To appear in a symposium on *Knowledge and its Limits* in *Philosophy and Phenomenological Research*

Replies to Commentators*

Timothy Williamson

Reply to Brueckner

The core of Tony Brueckner's critique in 'Knowledge, Evidence, and Skepticism according to Williamson' is his claim in section 5 that my account of perceptual knowledge has an unacceptable consequence. My reply will concentrate on that claim and largely ignore the rest of Brueckner's interesting discussion, for it is easy to check that the claim is essential to Brueckner's argument against my analysis of skepticism and evidence.

The alleged consequence at issue concerns a case in which Brueckner knows by seeing that his cup is red. According to Brueckner, I am committed to the implausible view that '[his] belief of the proposition that [his] cup is red is justified in virtue of [his] belief of the proposition that [his] cup is red!' Why does he take my account of perceptual knowledge to imply any such thing? He does not clearly explain, but his account offers some clues.

I defend the equation E=K, that one's total evidence is one's total knowledge. By that equation, since Brueckner knows that his cup is red, the proposition that his cup is red is part of his total evidence. Of course, if it is a normal case of vision, Brueckner will have far more visual knowledge than that: for example, of the cup's specific shade of red, and of its shape and location. All that other knowledge is also part of his total evidence. For simplicity, however, let us ignore all that other knowledge and treat the proposition that his cup is red as though it exhausted the relevant part of the total evidence. If anything, that simplification should help Brueckner. I also hold that beliefs are justified, when they are, by the subject's total evidence, that is, by the subject's total knowledge. When one's total evidence is *e*, the justified level of confidence (credence) in a hypothesis *h* is measured by the conditional probability of *h* on *e*. If one is less than fully rational, one may unjustifiably believe a proposition for subjective reasons unrelated or insufficiently related to the high degree to which one's evidence supports it; what one's total evidence determines is the less subjective matter of the level of confidence in *h* that is justified for one. In the present simple case, the proposition that his cup is red is part of Brueckner's total evidence, and therefore has conditional probability 1 on his evidence. His knowledge that his cup is red justifies his belief that his cup is red. If we want to use the somewhat obscure expression 'in virtue of' (Brueckner's, not mine), we can say: his belief that his cup is red is justified in virtue of his knowledge that his cup is red. Perhaps Brueckner's thought is that since knowledge entails belief, his knowledge that his cup is red *is* his belief that his cup is red (which constitutes knowledge in this case).

Consequently, by substitution of coreferring expressions in the position after 'in virtue of', his belief that his cup is red is justified in virtue of his belief that his cup is red. Thus, by the assumed equivalence of 'X is justified in virtue of Y' and 'Y justifies X', his belief that his cup is red justifies his belief that his cup is red. But surely such a belief is not self-justifying.

That argument is spurious. Presumably, the number 9 is composite in virtue of its divisibility by 3. If Brueckner's knowledge that his cup is red is his belief that his cup is red, then in a corresponding sense 9's divisibility by 3 is its compositeness. An inference like the one above would yield the implausible-sounding conclusion that 9 is composite in virtue of its compositeness. Such reasoning hardly refutes the harmless claim that 9 is composite in virtue of its divisibility by 3. An exact diagnosis would require a proper semantics of the 'in virtue of' construction, which would take us too far out of our way. If Brueckner has a point, he should be able to make it in other and clearer terms.

If I claim that Brueckner's belief that his cup is red is justified in virtue of his knowledge that his cup is red, and that his knowledge justifies his belief, I am claiming just this: the belief that his cup is red is justified for him *because* he knows that his cup is red. The 'because' here is explanatory rather than causal (9 is composite because it is divisible by 3). I grant that, necessarily, if he knows that his cup is red then he believes that his cup is red. Given the inference schema 'P because Q; necessarily, if Q then R; therefore P because R', we could conclude that the belief that his cup is red is justified for him because he believes that his cup is red, which is surely false. But the inference schema is clearly unsound. Suppose that I have a right to vote in the election because I am a citizen at least eighteen years old. Necessarily, if I am a citizen at least eighteen years old then I am at least five years old. It does not follow that I have a right to vote in the election because I am at least five years old.

Perhaps Brueckner had a very different argument in mind. If so, I have no idea what it was. Clearly, he suspects that my account of evidence as knowledge does not do justice to the evidential relevance of perceptual experience, even though I allow that perceptual experience may be a necessary condition of perceptual knowledge. However, he has completely failed to substantiate his suspicion.[1]

Reply to Conee

A condition is luminous if and only if whenever it obtains the agent is in a position to know that it obtains. In *Knowledge and its Limits* I argue that only trivial conditions, such as necessary or impossible ones, are luminous. In 'The Comforts of Home', Earl Conee finds my defence of a crucial premise of the argument unconvincing. Instead, he presents a different argument of his own against luminosity, but claims that the denial of luminosity lacks the epistemological significance that I claim for it.

1. Conee states the crucial premise of my anti-luminosity argument thus (for the exemplary case in which the agent is Smith and the condition is that one feels cold):

(1) If at some time during the sequence Smith knows herself to feel cold, then one millisecond later Smith feels cold.

As emphasized in the book (p. 97; all page references are to *Knowledge and its Limits*), I claim no sort of lawlike status for (1). Rather, (1) is a description of a specific hypothetical process, during which the agent's state changes very gradually from one of feeling cold to one of feeling hot. The conditional in (1) is to be read as merely truth-functional. Thus Conee is mistaken when he says that (1) 'is not plausible on its face, because it implies that knowledge that is entirely about the present depends on the future' and when he treats (1) as vulnerable to the objection that 'sudden death can intrude'. By hypothesis, Smith does not die during the process (although she could have died in that time). Nor does (1) imply any dependence of present knowledge on future states. In particular, it does not imply the counterfactual that if Smith had not felt cold one millisecond later, one millisecond before that she would not have known herself to feel cold. To make the point vivid in terms of possible worlds: perhaps in the nearest worlds in which Smith does not feel cold one millisecond later, a sudden change has intervened (unlike what happens in the hypothetically actual world in which the process occurs as originally described) and one millisecond before she did know herself to feel cold. No strange dependence of present on future is at issue.

Conee's misunderstanding of the status that I attribute to (1) is not the basis of his main objection to my argument for (1). That argument uses a reliability constraint on knowledge. Conee responds by sketching an epistemology of feeling cold that is supposed to satisfy the reliability constraint while falsifying (1). The idea is that, in the sense of ‘feeling cold’ in which the condition is an optimal candidate for luminosity, at the earlier time Smith knows that she feels cold by feeling a specific thermal phenomenal quality Q, where feeling Q entails feeling cold. A millisecond later she may be feeling a very slightly different specific thermal phenomenal quality Q*; even if feeling Q* does not entail feeling cold, that is supposed to show nothing about the reliability of feeling Q as a basis for the belief that one feels cold.

