sum is the intrinsic value of E and part (b) its instrumental value.

The principle avoids double counting. For instance, if E includes F and G, F causes G and G has fundamental value, then F has instrumental value. But when evaluating E, we should not add the instrumental value of F to the value of G, since the former is derived from the latter. The principle avoids that by computing the value of E solely from the value of E. Similarly, if E causes E, E causes E and E has fundamental value, then E has instrumental value. But when evaluating the E, we should not add the value of E to that of E. Again, the principle avoids doing so.

Combined with veritism, the idea that only true belief has fundamental value in epistemic matters, consequentialism yields a consequentialist variant of the "Truth Sole Source of Value" principle. ¹⁰ The resulting version of Meno's Paradox is the "swamping problem":

Evaluation principle–consequentialist (EP–c): If events E and E' are good solely in virtue of the value of events of type V that they include or cause, and if each includes or cause the same events of type V as the other, then conditions E and E' are equally good.

Truth Sole Source of Value—consequentialist (TSSV—c): Events of knowing and of believing the truth are good solely in virtue of the value of the events of believing a truth they include or cause.

Equal Theoretical Success—consequentialist (ES–c): An event of knowing that p and the corresponding event of believing the truth p include or cause the same events of believing the truth as the other.

Meno's Thesis is as before, and the four claims are logically contradictory here as well. Again, EP-c seems unassailable, and one of TSSV-c, ES-c or Meno's Thesis has to go.

A practical version of the paradox results from including good practical consequences as well as true beliefs in the source of value. That appears to be what Plato had in mind in the Meno.

The consequentialist version of "Truth Sole Source of Value" is weaker than the entailment one in the sense that it is more generous in its attribution of value. Accordingly, the consequentialist version of "Equal Success" is stronger: not only are the necessary consequences of knowledge and true belief the same, but their causal consequences are too.

The swamping problem has been aimed at reliabilism. In its simplest version, reliabilism about knowledge is the view that knowledge is reliably formed true belief. (Goldman, 1986) Reliably forming one's belief on some occasion has instrumental value insofar as it causes one to form a true belief. But if knowing p includes both reliably forming a belief about p and having the true belief that p, the value that accrues to reliably forming one's belief in virtue of causing a true belief cannot be added to that of the true belief itself. Provided that one's reliably forming that belief on that occasion does not cause further true beliefs, the Equal Success premise holds. The reliabilist then faces the consequentialist version of Meno's Paradox.

2. Liberal Solutions

Most current answers to Meno's Paradox are liberal in that they endorse at least a qualified version of Meno's Thesis. As we will see, they all reject the entailment version of "Truth Sole Source of Value". They can be further classified into those that reject consequentialist "Equal Success" and those that reject consequentialist "Truth Sole Source of Value".

2.1. Rejecting Consequentialist Equal Success

To reject the entailment version of Equal Success (ES), one has to argue that knowing p necessarily entails having some true belief that believing p does not. The prospects for such a claim are dim. For instance, a "Strong Higher-Order Belief Solution" would claim that knowing entails believing that one knows. Hence, knowing p entails having two true beliefs, whereas having a true belief only entails having one: the true belief that p and the true belief that one knows that p. The solution is implausible since knowing does not require believing that one knows.

Defenders of Meno's Thesis are therefore plausibly committed to rejecting the entailment version of "Truth Sole Source of Value". They may reject either its consequentialist version as well or consequentialist Equal Success. We consider the latter option here. The idea is that knowing has better consequences in terms of having true beliefs than simply having a true belief.

Goldman and Olsson's (2009, 24–31) "Conditional Probability Solution" is one such view. They say that knowing generally makes it more probable that one has further true beliefs than simply having a true belief, in the following sense: the probability of having further true beliefs given that one knows that p in a certain case is greater than the probability of having further true beliefs, given that one believes the truth p in that case. That is so because knowing entails having reliably formed one's belief, which typically goes with being able to reliably form further beliefs in the same way.

The proposal faces two problems. First, a consequentialist should use a notion of probability that is not only objective, as Goldman and Olsson (2009, 28) acknowledge, but also time-sensitive. Before the lottery is drawn, a ticket may have instrumental value because it has some chance of winning, but after the ticket lost, it no longer has any value. On the time-sensitive notion of probability, past has probability one and conditionalising on it has no effect. Hence at the time of knowing, the probabilities given that one knows and given that one has a true belief are the same. Moreover, it is unclear why one should consider the probabilities at some other time, and if so, which.

