IS SAFETY IN DANGER?

Fernando Broncano-Berrocal

Abstract: In "Knowledge Under Threat" (Philosophy and Phenomenological Research 2012), Tomas Bogardus proposes a counterexample to the safety condition for knowledge. Bogardus argues that the case demonstrates that unsafe knowledge is possible. I argue that the case just corroborates the well-known requirement that modal conditions like safety must be relativized to methods of belief formation. I explore several ways of relativizing safety to belief-forming methods and I argue that none is adequate: if methods were individuated in those ways, safety would fail to explain several much-discussed cases. I then propose a plausible externalist principle of method individuation. On the one hand, relativizing safety to belief-forming methods in the way suggested allows the defender of safety to account for the cases. On the other hand, it shows that the target known belief of Bogardus's example is safe. Finally, I offer a diagnosis of a common error about the kind of cases that are typically considered potential counterexamples to the necessity of the epistemic condition: proponents of the alleged counterexamples mistake a strong condition that I call super-safety for safety.

Keywords: methods of belief formation; Bogardus; knowledge; reliability; safety.

One knows a proposition only if one's belief could not easily have been false. This is, roughly, the safety condition for knowledge. The condition has been proven to be a valuable tool in the analysis of epistemic luck, the Gettier problem, lottery puzzles and skepticism, and hence, although not everyone is willing to endorse it, safety stands as one of the pillars of contemporary theory of knowledge, not only because of its fruitful applications, but also because of its resistance to counterexamples.

Recently, several commentators (Comesaña 2005; Kelp 2009; Neta & Rohrbaugh 2004) have offered a series of counterexamples that have the aim of showing that unsafe knowledge is possible. Tomas Bogardus (2012) has revealed a flaw in the argumentation provided by these authors. He argues that the cases proposed rest on a fallacious move: from the fact that S was at epistemic risk just before forming her belief, it is inferred the conclusion that S's belief was formed unsafely. In other words, according to Bogardus, the cases are not

counterexamples to safety because they locate epistemic risk *before* the target beliefs are formed.[1] For Bogardus, a genuine counterexample to the necessity of safety would be a case in which the agent in question is at epistemic risk *when* she forms the relevant belief and not merely before. Accordingly, he proposes the following case:

ATOMIC CLOCK

[T]he world's most accurate clock hangs in Smith's office at a cereal factory, and Smith knows this. The clock's accuracy is due to a clever radiation sensor, which keeps time by detecting the transition between two energy levels in cesium-133 atoms. This radiation sensor is very sensitive, however, and could easily malfunction if a radioactive isotope were to decay in the vicinity (a very unlikely event, given that Smith works in a cereal factory).

This morning, against the odds, someone did in fact leave a small amount of a radioactive isotope near the world's most accurate clock in Smith's office. This alien isotope has a relatively short half-life, but—quite improbably—it has not yet decayed at all. It is 8:20 am. The alien isotope will decay at any moment, but it is indeterminate when exactly it will decay. Whenever it does, it will disrupt the clock's sensor, and freeze the clock on the reading "8:22." (Don't ask why; it's complicated.) Therefore, though it is currently functioning properly, the clock's sensor is not safe. The clock is in danger of stopping at any moment, even while it currently continues to be the world's most accurate clock.

Smith is quite punctual, and virtually always arrives in her office on workdays between 8:20 and 8:25 am, though no particular time in that duration is more likely than any other to see her arrive. Upon entering her office, Smith always looks up at her clock and notes the time of her arrival.

---

1 I will take for granted that Bogardus succeeds in defending the safety condition from the counterexamples proposed by the aforementioned authors. However, to illustrate what is the problem with the inference made by these authors, let us consider an analogous problem identified by E.J. Coffman in the debate on luck. Coffman (2009: 503) argues that the following thesis (which he calls the luck infection thesis) has blatant counterexamples: if it is by luck that S is positioned to φ, then it is by luck that S has φ-ed. For example, it can be by luck that Kobe Bryant has found a basketball to make a slam dunk, but once he has the basketball it is not by luck that he makes it. Analogously, from the fact that it is by luck that S is in a position to know that *p* it does not necessarily follow that S luckily knows that *p*, or in other words, from the fact that S was at risk of not being in a position to know that *p* it does not necessarily follow that S's belief that *p* is formed unsafely. To see this, imagine that God flips a coin to decide whether to deprive Adam of vision entirely. Heads: God deprives Adam of vision; tails: God does not do anything. Thousandths of seconds after the coin lands tails, Adam looks at his watch and forms the belief that it is 8:00 am. On reflection, Adam was at risk of losing his good epistemic position *before* forming his belief. However, this does not mean that his belief is formed unsafely, because by the time he forms it, he is already in a good epistemic position and, intuitively, he can know that it is 8:00 am.

Today, in the actual world ("@"), that alien isotope has not yet decayed, and so the clock is running normally at 8:22 am when Smith enters her office. Smith takes a good hard look at the world's most accurate clock—what she knows is an extremely well-designed clock that has never been tampered with—and forms the true belief that it is 8:22 am (Bogardus 2012: 12-13).

In this paper, I will show two things. In particular, that ATOMIC CLOCK does not prove that knowledge and safety can come apart, as Bogardus thinks. In general, that known beliefs are safe even if the agents in question are at epistemic risk when they form them.

1 Setting the Scene

For the sake of the argument, I will use the following definition of safe belief:

SAFE BELIEF: S's belief that p formed in the actual world @ via a belief-forming method of type M is safe iff 1) it is true in @ and 2) in nearly all, if not all, close possible worlds in which S forms the belief that p via a belief-forming method of type M, that belief is true.[2]

The *safety condition* says that S knows that p only if S's belief that p is safe. Bogardus's argument against the necessity of safety is as simple as follows:

(1) Smith knows that it is 8:22 am via certain belief-forming method (e.g. looking at the clock) in @.

(2) Smith believes *falsely* that it is 8:22 am via the same type of belief-forming method in very many close possible worlds (i.e. Smith's belief is unsafe).

Therefore, unsafe knowledge is possible.