For the sake of argument, we may grant that Smith feels Q at the earlier time (*t*) and Q* a millisecond later, and that necessarily whoever feels Q feels cold. Conee calls feeling Q ‘an entailing basis’ for the belief at *t* that one feels cold and comments ‘Nothing could be a more reliable basis than that’. One might worry that the facts that she feels Q and that feeling Q strictly implies feeling cold would not enable Smith to know that she felt cold if she lacked cognitive access to those facts. However, my argument for (1) is an argument from reliability, not from cognitive access to reliability, so the question is whether Conee has provided a reliable basis for Smith’s belief at *t* that she feels cold.

Consider an analogy. Suppose that Jones has a clear and distinct experience E as of seeing 29 stars. On the basis of having E, Jones forms the belief that he is having an experience as of seeing 29 stars. In Conee’s sense, having E is an entailing basis for that true belief and is maximally reliable. However, for any natural number *n* between 20 and 40, when Jones has an experience as of seeing *n* stars, he usually forms a belief that he is having an experience as of seeing 29 stars. In most cases, that belief is false. The underlying psychological mechanism is the same for all those values of *n*. He makes no attempt to count but simply estimates the number from his general impression; forgotten events in his childhood caused a strong bias in favour of the number 29. When Jones, on the basis of having E, forms the belief that he is having an experience as of seeing 29 stars, does he know that he is having an experience as of seeing 29 stars? Surely not. Moreover, surely a natural justification for that answer is that his belief was formed in too unreliable a way. We can imagine Jones as reaching his conclusion by an implicit inference from (2) and (3) to (4):

(2) I am having an experience with this specific phenomenal quality.

(3) Any experience with this specific phenomenal quality is an experience as of seeing 29 stars.

Therefore:

(4) I am having an experience as of seeing 29 stars.

Assume that the demonstrative ‘this specific phenomenal quality’ in (2) and (3) works in a way that makes (2) epistemologically unproblematic (because it is guaranteed to pick out a quality of the experience that the speaker is having) and (3) a necessary truth (because it rigidly designates a quality that only experiences sufficiently similar to E possess). Nevertheless, Jones does not know (3), because his belief in (3) is formed in too unreliable a way.

Contrast the conclusion of the previous paragraph with what Conee says about the original case (footnote 12):

To have knowledge, it may be that the basis for the classifying belief must include some understanding of the entailment of the classification by the more specific feeling. Since the content of this understanding would be a necessary truth, this could not weaken the reliability of the basis.

Conee assumes that in assessing the reliability of a belief one should hold its content fixed. If *p* is a necessary truth, then it is impossible to have a false belief in *p*, so Conee counts any belief in *p* as maximally reliable. That conception of reliability is too narrow. For example, an irrational amateur mathematician may form mathematical beliefs in highly unreliable ways, even though about half of the believed propositions are truths, and therefore necessary truths (pp. 181-2). Such beliefs are unreliable in a sense relevant to whether they constitute knowledge. The importance of beliefs with other contents reached by a similar process for the assessment of reliability was emphasized in the original anti-luminosity argument (p. 101). Beliefs with slightly different bases also bear on reliability. In the case of Jones, his false belief on the basis of having an experience E*, which is in fact as of seeing 28 stars, that he is having an experience as of seeing 29 stars detracts from the reliability of his true belief on the basis of having E that he is having an experience as of seeing 29 stars. Similarly, in the case of Smith, a false belief a millisecond after *t* on the basis of feeling Q* that she is feeling cold would detract from the reliability of her true belief at *t* on the basis of feeling Q that she is feeling cold. Thus Conee's attempt to dismiss the reliability argument for (1) fails (see particularly his footnote 13).

Conee says nothing about the basis he postulates for Smith's belief on the basis of feeling Q that she is feeling cold that would make it any better off than Jones's belief on the basis of having E that he is having an experience as of seeing 29 stars. That spells the failure of his attempt to explain how counterexamples to (1) could occur during the hypothesized process.

Conee adds an unnecessary complication by arguing that ‘it is best to suppose that Smith’s grasp of the concept [of feeling cold] is not perfect’, on these grounds (footnote 10):

> If Smith has such a keen grasp of the concept, and feeling cold is luminous, then her confidence would drop off sharply when her thermal feeling is no longer in the cold range.

If this were granted, it would already suffice to refute the luminosity of the condition that one feels cold. For it is clear that people who fully grasp the concept of feeling cold can undergo exactly the kind of gradual diminution described of confidence that they feel cold. But in fact the conditional claim just displayed is by no means trivial. The luminosity of the condition that one feels cold does not automatically entail the luminosity of the condition that one does not feel cold. We have no compelling reason to accept the verificationist assumption that fully grasping a concept involves the capacity always to recognize whether it applies. Indeed, to establish Conee’s quoted claim, one would need to provide considerations similar to those in the anti-luminosity argument itself.

2. I turn from Conee’s critique of my anti-luminosity argument to his own anti-luminosity argument. The idea is that, in suitable circumstances, the grounds for Smith’s true belief

that she feels cold could be defeated by apparently expert testimony that she does not really feel cold but, for example, is merely suffering from a recently identified rare psychiatric disorder one symptom of which is an illusion of feeling cold; even if she maintained her belief that feels cold, it would be insufficiently justified to constitute knowledge. That the testimony is in fact utterly bogus, perhaps even incoherent, is consistent with its capacity to inflict epistemic defeat. Such considerations generalize easily and extensively.

I have considerable sympathy for Conee's anti-luminosity argument. It is related to, although not identical with, a fallibilist observation in the book: 'There is no limit to the conclusions into which we can be lured by fallacious reasoning and wishful thinking, charismatic gurus and cheap paperbacks' (p. 94). However, I envisaged the friend of luminosity replying that 'such mistakes are always rectifiable'. I have a similar concern about Conee's argument. The friend of luminosity might insist that, despite the defeaters, the agent is still in a position to know, and perhaps would know if she ignored the defeaters. I do not endorse such a reply. It requires a sense of 'in a position to know' too watered down to be of much use. However, I wanted an anti-luminosity argument that would be even more robust to the sense of 'in a position to know' (a notion needed more by the friends of luminosity, to ward off cheap counterexamples, than by its enemies). I also wanted an anti-luminosity argument that would work even on a less holistic conception of knowledge, and reveal limits to knowledge that are more intrinsic to the allegedly luminous conditions and our simplest cognitive relations to them. That is why I gave the kind of anti-luminosity argument that I did. But I do not regard the kind of anti-luminosity that Conee gives as negligible or insignificant. As he points out, it can be used to reach some conclusions for which my anti-luminosity is inapplicable, in particular concerning obstacles to knowledge that a condition obtains in cases far from those in which it does not obtain.