Second, the proposal's notion of instrumental value fails to distinguish means from symptoms. Suppose I need to take a certain drug that typically causes vomiting. The probability of me taking the drug conditional on me vomiting next week is higher than the unconditional probability of me taking the drug. Yet vomiting next week has clearly no instrumental value for taking the drug. Vomiting next week is an indication or symptom that I have taken the drug, not a means to it. Instrumental value should only accrue to means, not to symptoms. Goldman and Olsson's solution faces a similar worry. Having reliably formed a true belief that p is a symptom of being disposed to form certain beliefs reliably. The disposition itself is a means to forming further true beliefs, not having reliably formed a true belief that p. So while reliable dispositions typically have instrumental value beyond that of a single true belief, that is not obviously the case for instances of knowledge, which are mere symptoms of these dispositions. What Goldman and Olsson need to claim is that particular instances of reliable belief formation themselves cause or are likely to cause some further true beliefs. 12

Fortunately, events of knowing *p* often appear to cause further true beliefs themselves. Such cases can be put forward against consequentialist "Equal Success". Under a common interpretation, Plato's "Tethering Solution" claims that acquiring knowledge that *p* often causes one to maintain one's true belief that *p* in the face of misleading counter-evidence. (Williamson, 2000, 78; Olsson, 2007) Swinburne's (1999, 64) "Higher-Order Belief Solution" claims that when acquiring knowledge, we typically acquire the true belief that we know and related higher-order beliefs. On a "Same-Order Belief" solution, one could point out that in acquiring one piece of knowledge, one typically acquires several related true beliefs. For instance, I acquire knowledge that

there is an apple in a basket by seeing it. In doing so, I acquire true beliefs about the content of the basket, the colour and shape of the apple, and so on.

Analogous remarks can be made about the practical success caused by knowledge to answer practical versions of Meno's Paradox. (Williamson, 2000, 62)

These solutions may be resisted. One could claim that in each case, further true beliefs are caused either by the acquisition of knowledge that p—not by the state of knowing that p itself—or by the corresponding true belief that p. The objection turns on issues about the individuation of events of knowing and criteria for causation that we will not get into here.¹³

If successful, these solutions show that consequentialist "Equal Success" is false. A further step is to argue that in general, knowledge is more successful than true belief. It is not straightforward to do so insofar as what generality requires is unclear. But if it can be done, a generic version of Meno's Thesis can be vindicated.

These solutions cannot establish a universal version of Meno's Thesis, however. It is easy to imagine cases in which one's knowing *p* has no further consequences in terms of true beliefs or practical success. Hence an existential version of consequentialist "Equal Success" is hard to deny. Plugged into an existential version of the Meno's Paradox, it yields the conclusion that in some cases, knowing is no better than the corresponding true belief.

The intuition behind Meno's Thesis appears to be universal. A man on his deathbed wonders whether his son ever loved him. It appears better for him to come to know it on the basis of a reliable testimony than merely to believe the truth about it on the basis of a well-meaning friend's guess. Yet that piece of knowledge does not have any good consequences that the true belief would not have. The present solutions cannot vindicate the judgement that that piece of knowledge is better than the true belief alone.

To sum up, these solutions argue that in general, knowing has good consequences that having a true belief does not have. The claim is plausible and supports a generic version of Meno's Thesis, insofar as the notion of instrumental value makes sense. But the intuition behind Meno's Thesis is a universal one and they fail to vindicate or explain it.

2.2. Denying Truth Sole Source of Value

Other solutions reject the idea that the value of knowledge entirely derives from its entailing or causing true beliefs. One may simply deny that the value of knowledge is at all explained in terms of the value of true belief. The "Fundamental Value Solution" is the simplest such option: on that view, knowledge simply has fundamental value, which is irreducible to that of true belief. However, most solutions in the literature hold onto the idea that the value of knowledge derives from that of true belief. They simply deny that it is explained by entailing or causing it.

Goldman and Olsson's (2009, 31–5) "Value Autonomisation Solution" combines two ideas: "type instrumentalism", according to which something can have instrumental value in virtue of belonging to a relevant type of thing that generally has good consequences, and "value autonomisation", according to which things of instrumentally valuable types sometimes come to be regarded as intrinsically valuable. The view is illustrated with the case of good intentions. A particular good intention, they say, is seen to be good simply because good intentions generally have good consequences, and we end up regarding them as good in themselves. The proposal is that reliable belief formation—and hence knowledge—comes to be regarded as intrinsically valuable, irrespective of whether it causes further true beliefs.

Both aspects of the proposal are problematic. Type instrumentalism seems wrong. (Piller, 2009a, 131–2) If aspirin cannot cure my current headache, then taking an aspirin will not cause it to vanish. On consequentialist views relying on actual causal consequences, taking aspirin has no instrumental value whatsoever in my case. ¹⁴ At most it appears to have value, precisely because it is an action of a type that typically cures headaches. That seems to me right. By contrast, type instrumentalists say that taking the aspirin does have some instrumental value. They have a hard time explaining why I should not bother taking the aspirin after I learn that my headache is of a kind that it cannot affect.