(1) is not only intuitively true, but it is also considered true by many respectable theories of knowledge.[3] Therefore, we must grant that Smith knows that it is 8:22 am when she looks at the clock in @. Let us concentrate then on premise (2). In order to argue against Bogardus's criticism of safety, I aim to show why (2), as it stands, is false. First, let us state the reason we find it true at first glance. To that aim, it is useful to start with a modal definition of risk in general:

RISK: An event is at risk of occurring iff it would occur in a large enough proportion of close possible worlds.

How large should be the relevant proportion of close possible worlds in order for an event to be considered risky largely depends on the type of event at issue. In

---

2 This definition of safe belief is based on Duncan Pritchard's version of safety (2005: 163).
3 Bogardus is particularly interested in showing the truth of (1) and reviews, accordingly, a representative set of theories.

most cases, we would not say that an event is risky unless the proportion of close possible worlds in which it would occur is a large one. Plausibly, this applies to mundane events such as raining, someone dropping a coffee cup or a bulb blowing. However, we sometimes consider an event risky when the relevant proportion of close possible worlds in which it would occur is not large. For example, suppose that an agent plays Russian roulette with one bullet in the chamber of a revolver with a 6-shot capacity. Presumably, given the low probability of being shot (approximately 16%), the agent would not die in a large proportion of close possible worlds. Still, we would intuitively say that the agent is at risk of dying.[4]

Interestingly, the notion of risk also applies to the epistemic case. One type of epistemic risk is the risk that an agent's belief has of being false.[5] Following RISK, an agent's belief is epistemically risky in this sense just in case it would be false in certain proportion of close possible worlds. How large should that proportion be? It is not easy to give a precise answer, but we would not say that a belief is epistemically risky (at least in a significant sense) if the proportion of close possible worlds in which it would be false were a small one. Plausibly, epistemic risk can be defined in the following terms:

> EPISTEMIC RISK: S's belief that p is at risk of being false iff p is false in at least half the close possible worlds in which S believes that p.

According to EPISTEMIC RISK, an agent who uses a belief-forming method that is as reliable as tossing a coin is at risk of forming false beliefs: the proportion of guesses would be no better than chance. However, in other cases the risk of believing a false proposition is much higher. These are the kind of cases I am interested in. For what I ultimately want to show is that safety is consistent with cases of *high* epistemic risk. We can define high epistemic risk in the following way:

> HIGH EPISTEMIC RISK: S's belief that p is at high risk of being false iff p is false in *most* close possible worlds in which S believes that p.

_____

4 Here is a plausible explanation of why our intuitions about risk vary depending on the type of event: the threshold beyond which the proportion is large enough is a function of the significance of the event for the agent (I borrow the idea from Neil Levy (2011: 18), who introduces it in his analysis of the concept of luck). We can be more specific: the more significant an event is for an agent, the smaller needs to be the proportion of close possible worlds in which it would occur to be risky for the agent. Say that there is 16% probability that you die playing Russian roulette and 16% probability that it rains. Since the possibility of dying is very significant for you (much more than the possibility of it raining), it is reasonable to say that you are at risk of dying when playing Russian roulette even though you would not *easily* die. However, we would not say that you are at risk (at least at an equivalent risk) of getting wet when strolling due to the rain.

5 Other types of epistemic risk are the risk of believing a closely related false proposition, a proposition with gappy content, with paradoxical content or with no content at all. See Hiller & Neta (2007) and especially Manley (2007) for further discussion on the matter.

If we reconsider ATOMIC CLOCK, we notice that Smith's belief that it is 8:22 am is epistemically risky in this sense because her belief would be false in very many close possible worlds. In those worlds, the alien isotope decays disrupting the sensor and freezing the clock on the reading "8:22" before Smith looks at it at a time subsequent to 8:22 am. I will argue that, although Smith would easily have formed the false belief that it is 8:22 am, i.e. although her actual known belief is epistemically risky in the sense specified above, her belief is not unsafe (in the sense specified by SAFE BELIEF). In particular, the reason Smith's actual belief counts as safe is that:

> ~(2) in close possible worlds where Smith forms the false belief that it is 8:22 am, she does *not* form the belief via *the same type* of belief-forming method that she uses in @.

~(2) needs justification (one of the goals of the paper is to provide such a justification). For now, note that ~(2) is *compatible* with the fact that Smith would believe falsely that it is 8:22 am in many close possible worlds, i.e. with the fact that Smith's actual known belief is epistemically risky in the sense specified by HIGH EPISTEMIC RISK. Precisely, —here comes the explanation sought— premise (2) seems true at first glance because Smith's belief is epistemically risky in this sense.

Contra this intuition, I will argue that not every case of epistemic risk is a case of unsafe belief. In other words, I will argue that there are epistemically risky safe beliefs. More specifically, I will defend this claim by arguing that safety is a stringent condition that does not require match between belief and fact across close possible worlds *simpliciter*, but match between belief and fact across close possible worlds in which the agent acquires the relevant belief via the same type of belief-forming method that she uses in @.

First of all, we need to show why modal conditions need to be relativized to methods of belief formation. The following case illustrates why safety requires such a relativization:

GRANDMOTHER

A grandmother goes to the hospital to visit her grandson. The grandson's heart is in such a very bad condition that he could easily suffer a deathly cardiac arrest at any moment. If the grandson died, others would tell the grandmother he was alive to spare her upset. The grandmother enters the hospital room and sees that the grandson is alive.[6]

---

6 The case is inspired by Robert Nozick's grandmother case (Nozick 1981: 179). I have modified the original case in such a way that the target proposition could easily be false, as in ATOMIC CLOCK.

Intuitively, the grandmother knows that her grandson is alive when she looks at him. Nevertheless, the grandson could easily have died and the grandmother could easily have formed the false belief that he was alive (because others would have told her so). That is, the grandmother's actual belief is epistemically risky in the sense defined by HIGH EPISTEMIC RISK.

However, from the fact that her belief is epistemically risky in the specified sense, it does not follow that it is unsafe. According to SAFE BELIEF, the target belief counts as safe because close possible worlds in which the old woman believes that her grandson is alive via *the same type* of method that she uses in @ (a visual method) are worlds in which the grandson has not yet suffered a cardiac arrest and, consequently, lives. In other words, close possible worlds in which the grandmother believes falsely that her grandson is alive (most close possible worlds) do not make the grandmother's actual belief unsafe because those are worlds in which she forms the target belief by testimony, i.e. by a method of *a different type* to her actual method of belief formation.