Conee also suggests that some conditions that can be argued in his way to be non-luminous cannot be argued in my way to be non-luminous because they lack marginal cases. His candidate is the condition C', that one feels pain and consciously attributes to oneself being in severe pain, on the grounds that 'No questionable case of feeling pain would be consciously regarded as severe, and there do not appear to be marginal cases of self attributing severe pain' (footnote 19). Both conjuncts seem wrong. Against the first, a self-pitying, self-deceiving person who is a questionable case of feeling pain may consciously attribute to himself the state of being in severe pain. Against the second conjunct, self-attribution in the intended sense involves judging that one is in severe pain, not merely toying with the idea that one is in severe pain, and there are intermediate cases too; the degree of one's commitment may make one a marginal case of judging that one is in severe pain.

3. Conee argues that his anti-luminosity argument undercuts my equation E=K of one's total evidence with the total content of one's knowledge:

> While Smith appreciates a defeater, what is defeated is her experiential evidence to attribute feeling cold to herself. So she has that experiential evidence. It continues to be epistemic reason for her to think that she does feel cold. Thus, it is evidence for her, though it is not a fact that she knows.

Conee is right that Smith continues to have some evidence in favour of the proposition that she feels cold. Since she does not know that she feels cold (let us assume), her evidence does not include the proposition that she feels cold. But Conee ignores a host of eligible candidates for inclusion in Smith's evidence that are compatible with the equation E=K. Indeed, Conee's own earlier discussion supplies just such a candidate. For he imagines a pseudo-expert who distinguishes between how experience appears to the subject and how experience really is. Unlike me, Conee himself does not claim that it is metaphysically possible for experience to appear to the subject other than it really is. His point is that the pseudo-expert can exploit the merely epistemic possibility of a difference. Given this epistemic possibility for Smith, even if she is not in a position to know that she feels cold, she may still know that she appears to herself to feel cold. If so, her evidence includes the proposition that she appears to herself to feel cold. That evidence raises the probability that she feels cold. Thus she still has evidence in favour of the proposition that she feels cold.

Other pieces of propositional evidence are also available to Smith. In feeling cold, she has a particular experience ε, which she can describe in various specific ways that do not presuppose that she feels cold. She may still know that she has ε. If so, her evidence includes the proposition that she has ε. Again, she may know that she is strongly inclined to judge that she feels cold. If so, her evidence includes the proposition that she is strongly inclined to judge that she feels cold. Evidence of such kinds also raises the probability that she feels cold. Thus she can still have evidence of such other kinds in favour of the proposition that she feels cold. Of course, she might happen not to formulate those propositions consciously. But she is still in a position to know them, and thereby include them in her evidence.

In response, Conee might try to generalize his thought experiment, by having the pseudo-expert undermine Smith's knowledge that she appears to herself to feel cold, that she has experience ε, that she is strongly inclined to judge that she feels cold, and so on. That is not altogether easy to imagine. If the pseudo-expert says things of the form 'You appear to yourself to have property P, but really you don't' or 'You think that you have P, but really you haven't', he thereby allows Smith to know that she appears to herself to have P or to know that she thinks that she has P, and consequently to include such relevant propositions in her evidence. If the pseudo-expert must not say anything of that form, his task of undermining *all* Smith's relevant knowledge is much harder. By blinding her with science, he is supposed somehow to reduce her to a state of such total intellectual disarray that she is no longer in a position to know anything relevant at all about how she feels or what her experience is like. To claim that in such an intellectually devastated state she would still have her experience as evidence that she feels cold is grossly to underestimate the cognitive damage that would have been inflicted on her.

In his conclusion, Conee asserts that phenomenal qualities provide a comfortable cognitive home, even though they are non-luminous, because they are always there 'among our ultimate evidential resources'. But what use to you as evidence is a phenomenal quality when you are not even in a position to know that you have it? Conee reassures us that phenomenal qualities are still 'known to us by acquaintance', but in Conee's sense you can know *x* by acquaintance even if you are not in a position to know that *x* exists or appears to exist or is present to you or appears to be present to you or ….

When knowledge by acquaintance is so isolated from propositional knowledge, merely saying that phenomenal qualities are ‘known to us by acquaintance’ no longer helps to explain with what right they might play any more privileged a role in the theory of evidence than do perfectly unconscious qualities. Conee relies uncritically on some assumptions from a foundationalist epistemology such as one finds in Bertrand Russell’s *The Problems of Philosophy* (1912) while rejecting others: but classical foundationalism was an integrated system, from which one cannot simply tear fragments and take for granted that they will continue to have the significance which they had when they functioned within the original system.

Reply to Hawthorne

In ‘Knowledge and Evidence’, John Hawthorne presses some of the most important challenges to the epistemology of *Knowledge and its Limits*. My sections here are numbered in correspondence with the sections of his comments to which they respond.

1. Hawthorne raises a doubt as to how much my anti-luminosity argument achieves, even granted that it soundly shows that all luminous conditions are trivial. For, he claims, it does not refute the claim that all sufficiently simple phenomenal conditions have a property of *coziness*, which is very similar to luminosity. In response to problems about

indexical modes of presentation of conditions, Hawthorne’s final definition of ‘cozy’ is as a predicate of concepts that denote conditions, rather than of the conditions themselves. It is tantamount to this:

A concept of a condition is cozy if and only if in every case in which the concept determinately applies one is in a position to know that the condition obtains by an exercise of the concept.

For ease of comparison, we can use ‘luminous’ as a predicate of concepts too, defined like ‘cozy’ but for the deletion of the word ‘determinately’. Roughly speaking, to be cozy is to be luminous except for borderline cases. The underlying idea is that vagueness is the only obstacle to luminosity for elementary phenomenal concepts (in a suitable sense of ‘elementary’), but not for concepts of external conditions. Hawthorne does not endorse this idea: his point is that it seems to embody a large residue of the phenomenal conception of evidence that I oppose left uneliminated by the anti-luminosity argument.

Friends of the phenomenal conception might argue that vagueness is not a deep obstacle to knowledge of one’s own phenomenal states, because its source is in the concepts with which one classifies those states rather than in the states themselves. In that case, one might expect the phenomenal states to be at least capable of being classified under precise concepts too. For a precise concept of a condition, the ‘determinately’ operator in the definition of ‘cozy’ is redundant, and the concept is cozy if and only if it is luminous. We thus have three theses in play:

(5) Every luminous concept is trivial.

(6) Every elementary phenomenal concept is cozy.

(7) Every precise concept is cozy if and only if luminous.

Of these, (5) is the upshot of the anti-luminosity argument, which Hawthorne is granting; (6) is the proposed revision of the phenomenal conception of evidence; (7) is the corollary just noted of his definitions. Together, (5)-(7) easily imply this:

(8) Every precise elementary phenomenal concept is trivial.

That hardly seems a welcome result for the phenomenal conception of evidence. The obvious moral to draw from (8) is that the nature of the phenomenal is to elude precise characterization. Thus the vagueness of phenomenal concepts has a deep source in the phenomenal states themselves; it is not an accident of our conceptualization.