Second, value autonomisation sketches a psychological account of how we end up regarding merely some instrumentally valuable things as valuable in themselves. (Goldman and Olsson, 2009, 34) The account answers the question whether and why we value knowledge better than true belief, not whether knowledge is better than true belief. I find the psychologisation of the problem unhelpful. It fails to distinguish error theories, according to which we value knowledge more than mere true belief but mistakenly so, from vindication theories, according to which we value knowledge more than mere true belief because it is better.¹⁵

The "Value Autonomisation Solution" supports Meno's Thesis only when construed as a vindication theory. To do so, one needs to assume that once we value knowledge intrinsically, knowledge becomes intrinsically valuable. In other words, one should endorse a kind of subjectivism about value. This point is controversial but defensible. At that stage the solution is in effect the "Fundamental Value Solution" supplemented with a genetic and subjectivist story of how knowledge comes to acquire fundamental value. 16

The "Credit Value Solution" defended by Riggs (2002), Greco (2003, 2010) and others has two components. First, they argue that a type of value, namely "credit", occurs when some good arises from an agent's ability. Typical illustrations come from sports: a skilful catch is better than one occurring through a fluke. (Greco, 2003, 122) By the same token, true belief arising from one's ability is better than true belief *simpliciter*. Second, they claim that knowledge is true belief arising from a subject's ability. From this, a universal version of Meno's Thesis follows.

The solution faces three objections. First, not all knowledge may involve true belief arising from ability. Sometimes we want to find answers, but sometimes we simply want to be given answers. In both cases, we would end up knowing, but situations most naturally described in the second way are harder to picture as true beliefs arising from one's abilities. Testimony and calculators provide illustrations. If successful, they force credit theorists to retreat to a generic version of Meno's Thesis.¹⁷ Second, one may doubt whether knowledge implies true belief arising from ability in a sense that yields credit value. Swallowing is a success arising from ability too, and we gain much perceptual knowledge quite like we swallow. One may doubt whether successful swallowing is made better by the fact that it arises from an ability to do it. The doubt may be extended to knowledge we acquire in a quasi-automatic fashion. Again, that would force a retreat to a generic claim—but even that would be threatened if knowledge is typically acquired in such a way. Third, one may reject the very idea of credit value. Alleged cases of credit value are explainable in other ways. In sports, for instance, we value good players at least as much as good results. A skilful catch may appear better because it indicates a good player. By contrast, not all goods become better when brought about by one's ability. Suppose it would be good to clean the terrace in front of my house. If a rainfall happens to wash it, I should not regret not having done it myself. That is so even if the cleaning would have only involved effortlessly pushing a button. Hence, one may doubt that bringing about something good through one's one ability carries any additional value.

Zagzebski's (2003) "Good Motivation Solution" is a restrictive version of the "Credit" view. She takes credit to require a good resulting from one's being motivated by the good. A boy that brings back a lost cat deserves credit if he aimed at being helpful, not if he only aimed at getting a reward. She says that knowledge is valuable when it is desirable true belief resulting from one's virtuous intellectual acts motivated by a love of desirable truth (2003, 24). As she recognizes, the view only supports a generic version of Meno's Thesis. (Ibid.)

Piller's (2009a, 123–6; 2009b, 426) "Deontic Value Solution" claims that (a) one ought to choose what is likely to be good from one's point of view, (b) choosing as one ought to is itself good. Applied to belief, that means that one ought to form beliefs in ways that appear reliable from one's point of view, and that true beliefs so formed will be better than mere true belief.

The solution in effect ascribes value to internally justified true belief, and thus mismatches the universal Meno's Thesis. Moreover, there is a worry that the notion of deontic value is inconsistent or undefined. Start from the totality of what is good, call it Good. One ought to choose what is likely good. On the deontic view, it follows that choosing the likely good is good as well. Now the good is not just Good, but Good and choosing what is likely good. So either Good is not the totality of the good, contrary to assumption, or Good self-referentially includes choosing the likely good, but I doubt such a notion is well-defined.¹⁹

To sum up, we reviewed four main options to deny "Truth Sole Source of Value". The "Fundamental Value Solution" claims that knowledge is a fundamental good. The "Credit Value Solution" claims that knowledge has value as credit for something good, namely true belief. The "Good Motivation Solution" claims that knowledge has value when resulting from a motivation for something good, namely desirable true belief. The "Deontic Value Solution" claims that knowledge is good when it is believing as one ought to believe. The first two can support a universal version of Meno's Thesis. The latter two can at most support a generic version. The "Value Autonomisation Solution" is ambiguous between an error theory and a variant of the "Fundamental Value Solution".

I have mentioned some difficulties for these views, but they are not decisive. The generic version of Meno's Thesis may be defended by denying that the universal one is intuitive or by endorsing some error theory about the intuition behind the universal one—along the lines of the "Value Autonomisation Solution", for instance. The "Fundamental Value Solution" may appear unsatisfactory but I see no major problem with it. "Credit Value" theorists may have resources to answer the objections raised.