It is *prima facie* plausible to maintain that GRANDMOTHER and ATOMIC CLOCK are not structurally equivalent. Contra this intuition, I will argue that they have the same general structure by showing that the relevant beliefs of both agents, although epistemically risky, are not unsafe. More specifically, my reasoning will be as follows: 1) the agents come to know the target propositions via belief-forming methods of certain type in @. 2) In most close possible worlds, those propositions are false, i.e. the target actual beliefs are epistemically risky (as defined by HIGH EPISTEMIC RISK).[7] 3) However, this does not mean that they are unsafe, because close possible worlds in which the believed propositions are false are worlds in which the corresponding beliefs are formed via methods of a different type to the methods that the agents use in @.

In general, my strategy to diminish the force of Bogardus's counterexample is roughly that: to show that ATOMIC CLOCK is structurally equivalent to several cases of epistemic risk in which the target beliefs are known *and* safe. For that, we need a principled way of individuating methods of belief formation that allows us to give a principled explanation of why a method used by an agent in @ is of a different type to the method that she uses in close possible worlds in which she believes the same proposition falsely. With such a principle, we could give a uniform explanation of all cases structurally equivalent to ATOMIC CLOCK and, thus, we would be able to justify ~(2) in a non-question-begging way. It is important to emphasize that the treatment of ATOMIC CLOCK must be the

---

7 In ATOMIC CLOCK, the cause of epistemic risk is the isotope, which could easily stop the clock. In GRANDMOTHER, the source of epistemic risk is the bad condition of the grandson's heart, which could easily collapse.

same as of the rest of the cases (i.e. it must be principled); otherwise, our reply to Bogardus's counterexample would be *ad hoc.* Desirably, the principle of individuation should also offer the correct diagnosis of paradigmatic cases of unsafe belief. As we will see, an externalist principle of individuation will satisfy all these desiderata.

2 Relativization to Methods of Belief Formation

Mark Alfano (2009) discusses several ways of relativizing Nozick's modal conditions (sensitivity and adherence) to methods of belief-formation. Here, I will analyze several of the principles that Alfano discusses in order to evaluate whether they can be used to relativize safety to methods of belief formation. Let m_1 and m_2 be two belief-forming method tokens and M a belief-forming method type. Let us consider the following principle:

> (R1) m_1 and m_2 are of the same type M iff m_1 and m_2 are both cases of vision or olfaction or audition or taction or gustation or testimony or deduction or induction or memory.[8]

(R1) offers the correct diagnosis of GRANDMOTHER. In @, the old woman uses a belief-forming method based on vision (she looks at her grandson). In close possible worlds in which her grandson dies (most of them), she uses a belief-forming method based on testimony (others tell her that her grandson is alive). It thus follows from (R1) that the grandmother's actual belief-forming method is of a different type to the one she uses in most close possible worlds. Since, according to SAFE BELIEF, relevant close possible worlds are only those in which the agent forms the target belief via the same type of method that she uses in @, most close possible worlds are irrelevant to assess whether the grandmother's actual belief is safe. It is safe indeed, as in nearly all (if not all) close possible worlds where her grandson lives (not many of them) she *sees* that he is alive.

Note that (R1) offers an incorrect diagnosis (at least for our purposes) of ATOMIC CLOCK. In particular, it fails to explain why Smith's belief-forming method in @ is of a different type to the method that she uses in most close possible worlds. On the one hand, Smith looks at the working clock in @. Thus, the actual method token is a case of vision. On the other hand, in close possible worlds in which the isotope decays and stops the clock (most of them), Smith looks at the stopped clock and forms the false belief that it is 8:22 am. Therefore, the relevant method tokens in those possible worlds are cases of vision too. It

---

8 Adapted from Alfano (2009: 279). As Alfano explains, (R1) may need a few more disjuncts, but the idea should be intuitively clear. See Goldman (2009: 80-82) for relevant discussion on this way of individuating belief-forming methods.

thus follows from (R1) that all methods (in @ and in close possible worlds) are tokens of the same method type. Then, it follows from SAFE BELIEF that Smith's belief is unsafe, as in nearly all (if not all) close possible worlds she believes the same proposition by the same type of method and the proposition is false. Given the safety condition, her actual belief is not knowledge, which contradicts premise (1).

This result is certainly bad for our purposes of defending safety from Bogardus's attack. However, we cannot conclude (on pain of begging the question) that (R1) is unsatisfactory just because it fails to offer the diagnosis of the case that is convenient for our purposes. That is to say, we need to justify the tenability of principles of method individuation independently of Bogardus's case. To that aim, I will put aside, for the moment, ATOMIC CLOCK and Bogardus's contention that unsafe knowledge is possible and I will try to answer the following question: if one were in the position of the safety theorist, which is the best principle of method individuation that one could adopt, i.e. which principle would allow to account for more cases? To answer this question, we need to take into account a variety of cases and see whether the different principles of method individuation allow the safety theorist to explain them in keeping with our intuitions about whether they are cases of knowledge or not.