On an epistemic conception of vagueness, a deep source of vagueness in phenomenal states constitutes a deep source of ignorance in phenomenal states. That is exactly what friends of the phenomenal conception of evidence were trying to avoid. The concession goes so far into the heart of their position as to leave little motivation for holding on to some other vestige of Cartesianism. The view is too radically discredited to confer much epistemological credit on those of its consequences that have not been separately refuted.

Suppose, then, that there is no deep source of ignorance in phenomenal states themselves. By hypothesis, there is a deep source of vagueness in them, which presumably requires an anti-epistemic conception of vagueness. That is already a problematic theoretical commitment for the phenomenal conception of evidence. Moreover, (8) becomes even more puzzling if there is no deep source of ignorance in phenomenal states. Just what is the obstacle supposed to be in principle to forming non-trivial precise elementary phenomenal concepts?

Indeed, (8) is likely to strike friends of the phenomenal conception of evidence as simply false. On their conception, it should be possible to have a vivid memory of a phenomenal experience E, and on that basis to form the phenomenal concept C of having a phenomenal experience qualitatively exactly like E. C is a recognitional concept of E, informed by the memory; it is not the trivial concept *feeling THUS* that Hawthorne discusses. Since a process of gradual change can start with an experience qualitatively exactly like E, which therefore falls under C, and end with an experience qualitatively very different from E, which therefore falls under the negation of C, C is non-trivial in the sense of the anti-luminosity argument. C is a phenomenal concept, for its application-condition is specified in purely phenomenal terms. It is precise in every relevant respect. Finally, C is elementary; it is defined in simple terms, and we may assume that E itself is a simple experience (for example, of a uniform shade of colour); to classify C as not elementary would trivialize the revised phenomenal conception of evidence. Thus C is a counterexample to (8). Given (5) and (7), which Hawthorne does not contest, it is also a counterexample to (6), the revised phenomenal conception of evidence itself. When one has an experience qualitatively exactly like E on a later occasion, C applies determinately, because it is precise, while one may not be in a position to know that one is having an experience qualitatively identical with E, because for all one knows one is having an experience indiscriminably different from E, as one will in fact do a millisecond later; thus one is not in a position to know that the condition that one has an experience qualitatively exactly like E obtains by an exercise of C. Thus even the revised phenomenal conception of evidence fails.

2. Hawthorne goes on to argue that even if we forget about the phenomenal conception of evidence, the equation E=K in *Knowledge and its Limits* of one's total evidence with the total content of one's knowledge still has apparently counterintuitive consequences. He describes two possible worlds thus:

World 2A: I see a gas gauge that reads "Full". The gauge is accurate, and so I come to know that the gas tank is full.

World 2B: I see a gas gauge that reads "Full". The gauge is inaccurate, and so (since knowledge is factive) I do not come to know that the gas tank is full.

If E=K holds then, Hawthorne notes, the observer in World 2A has some evidence (the proposition that the gas tank is full) that the observer in World 2B lacks. But there is a strong inclination to think that the observers in 2A and 2B have exactly the same evidence.

As Hawthorne predicts, I reply that a perspective from which it is plausible that the observers in 2A and 2B have exactly the same evidence is also a perspective from which it is plausible that they have exactly the same knowledge. Both symmetry claims risk unacceptable skepticism, but the present point is just that claims of knowledge and evidence stand or fall together. We can check that point by considering two cases separately.

First, suppose that for all the observer in World 2A knows, he is in World 2B. Then we should retract the description of 2A as a world in which the observer knows that the tank is full. Although the gauge is accurate, that fact is not epistemologically available in the way required for seeing that it reads "Full" to yield knowledge that the tank is full. On this view, in neither world does the observer know that the tank is full. Equally, the observers have the same evidence (in other respects too, we may suppose). Thus there is both symmetry of evidence and symmetry of knowledge; the equation E=K stands.

Second, suppose that it is not the case that for all the observer in World 2A knows, he is in World 2B. Thus, in 2A, the danger of an inaccurate fuel gauge is too remote for 2B to be an epistemic possibility. The observer in 2A is in better epistemic circumstances than is the observer in 2B. Then it seems quite unsurprising that the observer in 2A has better evidence than the observer in 2B: for example, for the proposition that he will get home that night. Thus there is both asymmetry of evidence and asymmetry of knowledge; again, the equation E=K stands.

The idea that such cases involve symmetry of evidence without symmetry of knowledge is an artifact of erroneous philosophical theory.

3. In *Knowledge and its Limits*, I develop an account of evidential probability as probability conditional on the agent's total evidence. Given the equation E=K, evidential probability is probability conditional on whatever the agent knows. Thus whatever the agent knows has evidential probability 1. Hawthorne worries that this will yield counterintuitive results in decision theory. For consider any bet on which I win some amount if *p* is true and lose some amount if *p* is false. Imagine, as Hawthorne does, that the offering of the bet is not itself relevant evidence about the truth-value of *p*. Given my account of evidential probability, if I know anything incompatible with *p*, then the evidential probability of *p* is 0, so the expected value of the bet for me is negative, no matter how much I stand to win if *p* is true and how little I stand to lose if *p* is false. Thus I should reject the bet. But wouldn't it be sensible to accept the bet on the off-chance that I am wrong about *p*, if the potential gains sufficiently outweigh the potential losses? Hawthorne's example is of buying cheap life insurance for the coming year, which is betting on one's own death within that period, even though one knows that one will be on holiday in Blackpool a year from now (which is incompatible with being dead then).

In response to such examples, the sceptic says that one never knew *p* in the first place; the contextualist says that discussing the bet produces a context of utterance in which the word 'know' does not apply to one in respect of *p*; the subject-sensitive invariantist says that being offered the bet destroys one's knowledge of *p*. I will not resort to any of those ideas; I have criticized them elsewhere.[2] All of them involve larger concessions to skepticism than I am willing to make.[3]

It is important to realize that *no* decision theory based on expected utility, calculated according to the standard axioms and definitions of mathematical probability theory, will be everywhere consistent with what pretheoretic common sense predicts a sensible person would do. For let *p* be a complex logical falsehood; more specifically, imagine that *p* is the negation of a complex truth-functional tautology composed from mutually unrelated atomic sentences of unknown truth-value. Then the standard axioms imply that *p* has probability 0, because it is logically false, independently of any connection between knowledge and probability. Consider a bet on which you win ten million dollars if *p* is true and lose one cent if *p* is false. The expected utility of the bet is minus one cent. Nevertheless, a reasonable human might accept the bet: obviously so if she has no time to work out the truth-table for *p*, but even if she has worked it out and found *p* to be false on every line, she might reasonably estimate the chance of a computational error as high enough for the bet to be worth taking on. That is not to say that she would accept millions of such bets, all on *p*. Conversely, consider a bet on which you are horribly tortured to death if *p* is true and win a carrot if *p* is false. The expected utility of the bet is positive, the utility of a carrot. Nevertheless, few reasonable humans would accept the bet, even if they had worked out the truth-table.[4] The penalty for a small computational error is just too high. Reasonable humans have cognitive habits for managing their own fallibility which the probability calculus makes no attempt to reflect. Of course, one could try to construct a non-standard probability calculus in which truth-functional tautologies can have probability less than 1, but such modifications tend to make huge sacrifices in mathematical power for tiny gains in psychological realism.