My aim here is not to establish that these views are wrong, however. Rather, I want to undercut their central motivation, namely to uphold Meno's Thesis. To do this, I show how one can explain why one ought to prefer knowledge over mere true belief even if knowledge is not better than true belief. That knowledge is to be preferred to mere true belief seems to me all there is to explain in Meno's Paradox. Hence, the explanation puts the burden of proof on those who want to argue that knowledge is not only preferable to true belief, but also better

3. An Austere Solution

In this section, I put forward an austere answer to Meno's Paradox, the "Higher Expected Value" solution. It is austere in the sense that it rejects Meno's Thesis entirely. That part is straightforward: we reject the notion of instrumental value and endorse the entailment version of "Truth Sole Source of Value". As we have seen, the entailment version of "Equal Success" is overwhelmingly plausible. The full denial of Meno's Thesis follows.

The main thrust of the solution is to explain why knowledge ought nevertheless to be preferred to mere true belief. The view relies on a certain externalist view of rationality, which I introduce first. On the basis of that view, and the assumption that believing the truth is good, we can show that knowledge has a higher expected value of a certain kind. The higher expected value of knowledge explains why it ought to be preferred to mere true belief, in a sense that we will specify. That, I claim, is all there is to the intuition behind Meno's Thesis.²⁰

3.1. Two Pictures of Rationality

Let me first contrast two pictures of rationality.²¹ Both share the idea that facts give us reasons to act in certain ways and to be in certain states such as beliefs,

desires or emotions. (In introducing the two pictures we focus on acts, but what we say applies equally to states.) For instance, if eating walnuts would kill you, you have a reason not to eat walnuts, whether you are aware of it or not. (Parfit, 2011, 31) To reasons corresponds an objective use of "ought:" you ought not to eat walnuts, since they would kill you. (Broome, 2004)

The two pictures differ on how rationality relates to reasons. On the first, which we may call internalist, the rationality of one's acts depends only on one's beliefs: one acts rationally if the truth of what one believes would give one sufficient reason to act in that way. (Parfit, 2011, 111) The natural dividing line here is between true and false belief. One's acts are rational if they adequately fit one's beliefs. If those beliefs are true in addition, one acts as one has reason to; if they are false, one typically does not.

On the second picture, which we will call externalist, we distinguish full rationality and internal rationality. Full rationality depends on knowledge, while internal rationality depends on belief. An act is fully rational if and only if the facts one knows give one sufficient reason to act in that way. If one acts without knowing facts that give one sufficient reason to act in that way, one's act is not fully rational. What if one had every reason to believe something whose truth would give one a sufficient reason, but that belief was false? A doctor uses a needle which is adequately marked as sterile but has been contaminated due to some rare blunder of which she is unaware. She does not know that the needle is safe, and still, is it not rational for her to use it? Not fully rational, on the present view, but excusable. The doctor's act is wrong because she does not know that the needle is safe, but excusable because she could not know that she did not know it. (Hawthorne and Stanley, 2008, 573) The natural dividing line here is between knowledge and mere belief. An act is internally rational if it fits one's beliefs, but internally rational acts divide into fully rational ones based on known facts and excusable ones based on mere beliefs.

Three points bring out the contrast between the two pictures. First, there is a use of "ought" distinct from the objective one on which it is not the case that you ought to have helped me if you neither knew nor believed that I needed help. Call this the rational "ought". The two pictures disagree on what one ought to do in that sense. On the internalist picture, someone who (mistakenly) believes that smoking protects their health ought to smoke. (Parfit, 2011, 115) This is not so on the externalist picture. Second, the difference between knowledge and true belief is superficial on the internalist picture but not on the externalist one. On the former, whether one knows or simply has a true belief makes no difference. Provided that one is rational, one is acting as one really has reason to in both cases. On the externalist picture, there is a crucial difference. Only knowledge allows one to act for the reasons one in fact has. (Hyman, 1999) In the true belief case, one at most acts as one has reasons to, but not for reasons one has, that is, one's act matches the reasons one has but is not justified nor explained by them. Third, the externalist sees differences in justification among

internally rational acts that the internalist does not. On the externalist view, an internally rational act based on knowledge is fully justified, while an internally rational act based on mere belief is at most fully excused. On the internalist picture, they are equally justified.

3.2. The Higher Expected Value of Knowledge

The answer to Meno's Paradox I put forward endorses the externalist picture. The picture entails that one ought to prefer "believing when one knows" to "believing when one does not know", even if both beliefs are equally valuable. To show that, we need two distinctions.

First, we need a distinction between value proper and expected value. Consider the following case, due to Piller (2009a, 123):

You are given a choice between two bundles of lottery tickets, at no cost. The winning ticket is in one of the two bundles, but you have to way to know or guess which. The first bundle contains 100 tickets, the other 10. As it happens, the winning ticket is in the smaller one.