To begin with, as Alfano explains, (R1) fails to provide (as far as the purposes of the safety theorist are concerned) the correct diagnosis of the following case:

REDWOOD

Suppose Scotty sees and correctly identifies a redwood as a tree. Suppose further that there are plants nearby that Scotty would mistake for bonsai trees if he were to see them and that Scotty would have seen them if he did not see the redwood. Nevertheless, it seems, Scotty knows that he sees a tree (Alfano 2009: 276).[9]

If we relativized the safety condition to methods of belief formation individuated according to (R1), we would have to claim that Scotty's belief that the object in

---

9 The case was originally introduced by Alvin Goldman (1976: 779). Epistemologists have paid more attention to another of Goldman's cases, which is structurally equivalent to REDWOOD. In that case, an agent comes to know that the object in front of him is a dog when he looks at a dachshund despite the fact that he could easily have believed the same proposition when looking at a wolf (hence falsely). Goldman intends to show that the agent's inability to discriminate wolfs from dachshunds does not prevent him from successfully identifying dachshunds as dogs: "[The agent's] true belief fails to be knowledge if there is an alternative situation in which a non-dog produces the same belief by means of the same, or a very similar, appearance. But the wolf situation is not such an alternative: although it would produce in him the same belief, it would not be by means of the same (or a similar) appearance" (Goldman 1976: 779). By contrast, a situation with dachshund replicas would be such an alternative: a non-dog would produce the same belief by means of the same, or very similar, appearance.

front of him is a tree (when he looks at the huge redwood) is unsafe. Let us see why. Presumably, in many close possible worlds Scotty looks at the nearby plants and comes to believe that they are trees (although they are not). Looking at nearby plants is a case of vision. So is looking at the redwood. Thus, (R1) considers all the relevant method tokens of the same type (cases of vision). Then, given SAFE BELIEF, Scotty's belief that there is a tree (when he looks at the redwood) is unsafe and hence, given the safety condition, not knowledge.

We can thus conclude that (R1) is not a good option for the safety theorist. Let us consider another proposal that can be traced back to Goldman (1976: 779-780) and to Nozick (1981: 184):

(R2) m_1 and m_2 are of the same type M iff m_1 is experientially the same "from the inside" as m_2.

(R2) offers the correct diagnosis of GRANDMOTHER. In particular, it explains why the grandmother's belief-forming method in @ is of a different type to the one she uses in most close possible worlds. In @, the old woman uses a visual method. In close possible worlds in which her grandson dies (most of them), she is told that her grandson is alive. Testimony and vision are not experientially the same from the inside. Therefore, (R2) considers the grandmother's belief-forming method in @ of a different type to the method that she uses in close possible worlds where she forms a false belief. Then, given SAFE BELIEF, her actual known belief is safe.

In addition, (R2) also offers the correct diagnosis of REDWOOD. In particular, it explains why Scotty's actual belief-forming method is of a different type to the methods that he uses in most close possible worlds. In @, Scotty has a redwood-like experience. In close possible worlds where he looks at the nearby plants (presumably, most of them), he has bonsai-like experiences. Consequently, it follows from (R2) that his actual method is of a different type to the methods that he uses in close possible worlds where he looks at the nearby plants. In this way, given SAFE BELIEF, his actual known belief counts as safe, because close possible worlds where he uses an experientially equivalent method to form the belief that the object in front of him is a tree are worlds in which the object in front of him is a tree, namely a redwood.

Note that (R2) does not offer the correct diagnosis (at least for our purposes) of ATOMIC CLOCK. In particular, it does not explain why Smith's belief-forming method in @, where she looks at the working clock, is of a different type to the methods that she uses in close possible worlds in which she looks at the stopped clock (most close possible worlds). To see this, just note that in @ and in all close possible worlds Smith has equivalent clock-like experiences.

Be that as it may, the safety theorist should reject (R2) for the following reasons. Consider the following case:

MATRIX

Robots dominate the Earth. They hunt humans to use them as sources of bioelectrical and thermal energy. For that task, they use a special pistol that shoots a microchip to the human's head in such a way that it automatically connects her to a simulated reality called the Matrix. Humans do not even notice the change. Unbeknownst to Neo, he has just entered an area full of robots. Robots start shooting microchips to his head but for very improbable reasons they miss. In the meantime, Neo has picked up a stone and has formed the belief that he is holding a stone in his hand. If he had been hit, he would have dropped the stone and he would still have believed (in the Matrix) that he is holding a stone in his hand.

(R2) offers an incorrect diagnosis of MATRIX. Intuitively, Neo knows that he is holding a stone in his hand, even if in most close possible worlds he drops the stone and still believes the same proposition (Neo's actual belief is epistemically risky in the sense specified by HIGH EPISTEMIC RISK, but that does not prevent it from being knowledge). Of course, we do not want SAFE BELIEF to imply that Neo's actual belief is unsafe. And yet, this is the outcome if methods of belief formation are individuated as per (R2). For in @ and in all close possible worlds Neo has the experience that he is holding a stone and (R2) consequently implies that all belief-forming methods are of the same type.

Therefore, the safety theorist has good reason to reject (R2). Let us consider now an externalist proposal:

(R3) m_1 and m_2 are of the same type M iff the relevant external factors involved in the target belief's etiology when formed via m_1 are the same as when formed via m_2.[10]

Let me explain how to understand (R3). I follow John Greco (2005: 266) in thinking that a belief's etiology "concerns such things as the history of the belief and the reasons why it is held". Intuitively, relevant external factors involved in a belief's etiology may be perceived objects such as chairs, tables or hands. To get a

---

10 Goldman (2009: 81) discusses a principle of individuation like (R3): "[A] possible construal of 'bases' would include specific external objects involved in the method of belief acquisition. The basis of belief in the dachshund case might be *seeing the dachshund*, and the basis of belief in the wolf case might be *seeing the wolf*" (instead of talking about methods of belief formation, Goldman talks, more generally, about bases of belief). In addition, Timothy Williamson (2009: 307) seems to assume a principle of individuation like (R3) but, at some point, he also seems to hold something along the lines of the proposal of this paper, (R4), see e.g. Williamson (2009: 325 fn.13).

more accurate idea of how (R3) classifies types of belief-forming methods, the best thing we can do is to consider some examples.[11]

According to (R3), 1-3 are (all other things being equal) pairs of different types of belief-forming methods: 1) believing that there is a dachshund by seeing a dachshund vs. believing the same proposition by seeing a wolf; 2) believing that one is drinking pure, unadulterated water by drinking pure, unadulterated water from a glass vs. believing that one is drinking pure, unadulterated water by drinking water from a glass that has been doctored with undetectable toxins by conniving agents; 3) believing that one was shown n number of flashes after drinking regular orange juice vs. believing that one was shown n number of flashes after drinking a glass of orange juice with a tasteless mind-altering drug. These examples should suffice to get a general impression of how (R3) works: in each instance of belief formation there is some external factor in the history of the belief that differs from the corresponding instance.