Perhaps a perfectly rational agent, transparently incapable of computational error, would reject the former bet (with expected utility minus one cent) and accept the latter (with the expected utility of a carrot). If so, it does not follow that it would be reasonable for ordinary mortals to make the same choices. If a perfectly rational agent would answer 'Yes' to the question 'Are you a perfectly rational agent?', it does not follow that it would be reasonable for us to give the same answer.

Thus, quite independently of the knowledge-based approach in *Knowledge and its Limits*, we can expect a mismatch between the good cognitive habits of reasonable mortals and any notion of rationality derived from a mathematically developed decision theory. The mismatch does not mean that either the cognitive habits are 'wrong' or the notion of rationality is. It is naïve to treat a formal decision theory as a practical guide to life. The decision theory may be right about rationality, in a theoretically purified sense of the term, even if the long-term results of attempting to apply the decision theory would be disastrous. In the case of the two bets above, by applying any standard decision theory one will in fact avoid wasting a cent on the former bet and gain a carrot on the latter. The problem is that occasionally in other cases with a complex true proposition p^* in place of p one will make a computational error and mistakenly conclude that p^* is a truth-functional contradiction. Then, in attempting to apply a standard decision theory, one will reject the first bet and miss the opportunity of winning ten million dollars, but accept the second bet and be horribly tortured to death. Of course, in so doing one is misapplying the decision theory: but one must take the risk of misapplication into account when deciding whether to treat a decision theory as a practical guide to life. Indeed, given plausible auxiliary assumptions, a decision theory may imply that the expected utility of *attempting* to apply that very theory in practice is negative, and therefore counsel against it. To dismiss the theory on those grounds alone as self-defeating would be simplistic, for its aim need not be so crudely practical.

These morals are reinforced when one considers the relation between irrationality and knowledge of irrationality. For most acts D, the condition that it is irrational for one to do D is non-trivial in the sense of the anti-luminosity argument. A process of gradual change can take one from circumstances in which D is irrational to circumstances in which D is rational. Therefore, by the argument, there are cases in which D is irrational, although the agent is not in a position to know that D is irrational. This is an outcome of the anti-luminosity argument itself, independently of any decision-theoretic connection between knowledge and rationality. It is not an easy result to live with, because we are so strongly inclined to think that the demands of rationality must be accessible to the agent. But no non-trivial standard of rationality meets that constraint.[5]

When we contemplate cases in which D is irrational but the agent is not in a position to know that D is irrational, we are likely to feel conflicted, perhaps to say that D is irrational only in an 'external' sense of 'irrational' and that in a more 'internal', perhaps more 'genuine', sense of the word D is not irrational. For any sense of 'irrational', irrational(1), we can define a new sense, irrational(2), such that something is irrational(2) if and only if the agent is in a position to know that it is irrational(1). But irrationality(2) is just as vulnerable as irrationality(1) to the anti-luminosity argument. There will be cases in which something is irrational(2) although the agent is not in a position to know that it is irrational(2). Iterating 'know' makes no difference to the overall structure of the problem.[6]

When stakes are high, agents with good cognitive habits who do not know that D is irrational may do D although D is in fact irrational. The proper response is not to change the meaning of the word ‘irrational’ but to appreciate the bind in which reasonable agents must act.

Since Hawthorne is granting the anti-luminosity considerations, he should concede that something can be irrational even when one is not in a position to know that it is irrational. But he does not acknowledge this point in discussing his example. He writes:

> Williamson might claim that it *is* irrational to buy the life insurance, but one does not know that it is irrational because one does not know that one knows that one will be in Blackpool in a year’s time. Few will be brought around to this view.

He elaborates the latter sentence in a footnote:

> Apart from the obvious strain in the claim of irrationality, it is not clear that we should be too quick to concede that one does not know that one knows this or that proposition about the future.

Given that Hawthorne grants that there are cases in which it is irrational to buy the life insurance but one is not in a position to know that it is irrational, what does he expect such cases to be like?

We may agree with Hawthorne's suggestion that we often know that we know things about the future. But then, by the anti-luminosity argument again, there will be cases in which I know that I know that I shall be in Blackpool next year without being in a position to know that I know that I know that I shall be in Blackpool next year. In such cases, I may know that it is irrational to buy the life insurance without being in a position to know that I know that it is irrational to buy the life insurance, and so on for even higher iterations. If stakes are high enough, human agents with good cognitive habits may do something which they know to be irrational, when they do not know that they know that it is irrational, or do something which they know that they know to be irrational, when they do not know that they know that they know that it is irrational, or […].

Either I am in a position to have arbitrarily high iterations of knowledge that I shall be on holiday in Blackpool in a year's time or I am not. If I am, then my position is so epistemologically privileged that buying life insurance for the next year would indeed be a waste of money. If not, then there is some iteration of knowledge that I am not in a position to have. Consequently, I am not in a position to have some iteration of knowledge that buying life insurance for the next year is irrational. If the insurance is cheap enough, good cognitive habits may lead me to buy the insurance.

Some will be tempted to build an alternative account of rationality around the 'good cognitive habits'. But such habits are too loose and contingent on the accidents of human psychology to provide a systematic decision theory. Nor do they solve the underlying problem, for they are just as vulnerable as anything else to the anti-luminosity argument. Sometimes good cognitive habits would lead one to do something although one is not in a position to know that good cognitive habits would lead one to do it.

4. Hawthorne raises a further doubt about the equation E=K, as to whether evidential probabilities at a time supervene on what is known at that time. Here is his example:

> I look at a reliable newspaper and form a belief that Q that has evidential probability .8; meanwhile, you look at a less reliable newspaper and form a belief that R that has evidential probability .3. We both later forget where we read Q and R respectively, though we both remember that we read the relevant claim in a newspaper and that newspapers are for the most part reliable. Suppose Q and R both entail S. Intuitively, I have better evidence for S than you do. But there is no difference between us on the score of what is known.