Does the bigger bundle have any value? No, I submit. It will not get you a prize and a bundle of tickets is useless to you. Yet you ought rationally to take it. Why? Because it is the one you can most expect to have value. From your point of view, it is the one that is most likely to win the prize, that is, to have value. In short, we say that it has the highest expected value. But we should be clear that expected value is not value, just as an expected result is not a result.²²

Expected value is perspective-dependent in a way that value is not. Initially you should choose the bigger bundle; but if you later learn that it is a loser, you should not. That is not because it has some value that suddenly vanishes when you learn that, but because you expected it to have value but learned that it does not.

Second, we need a distinction between really expected value and merely apparently expected value. When facts one knows give one reasons to expect something, I say that one really expects that thing. When one has beliefs the truth of which would give one reason to expect something, I say that one apparently expects that thing. If those beliefs do not constitute knowledge, one merely apparently expects that thing. For instance, if a man is convinced that it will rain tomorrow by reading tea leaves, he merely apparently expects the rain, while a woman who is convinced by a good weather report really expects the rain. My use of "expect" is non-psychological. The man and the woman both have an attitude of expecting the rain. But in my use, the rain is only really expected by the woman. Expectations in my sense do not even require subjects to have an attitude of expecting.²³

In a probabilistic framework, we can define real expectations as probabilities given what one knows and apparent expectations as probabilities given what one believes. Merely apparent expectations arise when there is

a gap between the two. Known facts are a special case of real expectations: assuming that knowledge entails probability one, they are real expectations with certainty.

We now get the following result:

Knowledge has a Higher Really Expected Value For any p, as far as the value of believing the truth about p is concerned, the really expected value of p to the knower (at the time of knowing) is always higher that it is to the mere believer (at the time of believing).

Let us say that the value of believing the truth about p is one on some relevant scale. The value of a particular true belief that p may vary widely depending on its implications. But as far as the value of believing the truth about p is concerned, its ceteris paribus value is one on the relevant scale. Now suppose that I currently know p. What value should I expect believing p to have? The question may appear a bit unnatural, since believing p is not an act I can choose, and since I do in fact already believe p. But it is well-defined nonetheless. Given what I know, the probability that I believe the truth about p by believing p is one. So as far as believing the truth about p is concerned, the really expected value (to me, now) of believing p is one. Now suppose that I do not currently know p, though I believe it. What value should I expect believing p to have? Given what I believe, the probability that I believe the truth by believing p is one. So the apparently expected value (to me, now) of believing p is one. But barring a few exceptional cases, the probability given what I know that I believe the truth about p by believing p is less than one. So the really expected value (to me, now) of believing p is less than one. As far as the value of believing p is concerned, the really expected value of believing p to the knower at the time of knowing is higher than its really expected value to the mere believer at the time of believing.²⁴

Now on the externalist view of rationality, one ought to act and to be according to facts known to one. In particular, one ought to prefer according to one's really expected values, that is, according to what can be expected to be best given what one knows. But as we have shown, believing p has a higher really expected *ceteris paribus* value to the knower than to the believer. Hence necessarily, one who knows p ought to prefer believing p and is fully rational in believing p, *ceteris paribus*. By contrast, to one who merely believes p, believing p is not decisively preferable.²⁵

We thus have an explanation of why one ought to prefer believing when one knows over believing when one does not. To the knower, one's current belief that p is the one she maximally ought to prefer having, *ceteris paribus*. To the mere believer, one's current belief that p is not maximally the one she ought to prefer having, *ceteris paribus*. The *ceteris paribus* clauses indicate that these claims are restricted to what pertains to the value of believing the truth about p. Note that the explanation essentially relies on the externalist picture. On the internalist view, both would believe as they rationally ought to.

The explanation puts pressure on defenders of Meno's Thesis. It only assumes that believing the truth has value and it is compatible with the denial that knowledge has any more value than true belief. It nevertheless explains why it is *ceteris paribus* rational to prefer knowledge over mere true belief, in the sense that one ought to prefer believing when one knows.²⁶ But I doubt there is anything more to the intuition behind Meno's Thesis than the idea that knowledge is *ceteris paribus* universally preferable to mere true belief. Hence the explanation puts a burden of proof on those who claim that knowledge also genuinely better than mere true belief.

3.3. The Perspective Worry and the Circularity Worry

Let me discuss two worries for the "Higher Expected Value" solution. Throughout the discussion, we omit the qualifications that a belief's "value" and "preferability" refer to its value and preferability insofar as believing the truth about the proposition believed is concerned, and that "expectation" refers to one's real expectations.