Let us analyze how (R3) judges the cases scrutinized so far. In REDWOOD, MATRIX and GRANDMOTHER the external factors involved in the etiologies of the actual beliefs of Scotty, Neo and the grandmother are different from the external factors involved in the etiologies of the beliefs that they would form in most close possible worlds. Compare the relevant pairs of belief-forming methods: 1) believing that there is a tree by seeing a huge redwood (@) vs. believing that there is a tree by seeing a shrub (most close possible worlds); 2) believing that one is holding a stone by seeing a stone in one's hand (@) vs. believing that one is holding a stone by being connected to a simulated reality; 3) believing that one's grandson is alive by seeing him alive (@) vs. believing it through the testimony of someone who has seen him dead (most close possible worlds). (R3), therefore, rules that the methods of belief formation used by Scotty, Neo and the grandmother in most close possible worlds are different from the methods that they use in @. Their actual true beliefs can in this way count as safe.

Note, in addition, that (R3) does not seem to offer the correct diagnosis (at least for our purposes) of ATOMIC CLOCK. In @, Smith looks at a clock whose mechanism is in motion and forms the true belief that it is 8:22 am. In close possible worlds where Smith forms the false belief that it is 8:22 am (most of them), she looks at the same clock, which is stopped, however. Since the most relevant factor of the etiologies of the relevant beliefs in all worlds is the clock and its reading "8:22", there does not seem to be much reason to think that (R3)

---

11 The following examples are discussed by Williamson (2009: 307) and are listed by Dani Rabinowitz (2011), a list that I copy (almost literally) in the next paragraph.

should consider Smith's actual method of a different type to the methods that she uses in most close possible worlds.

At any rate, the safety theorist has other reasons to reject (R3). In particular, it fails to offer the correct diagnosis of the following well-known case:

FAKE BARNS

Henry forms the true belief that the object in front of him is a barn. Although the object is a genuine barn, Henry does not know it because the environment is populated with indistinguishable barn replicas that would easily have led him to form false beliefs in the same proposition.[12]

In @, Henry looks at a genuine barn and forms the belief that the object in front of him is a barn. In close possible worlds where he believes the same proposition falsely (most of them), he looks at barn replicas. Since the etiologies of the relevant beliefs in most close possible worlds involve objects (barn facsimiles) that are different to the most relevant object of the etiology of the target belief in @ (a genuine barn), Henry's actual belief-forming method is, according to (R3), of a different type to the methods that he uses in most close possible worlds. However, this result is disastrous for the safety-theorist: if she cannot fix that in close possible worlds Henry acquires his beliefs via the same type of belief-forming method, she is compelled to regard Henry's actual belief as safe, which goes against the standard diagnosis of the case. Therefore, (R3) is useless for the safety theorist.

Let us take stock. The cases discussed thus far show that the safety theorist is in need of a principle of individuation that can guarantee, first, that in FAKE BARNS (a case where the agent lacks knowledge) the actual belief-forming method is of *the same type* as the methods that the agent uses in close possible worlds where the relevant beliefs are false (most close possible worlds). In that way, the agent's actual belief can count as *unsafe*. Second, the principle of individuation must guarantee that in GRANDMOTHER, REDWOOD and MATRIX (all of them cases of knowledge) the target belief-forming methods in @ are of *a different type* to the methods that the agents use in close possible worlds where their beliefs are false (most of them). In this way, their actual beliefs can count as *safe* and thus qualify for knowledge.

3 An Externalist Principle of Individuation

In what follows, I will develop an externalist principle of method individuation that will offer a solid and principled justification for ~(2). To give shape to the

_____

12 The case is by Carl Ginet and appears in Goldman (1976).

proposal, we need to introduce a common distinction in epistemology (due to McGinn 1984) between two notions of reliability: *local* and *global* reliability. According to McGinn, a locally reliable belief-forming method is a method that is reliable with respect only to the proposition believed, while a globally reliable belief-forming method is a method that is reliable with respect to a range or a field of propositions.

In my view, the distinction between local and global reliability should be drawn not only in terms of the range of propositions about which a belief-forming method is reliable, but also in terms of the range of circumstances about which it is regarded as reliable. After all, when we estimate the degree of reliability of a belief-forming method we not only locate the propositions but also the circumstances relative to which the ratio of true beliefs is highest.

Accordingly, the set of circumstances with respect to which a belief-forming method is *locally* reliable includes only those in which the target proposition is believed. By contrast, a *globally* reliable belief-forming method is a method that is reliable with respect to a field of propositions and a range of circumstances. Circumstances here are understood as sets of external conditions of multiple environments (e.g. good light conditions) plus the internal conditions of the agent in those environments (e.g. being in good shape, sober, and so on).[13] Fields of propositions may be, for example, propositions about certain color or propositions of the type 'n is an even number', where 'n' is replaced by any number.

The fact that two belief-forming method tokens are *globally* reliable to the same degree implies that both of them tend to produce similar cognitive performances concerning the same kind of propositions and circumstances. This is a good indication, but not a sufficient reason (as we will see later), to think that they are the same type of method. For example, when two visual-based methods of belief formation are of the same type their degree of global reliability will be highest and, plausibly, the same with respect to propositions about, say, the color red (rather than propositions about tones of voice) and with respect to circumstances in which the light conditions are good (rather than, say, circumstances without ambient noise). We can accordingly formulate the following necessary condition, which will be part of our principle of individuation:

---

13 See Sosa (2010: 465-7) for relevant discussion on this particular way of characterizing circumstances.

(i) m_1 and m_2 are of the same type M, only if they are *globally* reliable to the same degree with respect to the same field of propositions and the same range of circumstances.

The notion of knowing via a method was introduced by Nozick (1981) to protect his analysis of knowledge from counterexamples similar to GRANDMOTHER. One might think that there is such a thing as a commonsense notion of method of belief formation. This notion is plausibly captured by a principle like (R1), which individuates methods in terms of the modality employed in forming the belief. If we do not include in our principle of individuation some condition along these lines, we may have to deal with the objection that individuating methods *just* in terms of their global reliability does not take into consideration our ordinary intuitions about belief-forming methods. It would be desirable, therefore, to include such a condition in our principle of individuation.