The example is somewhat under-described. It is not clear from the initial description that 'there is no difference between us on the score of what is known'. Is each of us fully aware of the other's mental state? If not, and I believe Q but not R while you believe R but not Q, then I may know that someone believes Q and not know that someone believes R, while you know that someone believes R and fail to know that someone believes Q. If each of us is fully aware of their own and the other's mental state, that may in effect pool our evidential resources, which makes it less clear how I can have better evidence for S than you have. Again, the example sounds like one in which I know that I once had fairly good evidence for S (which is itself some evidence for S) while you believe without knowing that you once had fairly good evidence for S. If Hawthorne merely stipulates that there is no difference between us on the score of what is known, the problem is that we are unsure how to imagine such a case, and there is nothing very intuitive about the judgement that I have better evidence for S than you have. Indeed, one wonders what this ‘better evidence’ is supposed to be. Do I have some evidence that you lack, or *vice versa*? If so, what is it? If we have exactly the same evidence, how does it support S more in my case than in yours?

Hawthorne has not made it plausible that there can be differences in evidential probability without corresponding differences in knowledge. I continue to maintain that evidential probability supervenes on knowledge.

5. Hawthorne presents a final puzzle. He says that each of these three propositions seems very plausible (I have changed his numbering):

(9) We know all sorts of propositions about the future.

(10) For pretty much any empirical proposition about the future, there is at least a small objective chance that it will not obtain.

(11) Propositions deduced from a set of known premises are themselves known.

We are to agree, at least for the sake of argument, that we can assert (10) on the basis of quantum mechanics. According to Hawthorne, (9)-(11) ‘would seem to entail’ this:

(12) There can be truths of the form: S knows that P about the future and it is overwhelmingly [objectively] likely that not P.

Hawthorne explains the seeming entailment by saying ‘small objective chances add up to big ones’.

The first point to note about the puzzle is that it is not specific to my account of knowledge. Probably the great majority of philosophers would be inclined to reject (12) but accept each of (9)-(11), at least when suitably cleaned up and presented separately (in the case of (10), provisionally on a suitable interpretation of quantum mechanics). The puzzle is a puzzle for almost everyone. The question is whether my account has special resources for solving it not available to other accounts.

The second point to note is that (9)-(11) do not entail (12). To see this, let us construct a coherent toy model of knowledge on which (9)-(11) are true while (12) is false.

For convenience, work in terms of possible worlds. Think of propositions as sets of worlds. For each time *t* and world *w*, $\mathrm{OC}_{t,w}$ is the probability distribution that gives the objective chances of propositions at *t* in *w*. Thus if the proposition *p* is true at every world exactly like *w* up to *t* then $\mathrm{OC}_{t,w}(p) = 1$, because the past is certain. If *p* is true at some such worlds but false at others then $\mathrm{OC}_{t,w}(p)$ may be strictly between 0 and 1, because the future is uncertain. We may therefore assume that (10) holds in the model.

Suppose that for each time *t* and world *w* there is a similarity sphere $\$_{t,w}$ of worlds around *w*, differing from *w* in basic physical respects at most in what happens after *t*, such that $0.85 \leq \mathrm{OC}_{t,w}(\$_{t,w}) \leq 0.95$ (in what follows the qualification ‘in basic physical respects' will usually be left tacit). Thus at t in w the objective chance that one is in a world similar enough to w to belong to the similarity sphere is roughly 0.9. The numbers 0.85, 0.9 and 0.95 are chosen for illustrative purposes only; many other similar triples would serve the point just as well. Now make the grossly simplifying assumption that a necessary and sufficient condition for p to be known at t in w is that p is true at every world in the similarity sphere $\$_{t,w}$ (in other words, that $\$_{t,w}$ entails p). Thus one knows exactly what the past and present are like and roughly what the future is like. The condition ensures that if at t in w one knows the premises of a valid deductive argument, then at t in w one also knows its conclusion. Thus (11) holds in the model.

Claim (9) also holds in the model, because $\$_{t,w}$ entails all sorts of propositions about the future. It is not itself about the future, for it is true at w yet false at a world exactly like w in what happens after t but different in what happens before. However, let $p_{t,w}$ be the proposition which is true at a world x if and only if $\$_{t,w}$ is true at some world y exactly like x in what happens after t. Thus $p_{t,w}$ is about the future at t in the sense that if two worlds x and x^* are exactly alike in what happens after t then $p_{t,w}$ is true at x if and only if $p_{t,w}$ is true at x^*, for a world y is exactly like x in what happens after t if and only if it is exactly like x^* in what happens after t. To calculate the objective chance of $p_{t,w}$, let $h_{t,w}$ be the history of w up to t; more exactly, $h_{t,w}$ is true at a world x if and only if x and w differ at most in what happens after t. Since objective uncertainty is confined to the future, $\mathrm{OC}_{t,w}(h_{t,w}) = 1$. Now $h_{t,w}$ & $p_{t,w}$ entails $\$_{t,w}$, for if $h_{t,w}$ & $p_{t,w}$ is true at x then $\$_{t,w}$ is true at some world y exactly like x in what happens after t, so $h_{t,w}$ is also true at y, so x and y differ at most in what happens after t; thus y must be x itself, in which case $\$_{t,w}$ is true at x. Conversely, $\$_{t,w}$ entails both $h_{t,w}$ (by hypothesis) and $p_{t,w}$ (because every world is exactly like itself in what happens after t); by the latter entailment, $p_{t,w}$ is known at t in w. Consequently, $\mathrm{OC}_{t,w}(p_{t,w}) = \mathrm{OC}_{t,w}(h_{t,w} \mathbin{\&} p_{t,w}) = \mathrm{OC}_{t,w}(\$_{t,w}) \leq 0.95$. Thus $p_{t,w}$ is a known proposition with an objective chance of at most 0.95. All its logical consequences are also known at t in w.

The disturbing claim (12) fails in the model. For if p is known at t in w then $\$_{t,w}$ entails p, so $0.85 \leq \mathrm{OC}_{t,w}(\$_{t,w}) \leq \mathrm{OC}_{t,w}(p)$. Thus no proposition whose negation is objectively overwhelmingly likely is ever known.

Of course, the foregoing model of knowledge is far too coarse-grained, since it automatically counts all necessary truths, truths about the past and necessary consequences of known truths as known. But it serves to establish the logical point that (9)-(11) do not entail (12). Moreover, we can evidently impose further necessary conditions on knowledge (such as that it entails belief) in ways that preserves (9) and (11). Since (10) will still hold and (12) fail, we can thereby construct more realistic models of (9)-(11) without (12).

The model does require that what is known at t in w is not determined by the history of w up to and including t. For since $\mathrm{OC}_{t,w}(\$_{t,w})$ is not 1, some world w^* in which $\$_{t,w}$ is false and therefore not known at t is objectively possible at t in w. Even though w and w^* have the same history up to and including t, $\$_{t,w}$ is known at t in the former but not in the latter. Although that is a difference between w and w^* in what is the case at t, it does not count as a difference in their history up to and including t because ‘history’ was defined above in a way that effectively restricts it to basic physical respects. But this feature of the model is required by (9) and (10), independently of (11). For if at t in w one knows a proposition p about the future whose objective chance at t in w is less than 1, then some world *w*** in which *p* is false and therefore not known at *t* is objectively possible at *t* in *w*. Even though *w* and *w*** have the same history up to and including *t*, *p* is known at *t* in the former but not in the latter. Knowledge at a time does not supervene on basic physical history up to and including that time. If one does not like that result, indeterminism will force one to give up most knowledge of the future.