First, the "Perspective Worry". We have shown that knowledge is preferable to mere true belief to the knower at the time of knowing. But knowing appears preferable even when viewed from the past or the future. Can we explain that? Several replies can be made. First, the expected value of knowledge is plausibly generally higher. The expectation that one will believe the truth is higher, if one acquires knowledge than if one reads tea leaves. But that holds in general only; in particular, it does not apply to past mere true belief which is presently known to be true. Second, one may argue that being fully rational has value in itself. Since knowing implies a *ceteris paribus* fully rational belief while mere true believing does not, knowing has higher expected value from all perspectives. That is in effect Piller's "Deontic Value" solution. As we have seen, its consistency is dubious. Third, one may appeal to Broome's (1999) idea that in addition to aiming at what has most value, rationality requires various forms of coherence such as intending the means to one's ends or believing that one will do what one intends to do. Those requirements stand on their own and do not derive from or imply a value of being coherent. We may think that they include the requirement to prefer one's future and past self to be rational. Insofar as having a mere true belief is being in a state that does not maximally fit one's real expectations, it is a less than fully rational state. Rationality would thus require preferring not to be in such a state at any time. Obviously, the idea needs to be worked out in more detail. But the line seems to me promising.

Second, the "Circularity" worry. The externalist picture of rationality I am relying on puts a heavy normative load on the difference between true belief and knowledge. Is it not circular to assume it while trying to explain why knowledge is preferable to true belief? It would be so if the fact that one ought to prefer according to what one knows—rather than according to what one truly believes—was itself explained by the fact that knowing is preferable to having a mere true belief. But it is not. In fact, even if we assumed that

knowledge was not only preferable to but better than having a mere true belief, that would hardly explain why we ought to act in the light of our knowledge rather than our beliefs or true beliefs. The normative role of knowledge is either primitive or explained otherwise. Hence the solution is not circular.

That being said, the solution is conditional on the view that knowledge does play a special normative role that mere true belief does not play. I am fairly optimistic about its prospects, but it is beyond the scope of the present paper to defend it.²⁷

4. Conclusion

When one knows, one ought to believe what one believes. When one has a mere true belief, one ought not to believe what one believes. Or, more cautiously: it is not the case that one maximally ought to believe what one believes. There are several routes to those claims. The one we followed did not assume that knowledge was the aim of belief or that it had more value than true belief. Rather, it simply assumed that believing the truth was a good thing and it endorsed a view of rationality on which one ought to act and to be according to what can be expected on the basis of what one knows. Thus we can explain why knowledge ought to be (universally, *ceteris paribus*) preferred to true belief without claiming that it is better than true belief. That is the "Higher Expected Value Solution" to Meno's Paradox.

The solution leaves several issues open. First, it assumes an externalist picture of rationality on which knowledge plays a central normative role that true belief does not play. This view needs further exploration and defence. Second, it remains to be seen whether the solution can explain why past and future knowledge ought to be preferred over mere true belief. We have outlined an option that appeals to requirements of rationality that are neither based on nor implying values, but such requirements are controversial.

However, the proposal is substantial enough to put pressure on defenders of Meno's Thesis. In particular, it puts pressure on solutions to Meno's Paradox that introduce novel kinds of derived value such as type instrumental value, credit value or deontic value. In the light of our solution, such notions are suspect of confusing expected value and value proper.²⁸

Notes

- 1 The label *Meno's Paradox* is commonly used for an unrelated puzzle also stated in the Meno (80d).
- 2 See Sosa (2001), Kvanvig (2008) and Piller (2009b) for further discussion.
- 3 See Jones (1997, 434).
- 4 See Leslie (2008) for further discussion.
- 5 See Kvanvig (2008, 201–2). Kvanvig's terminology of "prima facie value" and his analogy with defeasible reasoning tends to confuse ceteris paribus value proper with prima facie appearance of value. We clarify the distinction below.

- 6 For further discussion of the value of true belief and a defence of the idea that it is a fundamental good, see Kvanvig (2003, Chap. 2).
- 7 See Williamson (2000, 52).
- 8 See Piller (2009a, 125) for further discussion. On his view the very fact that instrumental "value" raises double counting issues (see below) indicates that it is not value at all.
- 9 Not all notions of consequentialism in the literature fit that definition.
- 10 See Goldman and Olsson (2009, 24) on the idea that the swamping problem assumes verifism
- 11 On the stricter, actualist notion of instrumental value sketched in the previous section, a ticket that does not win has not and *did not* have instrumental value at any point. It merely had a chance of having instrumental value. Goldman and Olsson's solution relies on a more liberal version of consequentialism.
- 12 Piller (2009a, 128–9) makes related points. But I would not endorse his claim that one is equally likely to form further true beliefs given Goldman and Olsson's assumptions alone than given these assumptions and the fact that one has formed a true belief reliably.
- 13 See Williamson (2000, 2.4 and 3.4-3-7) for a useful discussion.
- 14 The same holds for consequentialist views relying on objectively probable causal consequences.
- 15 Goldman and Olsson's (2009, 34) defence of the psychological shift is unclear to me. They first claim that our evidence on the value of knowledge wholly consists in intuitions about its value an objectionable claim, but let it pass here. They then say that (a) we are not methodologically entitled to assume that our intuitions track the facts, though they do not want to deny it either, and (b) if these intuitions are misguided, "the epistemological enterprise needn't collapse". What they mean by "epistemological enterprise" is presumably what the following sentences refer to as "explaining epistemic intuitions" and "understanding the sources of epistemic intuitions". As they rightly point out, that project "may profit from understanding human psychology".