However, we have seen that (R1) individuates methods in an incorrect way for the safety theorist, as REDWOOD shows. But this does not mean that there is nothing we can preserve from (R1). Note that (R1) says two things: (1) that if m_1 and m_2 are both cases of vision or olfaction or audition, and so on, they are the same type of belief-forming method and (2) that if m_1 and m_2 are the same type of method, then they are both cases of vision or olfaction or audition, and so on. Cases like REDWOOD prove (1) wrong for the purpose of formulating a successful safety condition, but no case so far considered shows that (2), the necessity claim, is not suitable for that aim. Accordingly, we can include a requirement similar to (2) in our principle of individuation so as to take into account the modality employed in the formation of the belief. More specifically:

(ii) m_1 and m_2 are of the same type M, only if they are both based on vision or olfaction or audition or taction or gustation or testimony or deduction or induction or memory.

However, the conjunction of conditions (i) and (ii) may not be sufficient to individuate methods of belief formation in a correct manner. Consider REDWOOD. In @, Scotty uses a method token (m_1) that allows him to identify the redwood as a tree. It is a stipulation of the case that m_1 is globally reliable, and also that Scotty's actual circumstances (where there is a redwood within his field of vision) belong to the set of circumstances with respect to which m_1 is globally reliable. In the counterfactual scenario, Scotty looks at the nearby plants (e.g. shrubs) and thinks that they are bonsai trees using method token m_2. As the case is described, nothing prevents us from assuming that m_2 is globally reliable in the same way as m_1, i.e. to the same degree and with respect to the same propositions and circumstances. Accordingly, if the conjunction of (i) and (ii) is all it takes to

individuate methods, why should we think that m_1 and m_2 are not the same type of method? After all, they are both based on vision and globally reliable in the same way.[14]

At this point, it is clear that we need to supplement conditions (i) and (ii) with a further requirement. Let me motivate a condition that will turn our principle of individuation into a distinctly externalist proposal. Quite often, the expressions 'belief-forming method' and 'belief-forming process' are interchangeably used in the epistemological literature. Nevertheless, there are subtle differences between the notions to which these expressions refer. By 'belief-forming process' is usually meant, as Goldman (1986: 297) puts it, "a *basic* psychological process, something inherent in the fundamental architecture of the human cognitive system". A basic psychological process may certainly be part of a belief-forming method, but the notion of method is broader, as I will argue next. Crucially, methods of belief formation are, as Nozick (1981) originally introduced them, *ways of believing*.

In general, when we individuate ways of φ-ing, we not only take into account the abilities needed to φ, but also conditions of the circumstances that are relevant to the deployment of such abilities. For example, a way of climbing the Mount Everest is to make such-and-such moves without supplemental oxygen via the southeast ridge under such-and-such climate conditions. Ways of believing are no exception. In the same way as expert mountaineers classify ways of climbing in terms of the use (or not) of supplemental oxygen, the route followed or the season of the year, we are allowed to appeal to conditions of the circumstances that are relevant to belief formation in order to individuate ways of believing. In particular, relevant conditions are those with respect to which a way or method (whether based on one or another modality) is (to a certain degree) globally reliable.

Accordingly, the third condition of our principle of individuation can be stated as follows (let m_1 be the target belief-forming method):

> (iii) m_1 and m_2 are of the same type M, only if the circumstances in which the target belief is formed via m_2 are in the set of circumstances with respect to which m_1 is globally reliable.

Conditions (i), (ii) and (iii) make up a plausible externalist principle of individuation:

> (R4) m_1 and m_2 are of the same type M iff (i) m_1 and m_2 are *globally* reliable to the same degree with respect to the same field of propositions and the same range of circumstances; (ii) they are both based on vision or

---

14 I am grateful to an anonymous reviewer for pressing this point.

olfaction or audition or taction or gustation or testimony or deduction or induction or memory; (iii) the circumstances in which the target belief is formed via m_2 are in the set of circumstances with respect to which m_1 is globally reliable.

Let us analyze REDWOOD again. The method that Scotty uses in @, m_1, is some vision-based globally reliable method that enables him to identify redwoods as trees. Scotty's actual circumstances belong to the set of circumstances with respect to which m_1 is globally reliable. In this big set of circumstances we can find circumstances where the light conditions are good, where there are (within Scotty's field of vision) objects with trunks, branches and leaves with certain characteristics such as having such-and-such shape or being at a considerable distance from the ground, as well as many other relevant features. In the counterfactual scenario where Scotty uses method token m_2, there are (within his field of vision) objects with trunks, branches and leaves at a little distance from the ground, and so on.

The distance of the leaves from the ground, the width of the trunk as well as many other relevant factors of the circumstances of m_2 differ from the factors that shape the kind of circumstances with respect to which m_1 is globally reliable. Thus, condition (iii) is not met, which means that m_1 and m_2 are not the same type of method. Consequently, close possible worlds in which Scotty believes that there is a tree when he looks at, say, a shrub are not relevant to evaluate whether his actual true belief is safe. Since in close possible worlds where he looks at the redwood (presumably, few close possible worlds), he correctly identifies it as a tree, his actual true belief is safe, according to SAFE BELIEF.

Let us consider now GRANDMOTHER. In this case, condition (ii) is not met. In @, the old woman uses a vision-based belief-forming method. By contrast, in close possible worlds in which her grandson dies (most of them), she uses a testimony-based method. Thus, her actual method is of a different type to the method that she uses in most close possible worlds, which means that close possible worlds in which she believes that his grandson is alive falsely are not relevant to evaluate whether her actual true belief is safe. Since in close possible worlds where her grandson is still alive (few close possible worlds), she comes to believe that he is alive by looking at him, her actual true belief is safe, according to SAFE BELIEF.

In MATRIX, condition (ii) is not met either. In @, Neo forms the true belief that he is holding a stone in his hand by looking at his hand. Thus, he uses a vision-based belief-forming method. In close possible worlds where he forms the belief that he is holding a stone in his hand and the belief is false (most of them),

he is connected to the Matrix and a supercomputer feeds him false information. Whatever this method is, it is not based on vision, and, therefore, it is of a different type to Neo's actual method.