If one combines (9) and (10) with my account of evidential probability, the upshot is that an evidential probability of 1 does not entail an objective chance of 1. If *p* is known but has a small positive chance of being false, then its evidential probability is 1 but its objective chance is less than 1. The converse point is, of course, uncontroversial. An evidential probability of less than 1 does not entail an objective chance of less than 1. If *p* is a completely unknown truth about the past, its evidential probability is less than 1 but its objective chance is 1. Evidential probability and objective chance are very different things.

The foregoing remarks obviously do not pretend to be a full epistemology of truths about the future. But they do show that Hawthorne's final puzzle as he poses it can be solved. The solution draws on the epistemology of *Knowledge and its Limits*. For the principle that something is known in a world *w* only if it is true in every possible world similar enough to *w* to belong to the relevant similarity sphere is a *margin for error* principle in the sense of the book (p. 129). Such principles are crucial to understanding the failures of luminosity. Consequently, Hawthorne's puzzle provides support for my account of knowledge.

Reply to Yablo

In 'Prime Causation', Steve Yablo discusses my account of the value of prime conditions, conditions that cannot be factorised into the conjunction of purely internal and purely external conjuncts, in the causal explanation of action. He does so in a congenial spirit. In *Knowledge and its Limits* I cited his work on the proportionality of causes to effects in support. His new discussion provides a more fine-grained analysis of the trade-off between proportionality and naturalness in the competition amongst candidate causes. He writes 'I complain essentially that Tim is right for more reasons than he gives'. Nor is his friendly fire the sort that causes unintended casualties. Nevertheless, some differences of approach still divide us. I will comment on several of them.

It is convenient to begin by laying out Yablo's 'Proportionality Theory of Causal Relevance'. Here is the main idea:

> A property P of *x* is causally relevant to effect *y* iff
>
> (a) had *x* occurred without P, *y* would not have ensued.
>
> (b) P is not egregiously weak or egregiously strong.

Yablo toys with adding a third conjunct (c), to the effect that P is not egregiously strong- or weak-making; it is irrelevant to what follows here. What do 'egregiously weak' and 'egregiously strong' mean? Yablo defines them thus:

> *x*'s property P is egregiously weak (relative to *y*) iff some more natural stronger property of *x* is better proportioned to *y* than P is.

> *x*'s property P is egregiously strong (relative to *y*) iff some as natural weaker property of *x* is better proportioned to *y* than P is.

In a footnote, Yablo softens these definitions with a suggestion from Alex Byrne: 'Strictly we should allow for small drops in naturalness to be compensated by large enough gains in proportionality'. But what does 'better proportioned' mean? Yablo offers these constraints:

> A weaker property of *x* is better proportioned to *y* than a stronger one iff *y* would still have occurred, had *x* occurred with the weaker property but not the stronger one.

> A stronger property of *x* is better proportioned to *y* than a weaker one iff *y* would not have occurred, had *x* occurred with the weaker property but not the stronger one.[7]

Although Yablo is not very explicit as to what kind of entity the variables '*x*' and '*y*' are supposed to range over, his use of the word 'occurred' suggests that they are events; that is what I will assume.

One difference between Yablo and me is that he proposes strictly necessary and sufficient conditions for causal relevance in apparently non-circular terms: causal concepts do not obviously occur in his clauses (a) and (b) or the associated explanations of ‘egregiously weak/strong’ and ‘better proportioned’. That contrasts with the general scepticism in *Knowledge and its Limits* about the programme of identifying strictly necessary and sufficient conditions in non-circular terms for the correct application of philosophically significant concepts. It is not only the concept of knowledge that is unanalysable. My comments on causal relevance were certainly not intended to provide such conditions: non-circular necessary and sufficient conditions for X are not a prerequisite for informative theorizing about X. In particular, I see no reason to believe that such conditions can be stated for causal relevance (on this matter, appeals to Hume leave me unmoved). However, Yablo is not arguing on general grounds that there must be some such conditions; he is presenting a particular candidate. Merely to point to the poor record of previous candidates would not be neither a satisfying response.

One worry about Yablo’s analysis concerns essential properties of events. By calling P an ‘essential property’ of an event *x* here I simply mean that necessarily, if *x* occurs then *x* has P. Suppose that P is an essential property of *x*, and that some property Q is at least as natural as P and weaker than it: having P is sufficient but not necessary for having Q. Since *x* could not have occurred without P, the counterfactual ‘*y* would still have occurred, had *x* occurred with the weaker property [Q] but not the stronger one [P]’ has an impossible antecedent. Presumably, counterfactuals with impossible antecedents are vacuously true. Thus, by Yablo’s first constraint on ‘better proportioned’, Q is better proportioned than P to any event *y* (by his second constraint, P is also better proportioned than Q to *y*, but that is by the way). By hypothesis, Q is as natural as P, so P is egregiously strong (relative to any event *y*) by Yablo's definition. Consequently, given his Proportionality Theory of Causal Relevance, P is not causally relevant to any effect *y*. Is that result plausible? Let *x* be the eruption of Vesuvius in the year 79, *y* the destruction of Pompeii, P the property of being a volcanic eruption and Q the property of being self-identical. Presumably, P is an essential property of *x*; that event could not have occurred without being a volcanic eruption. Q is weaker than P: being a volcanic eruption is sufficient but not necessary for being self-identical. Q is arguably at least as natural as P, or more natural: the boundaries in extension of being self-identical seem less arbitrary than those of being a volcanic eruption (consider borderline cases for 'is a volcanic eruption'). Given these assumptions, Yablo's theory implies that the property of *x* of being a volcanic eruption was not causally relevant to effect *y*, the destruction of Pompeii. Surely that is the wrong result. Of course, one could quarrel with details of the example, but that seems an unpromising strategy: it is hard to believe that *no* essential property of an event stronger than some no less natural property is ever causally relevant to an effect.

Perhaps some tweaking of the background ontological framework of Yablo's theory would solve that problem. In *Knowledge and its Limits*, the discussion proceeds in terms of conditions rather than events, and no analogous problem seems to arise. But for present purposes, I will now bracket the foregoing issue and continue to use the language of events.

Probably a more important difference is that Yablo's theory is formulated in terms of counterfactual conditionals, whereas my discussion exploits the mathematical and conceptual resources of probability theory. Since our purposes differ somewhat, direct comparison is hard; nevertheless, some comments can be made. For simplicity, we can concentrate on Yablo's commitment to his conjunct (a) as a necessary condition of causal relevance, since that is independent of his tentative suggestions for fine-tuning the definition of 'egregiously weak/strong' and adding a third conjunct (c). Is a property P of *x* causally relevant to effect *y* only if, had *x* occurred without P, *y* would not have ensued?