These comments seem to amount to the following: even if Meno's Thesis turns out to be false, it is worth explaining why we find it appealing. In short, error theories are available. Indeed, it is natural to construe their solution as an error theory explaining how we end up valuing intrinsically some valueless means—as Piller (2009a, 132) points out, Mill gives just such a story about money. (*Utilitarianism*, chap. 4, 36) Yet Goldman and Olsson ostensibly introduce these points in defence of the idea that human psychology can shed light on a question that "concerns objective value, not how people *assign* valuational status", namely "whether a state of affairs with the status of being merely instrumentally good can objectively change its status to being non-instrumentally, or fundamentally, good". (Ibid.) I fail to see how their points support that unless they mean that human psychology "sheds light" on the question negatively, by explaining away judgements of objective value. Be that as it may, they then describe value autonomisation in ways that can—but need not—be construed non-psychologically: "some states of affairs that at one

time are assigned merely instrumental value are 'promoted' to the status of independent, or fundamental, value" (34), "while many such processes are originally regarded as merely instrumentally valuable to true-belief attainment, they are later upgraded to the status of independent value." (35; my italics) "Promoted" and "upgraded" can equally denote an objective change in value or a mere change in our judgements of value.

In the end, Goldman and Olsson appear aware that the autonomisation story can be construed as either an error or vindication theory but they seem to try to avoid commitment either way. That leaves unclear whether they ultimately endorse Meno's Thesis. Goldman (2011, §4) gives it a clear vindication reading, however.

- 16 See Parfit (2011, Chap. 3) for a characterisation and criticism of subjectivism about value. As it is formulated, Goldman and Olsson's solution would only lead to the claim that reliably formed true belief is better than true belief. If reliably formed true belief is not sufficient for knowledge, this falls short from showing that knowledge itself is better. But the story is easy to adapt to reach the latter conclusion, by starting from the remark that knowing is a good means to believe the truth.
- 17 See Lackey (2004) and Greco (2010, 80-4) for a reply.
- 18 Piller does not say whether likelihood given one's point of view is likelihood given what one knows or what one believes. He uses the indeterminate term "awareness" (2009a, 134).
- 19 A related issue is that the Deontic Value view warrants a regress in what one ought to do. Start from some good. One ought to choose the likely good. But then choosing the likely good is itself good. So one ought to choose the likely (good and choosing the likely good). But that itself is now good, and so on.
- 20 The solution agrees with Piller's (2009a) on several points, notably on the idea that the fact that knowledge is preferable to true belief is not grounded in knowledge's being better than true belief. (Piller, 2009a, 123) But they disagree on others: Piller's solution is internalist (2009a, 134) and liberal (2009a, 126), while the "Higher Expected Value Solution" is externalist and austere.
- 21 I label the two pictures "internalist" and "externalist", respectively. They have some affinity with the internalist/externalist divide set up by Williamson (2000, esp. §2.1, §8.7, §9.1, §9.3). But it is beyond the scope of the present paper to discuss whether they reflect the various "internalist/externalist" divides of the literature. The reader who finds my labels misleading is urged to ignore them.
- 22 In decision theory, representation theorems show that an agent meeting certain conditions of rationality prefers acts that have the highest expected utility, where the latter is computed on the basis of her preferences over possible outcomes and her expectations of each act to have these outcomes. The result is often summarized in the claim that the utility of an act is its expected utility. If utility was value, that would contradict the idea that expected value is not value. But as the spelled out version makes clear, the theorem is about rational preferences, not value. Utility is thus better understood as rational preferability rather than value. When values are known, things are rationally preferred according to their values. When values are not known, things are rationally preferred according to their expected value.

The notion of utility ignores the distinction. For instance, a rational agent may be indifferent between a ticket that returns \$1 for sure and a ticket that she judges to be 1/10 likely to return \$10 and nothing otherwise. In decision-theoretic terms, the utility of the two tickets is the same for her. But that does not mean that their *value* is the same. In fact, she knows that their value is different: the second is either worthless or ten times better than the first. Only their expected value is the same. See Piller (2009b, 423–5) for further discussion.