One could modify the case in such a way that the supercomputer fed Neo false information via some vision-based method, so that condition (ii) would be met. In that scenario, however, it is dubious that condition (iii) would hold. For circumstances in which Neo comes to believe that he is holding a stone being plugged in to the Matrix are not the type of circumstances with respect to which his actual vision-based method is globally reliable. After all, when we estimate the *global* reliability of an ordinary visual method (like Neo's actual method), we do not take into consideration circumstances in which one is plugged in to a supercomputer. Therefore, close possible worlds in which Neo is connected to the Matrix (most of them) are not relevant to assess whether his actual true belief is safe. Since in close possible worlds where he is not in the Matrix (few of them), he looks at his hand and forms the true belief that he is holding a stone, his actual belief is safe, according to SAFE BELIEF.

Consider FAKE BARNS. The standard diagnosis of the case says that Henry does *not* know that the object in front of him is a barn, although he is looking at a genuine one. (R4) explains why Henry's actual method is of the *same* type as the method of belief formation that he uses in close possible worlds where he looks at barn replicas. To see this, I will consider each condition separately. Let m_1 be Henry's actual method and m_2 the method that he uses in close possible worlds in which he believes that there is a barn in front of him when looking at a barn replica.

Condition (ii) obviously holds, since both m_1 and m_2 are vision-based methods. Condition (i) is also met. The fact that neither m_1 nor m_2 are *locally* reliable (reliable relative to the specific circumstances where they are used) does not prevent them from being *globally* reliable to the same degree and with respect to the same propositions and circumstances. In particular, m_1 and m_2 are globally reliable methods for discriminating barns from other buildings such as houses, garages, airport hangars or skyscrapers. As the case is described, there is no reason to think that Henry does not have that ability and that he does not retain it when he fails to know in the Fake Barn Country that the object in front of him is a barn, viz. when his ability is not *locally* reliable.

On the other hand, circumstances where there is an object with the characteristics of a barn within his field of vision, where the light conditions are such-and-such, where Henry is at distance n, and so on, belong to the set of circumstances with respect to which the ability is globally reliable. His actual

circumstances, where he uses m_1, are certainly in that set (he is in front of a prototypical barn, the light and distance conditions are good), and so is the counterfactual scenario where he uses m_2: the light and distance conditions are equally good and he is in front of a faithful barn replica, i.e. in front of an object with all the features that typically allow him to differentiate barns from other buildings reliably. Consequently, condition (iii) holds.

Since all conditions of (R4) are met, it follows that m_1 and m_2 are the same type of belief-forming method. In addition, since in most close possible worlds where Henry believes that the object in front of him is a barn via the same type of method his belief is false, his actual belief is unsafe, according to SAFE BELIEF, and not knowledge, given the safety condition, which agrees with the standard diagnosis of the case.

Let us consider now the alleged counterexample to safety, ATOMIC CLOCK. (R4) explains why Smith's actual belief-forming method is of a different type to the method that she uses in most close possible worlds. To see this, I will analyze each condition separately. Let us start with condition (ii). In @, Smith looks at the working clock and forms the true belief that it is 8:22 am. In most close possible worlds, she looks at the stopped clock and believes the same proposition falsely. Both belief-forming method tokens are vision-based, therefore condition (ii) holds. Does condition (i) hold? That is, are method token m_1 (the method of forming beliefs by looking at the working clock) and method token m_2 (the method of forming beliefs by looking at the stopped clock) globally reliable to the same degree with respect to the same field of propositions and the same range of circumstances?

Intuitively, the answer is negative. Let us see this in more detail. Plausibly, m_1 and m_2 are relativized to the same field of propositions, namely propositions about the time (rather than, say, propositions about colors). Are m_1 and m_2 globally reliable to the same degree with respect to the proposition that it is 8:22 am? If Smith used m_1 in a range of different situations to form the belief that it is 8:22 am, her belief would be true in most of them. Obviously, the same cannot be said about m_2. Therefore, condition (i) does not hold.

Note, in addition, that m_1 has such a high degree of reliability only relative to certain set of circumstances. The kind of circumstances contained in this set are, for example, circumstances in which the light conditions are good, in which the screen of the clock is not covered by an opaque layer of dust, in which the clock is not under strong magnetic influence or under water, in which the voltage level of the battery is not low (with low batteries clocks slow down) and, more importantly, circumstances in which a radioactive isotope has not decayed in the

vicinity causing the clock to stop. In the circumstances in which Smith uses m_2, a radioactive isotope has decayed disrupting the clock's sensor and stopping the clock. Therefore, those circumstances are not in the set of circumstances with respect to which m_1, Smith's actual method, is globally reliable. Therefore, condition (iii) does not hold.

Note that it is plausible to maintain that m_1 and m_2 are not the same type of method because we have conceived methods of belief formation as *ways of believing*. Thus, in the same manner as climbing the Mount Everest by using a functioning oxygen bottle is a different *way of climbing* to climbing the mountain with a broken oxygen bottle, coming to believe that it is 8:22 am by looking at a perfectly working clock is a different *way of believing* to coming to believe that it is 8:22 am by looking at a stopped clock (no matter that the clock or the oxygen bottle are the same objects when they work and when they do not).[15]

In sum, m_1 and m_2 are different types of methods of belief formation, according to (R4). In this way, close possible worlds in which Smith forms the belief that it is 8:22 am by looking at the stopped clock are not relevant to assess whether her actual true belief is safe. Since in close possible worlds in which she forms the same belief via the same type of method (few close possible worlds), the believed proposition is true, Smith's actual true belief is safe, according to SAFE BELIEF. Therefore, all ATOMIC CLOCK shows is that safety needs to be relativized to methods of belief formation. We can thus conclude that Bogardus has not made a case for the claim that unsafe knowledge is possible.