Suppose that some events are objectively chancy, without underlying deterministic mechanisms, and that had *x* occurred without P, the objective chance of *y* ensuing would have been much lower than it actually was, but still far from zero. Imagine *y* as an unremarkable event, of the same kind as the alternative effects that would have ensued if *y* had not. Then presumably it is not true that had *x* occurred without P, *y* would not have ensued (which is not to say that it is true that had *x* occurred without P, *y* would have ensued). But it does not at all seem to follow that P was not causally relevant to *y*. P might be exactly the property of *x* that one would naturally think of as causally relevant to *y* because it made *y* objectively almost certain to ensue. It is far from clear that counterfactual-based theories such as Yablo's give us what we need to understand causal relevance in an indeterministic world.

The necessity of Yablo's counterfactual condition (a) for causal relevance can be challenged even with respect to deterministic worlds. Let the events *x* and *y* be a bank robbery and the capture of the bank robbers respectively, and P be the property of *x* of being such that the getaway car fails to start. We can easily imagine P being causally relevant to *y*. Suppose that in masses of close worlds the bank robbery occurs, the getaway car starts and the bank robbers are not captured. However, in a very few close worlds, the car starts falteringly and the bank robbers are still captured. In fact, if the bank robbery had occurred without the car failing to start, it would or at least might have started falteringly, with the ensuing capture of the robbers. ‘Might’ counterfactuals are to be read here as the duals of the corresponding ‘would’ counterfactuals, so that ‘If *x* had occurred without P, *y* might have ensued’ is equivalent to ‘It is not the case that if *x* had occurred without P, *y* would not have ensued’. Thus it is not the case that if the bank robbery had occurred without the failure of the getaway car to start, the capture of the bank robbers would not have ensued. Does that show that the failure of the getaway car to start was not causally relevant to the capture of the bank robbers? Not in any ordinary sense of ‘causally relevant’; Yablo’s aim is not merely to stipulate a new sense for the phrase. Probabilistic notions have more promise here, even though one should not expect them to provide strictly necessary and sufficient conditions for causal relevance. In the circumstances, the probability that the bank robbers are captured is very high given that the car fails to start and very low given that it starts. For although the car might have started falteringly if it had started at all, the probability that it starts falteringly given that it starts is very low. The comparative naturalness of the property of being such that the getaway car fails to start also helps to compensate for the untruth of (a). The alternative property of being such that the getaway car fails to start unfalteringly seems, if anything, slightly less natural, since it lumps together the case in which it starts falteringly with the rather different case in which it fails to start at all. But why resist the inclination to say that both properties are causally relevant? Some philosophers (not Yablo) would press the naïve question ‘Which property did the causal work?’ as though it were simply the general form of ‘Whose finger was on the button?’: but properties are not rival agents, and nothing in the notion of causal relevance undermines the hypothesis that many different properties were causally relevant to a given effect.

Pre-emption constitutes a more general challenge to the necessity of (a) for causal relevance. Yablo refers the reader to other work of his for a treatment of that problem.[8] The bank robbery is not a classic case of pre-emption; to discuss whether anything Yablo says about pre-emption would help with the bank robbery would take us too far afield. At any rate, Yablo's condition (a) as stated in his present paper is unnecessary for causal relevance.

Much that Yablo says about causal relevance his examples seems illuminating and right. But one can appreciate such insights without expecting to build them into a non-circular statement of strictly necessary and sufficient conditions for causal relevance, perhaps rather as one can appreciate the insights of a good art critic into the real aesthetic qualities of artworks without expecting to build them into a non-circular statement of strictly necessary and sufficient conditions for artistic greatness. For both, insight depends on bringing the particular case under general concepts, but always against a background of circumstances too rich for us to express in a finite non-circular formula.

Notes

[*] I thank the four commentators for the time and energy they spent on my book, both them and other participants in the 2003 Rutgers conference on epistemology for their pertinent questions, and the Centre for Advanced Study at the Norwegian Academy of Science and Letters for its generous hospitality during the writing of most of these replies. For uniformity rather than stand-offishness, I use surnames throughout.

[1] In his section 3, Brueckner suggests a different criticism of my account of evidence, concerning spontaneous generation. However, it depends on Dretske's view that evidence can enable one to know a truth without enabling one to know an obvious consequence of that truth. Since he offers no reason why I should accept that highly controversial view, I discuss this criticism no further.

[2] I discuss scepticism in chapter 8 of *Knowledge and its Limits*, contextualism in 'Knowledge, Context and the Agent's Point of View', forthcoming in G. Preyer and G. Peter (eds.), *Contextualism in Philosophy* (Oxford: Oxford University Press) and 'Contextualism, Subject-Sensitive Invariantism, and Knowledge of Knowledge', *Philosophical Quarterly*, forthcoming, and subject-sensitive invariantism in the latter.

[3] Hawthorne cites a passage (p. 255) in which, discussing the connection between knowledge and assertion, I maintain that I should not assert that I shall not be knocked

down by a bus tomorrow, because I do not know that I shall not be knocked down by a bus tomorrow. He worries that I might be committed to a more general scepticism about knowledge of the future. No such commitment was intended. I was commenting on the particular example, perhaps with unnecessary pessimism about bus drivers. The important point is that from a perspective on which it is plausible that I should not assert that I shall not be knocked down by a bus tomorrow (for epistemic reasons rather than to avoid tempting fate), it is equally plausible that I do not know that I shall not be knocked down by a bus tomorrow. In any case, as Hawthorne notes, epistemological issues about insurance are not confined to knowledge of the future; one might be offered insurance against past contingencies which one has not yet specifically checked.

4 The remark mentioned in a footnote by Hawthorne about (the British equivalent of) not betting the farm against a dime on excluded middle was intended with a similar purpose. The issue is not *knowledge* of excluded middle, since it has probability 1 by the probability axioms without needing to be conditionalized on.

5 A useful indication of the robustness of the anti-luminosity argument in the present context is that analogues of it go through for concepts of evidential probability in place of the concept of being in a position to know, independently of any assumptions about the link between evidential probability and knowledge; see my 'Probabilistic Anti-Luminosity', forthcoming in Q. Smith (ed.), *Epistemology: New Philosophical Essays* (Oxford: Oxford University Press).

[6] If one insists that the agent must have all finite iterations of knowledge, one risks imposing a condition that human agents cannot meet (see also pp. 121-3).

[7] I have reworded Yablo's constraint on 'worse proportioned' in terms of 'better proportioned' to make its bearing on egregiousness more perspicuous.

[8] S. Yablo, 'De Facto Dependence', *Journal of Philosophy* 99 (2002): 130-148.