- 23 A less ambiguous term may have been preferable, but I know of none.
- 24 The exceptions are cases in which the probability of believing the truth by believing p given what one knows is one, even though one does not know p itself. First, there is the case of a mere believer who knows things that entail p without knowing p itself. She fails to draw the consequence from what she knows but believes p on other grounds. Second, in some cases involving infinite numbers of outcomes, the probability of p given what one knows can be one even though what one knows fails to entail p. For instance, it is often assumed that the probability that one will not pick the (geometrical) mid-point of a line given that one picks a point (genuinely) randomly on it is one. In my view the cases reflect a limit of the probabilistic framework rather than a limit in the present argument. When p is a consequence of what one knows that one has failed to draw, one does not really expect p in the way that one does if one knows p. Similarly, if probability one is not enough for knowledge, then one does not really expect p when p is merely probability one given what one knows in the way that one does if one knows p. Be that as it may, we can ignore these exceptional cases in the present discussion.
- 25 How far one still ought to believe *p* and whether one is still fully rational in believing *p* in cases in which one merely believes *p* depends on the disvalue of believing the false about *p* of the value of suspending judgement. If suspending judgement has no value and the disvalue of believing the false is finite, there will be cases in which one known facts give one most reason to believe *p* as far as believing the truth about *p* is concerned. But there is still a contrast between that case and the case of the knower. In the case of the knower, believing *p* is absolutely preferable insofar as believing the truth about *p* is concerned, while in the case of the mere believer, it is only relatively preferable.
- 26 If knowing has the same value as believing truly, then the expected value of knowledge cannot be higher than that of true belief—on the plausible assumption that one always expects believing at least as much as knowing, since the latter entails the former. In that sense our result is not strictly speaking that one ought to prefer knowing over believing truly. It is that one ought to prefer one's belief when one knows to one's belief when one merely believes. Belief that constitutes knowledge ought to be preferred to belief that does not, not because knowing itself has more expected value, but because believing has more expected value when one knows.
- 27 See Hawthorne and Stanley (2008) for a defence of that role with respect to action.
- 28 Thanks to Anne Meylan, Pascal Engel, Fabian Dorsch, Stephane Lemaire, and audiences at Amsterdam, Fribourg, the Logos research group (Barcelona) and Rennes for useful discussions of the ideas of this paper.

References

Broome, J.

1999 "Normative Requirements." Ratio 12: 398-419.

2004 "Reasons." In Reason and Value: Essays on the Moral Philosophy of Joseph Raz. Wallace, R. J., Smith, M., Scheffler, S., and Pettit, P., eds. Oxford University Press.

Goldman, A.

2011 "Reliabilism." In The Stanford Encyclopedia of Philosophy (Spring 2011 Edition). Zalta, E. N. ed. http://plato.stanford.edu/archives/spr2011/ entries/reliabilism/.

Goldman, A. and Olsson, E.

2009 "Reliabilism and the Value of Knowledge." In Epistemic Value. Pritchard, D., Millar, A., and Haddock, A., eds. Oxford University Press.

Goldman, A. I.

1986 Epistemology and Cognition. Harvard University Press.

Greco, J.

"Knowledge as Credit For True Belief." In Intellectual Virtue: Prespec-2003 tives from Ethics and Epistemology. Depaul, M., ed. Oxford University Press.

2010 Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity. Cambridge University Press.

Hawthorne, J. and Stanley, J.

2008 "Knowledge and Action." Journal of Philosophy 105(10): 571–590.

Hyman, J.

1999 "How Knowledge Works." Philosophical Quarterly 49: 433–451.

Jones, W.

1997 "Why Do We Value Knowledge?" American Philosophical Quarterly 34(4): 423-439.

Kvanvig, J. L.

2003 The Value of Knowledge and the Pursuit of Understanding. Cambridge University Press.

2008 "Pointless Truths." Midwest Studies in Philosophy 22: 199–212.

Lackey, J.

2004 "Review of Michael DePaul and Linda Zagzebski, eds. Intellectual Virtue: Perspectives from Ethics and Epistemology." Notre Dame Philosophical Reviews Available at: http://ndpr.nd.edu/review.cfm?id=1462.

Leslie, S.-J.

2008 "Generics: Cognition and Acquisition." *Philosophical Review* 117(1):

Olsson, E.

2007 "Reliabilism, Stability, and the Value of Knowledge." American Philosophical Quarterly 44: 343-55.

Parfit. D.

2011 On What Matters, Vol. 1. Oxford University Press.

Piller, C.

2009a "Reliabilist Responses to the Value of Knowledge Problem." *Grazer Philosophische Studien* 79: 121–35.

2009b "Valuing Knowledge: a Deontological Approach." *Ethical Theory* and Moral Practice 12: 413–28

Riggs, W. D.

2002 "Reliability and the Value of Knowledge." *Philosophy and Phenomenological Research.* 64(1): 79–96.

Sosa, E.

2001 "For the Love of Truth?" In *Epistemology: Essays on Epistemic Virtue* and *Responsibility*. Fairweather, A. and Zagzebski, L., eds. *Virtue* Oxford University Press.

Swinburne, R.

1999 Providence and the Problem of Evil. Oxford University Press.

Williamson, T.

2000 Knowledge and its Limits. Oxford University Press.

Zagzebski, L.

2003 "The Search for the Source of Epistemic Good." *Metaphilosophy* 34(1–2): 12–28.