---

15 One might be concerned about the apparent structural similarity between ATOMIC CLOCK and FAKE BARNS. In FAKE BARNS, the prototypical features that allow Henry to identify an object as a barn are shared both by genuine and fake barns. This is partly the reason circumstances with barn replicas belong to the set of circumstances with respect to which Henry's actual visual method is globally reliable (to belong to that set, the light conditions of the circumstances must be good as well, the distance must be appropriate, and so on). In ATOMIC CLOCK, one might argue, the prototypical features that allow Smith to read clocks are shared by the working and by the stopped clock (both read 8:22 am). Should then circumstances in which the clock is stopped be part of the set of circumstances with respect to which Smith's actual method is globally reliable? The answer is negative. Smith's actual method is truth-conducive because the clock is a *reliable indicator* of the time (according to Bogardus, it is the world's most accurate clock). Consequently, circumstances in which the clock is stopped are not the kind of circumstances with respect to which Smith's actual method is globally truth-conducive (or reliable).

4 Unsafe, Safe and Super-Safe Belief

I would like to conclude with some general remarks on the notions of safety and epistemic risk with the hope that they help to dissipate a common misunderstanding about the kind of cases that are considered potential counterexamples to the necessity of the epistemic condition. Consider the following case of knowledge:

> HANDS
>
> George is in his room. There are good light conditions. Nothing is epistemically amiss. He looks at his hands and forms the belief that he has hands.

Intuitively, George knows that he has hands. On the one hand, his belief is safe (as defined by SAFE BELIEF), because in nearly all (if not all) close possible worlds in which he forms the same belief via the same type of method that he uses in @, the belief is true. On the other hand, George's belief is *not* epistemically risky (as defined by EPISTEMIC RISK), because in most close possible worlds in which George believes that he has hands, he has hands. Call a belief *super-safe* just in case it is (1) safe and (2) not epistemically risky (in the specified senses of safety and epistemic risk).

The target beliefs in GRANDMOTHER, REDWOOD, MATRIX and ATOMIC CLOCK all amount to knowledge. As we have seen, they are safe because in nearly all (if not all) close possible worlds in which the agents come to believe the same propositions via the same type of methods that they use in @, the propositions are true. However, although safe, the beliefs in question are not super-safe, because they are epistemically risky: in most close possible worlds in which the agents come to believe the same propositions (no matter by which *type* of belief-forming method), the propositions are false.

The following table compares the different modal profiles of unsafe, safe and super-safe beliefs ('@' is the actual world and 'W_n' close possible worlds):

	@	W_1	W_2	W_3	W_4
UNSAFE BELIEF	p	~p	~p	~p	~p
	S uses M	**S uses M**	**S uses M**	**S uses M**	**S uses M**
	Bp	Bp	Bp	Bp	Bp
SAFE BELIEF	p	~p	~p	p	~p
	S uses M	~(S uses M)	~(S uses M)	**S uses M**	~(S uses M)
	Bp	Bp	Bp	Bp	Bp
SUPER-SAFE BELIEF	p	p	p	p	p
	S uses M	**S uses M**	**S uses M**	**S uses M**	**S uses M**

	Bp	Bp	Bp	Bp	Bp

I surmise that the reason some epistemologists have thought that the safety condition is not necessary for knowledge is that they have mistaken super-safety for safety, in such a way that they have proposed cases of knowledge with high levels of epistemic risk involved, thinking them to be obvious counterexamples to safety, when in reality they are only counterexamples to the view that super-safety is necessary for knowledge. However, while super-safety is very desirable from an epistemic point of view (agents with super-safe beliefs are in modally robust epistemic positions), only safety is considered a necessary condition for knowledge.

The tendency to mistake super-safety for safety is due to a misunderstanding about how to individuate methods of belief formation. In particular, it is thought that two belief-forming method tokens m_1 and m_2 are of the same type *if* they are both cases of vision or audition or deduction, and so on. However, I have shown that this way of individuating methods leads to a wrong diagnosis of certain cases (it is the other direction that is correct: m_1 and m_2 are of the same type *only if* they are instances of the same modality). The mistake lies in thinking that if m_1 and m_2 instantiate the same modality, they will always be the same type of method regardless of the circumstances in which they are used. However, I have shown that if m_1 is *globally* reliable to certain degree, m_2 is not the same type of method unless the circumstances in which m_2 is used are contained in the set of circumstances with respect to which m_1 is globally reliable.

There is something compelling about this thought. For example, Terry the taxi driver might be a very reliable way for you to go home, unless Terry is completely drunk, in which case Terry's taxi is a terribly unreliable way of getting home. The degrees of reliability here are so different that we judge that these are different ways of getting home, even though one is also tempted to judge that they are instances of the same type of method (Terry's taxi) used in different circumstances (driving sober vs. driving drunk).[16] Although in some conversational contexts it might be acceptable to claim that driving drunk is the same way of driving a car as driving sober, such a coarse-grained individuation of ways of φ-ing is not adequate in general. The fact that you take Terry's taxi (rather than, say, Larry's) is only a necessary condition for individuating your way of getting home.

The same applies to the epistemic case: it is a mistake to think that if m_1 and m_2 include the use of the same epistemic device (e.g. a clock or a

---

16 I am indebted to an anonymous reviewer for suggesting me this example.

thermometer), then they are the same type of belief-forming method regardless of the circumstances in which the device is used. In the same manner as shooting with a bow whose string is tight is not the same way of shooting as shooting with a bow whose string is completely loose, or in the same manner as playing a guitar that is correctly tuned is not the same way of playing as playing with an out of tune guitar, believing that p by means of a properly functioning epistemic device is not the same way of believing that p as coming to believe that p by means of a broken epistemic device.[17] Accordingly, in the same way as the success of a shot is safe even though the archer could easily have lost her method of shooting (e.g. even though her bow could easily have broken), the success of a belief counts as safe even though the agent could easily have lost her method of belief formation (e.g. even though the epistemic device used could easily have broken). In conclusion, safety is not in danger.

Acknowledgements

References

Alfano, M. (2009). Sensitivity theory and the individuation of belief-formation methods. Erkenntnis, 70: 271-281.

Bogardus, T. (2012). Knowledge under threat. Philosophy and Phenomenological Research, doi:10.1111/j.1933-1592.2011.00564.x

Coffman, E. J. (2009). Does luck exclude control? Australasian Journal of Philosophy, 87(3): 499–504.

Comesaña, J. (2005). Unsafe knowledge. Synthese, 146: 395-404.

Goldman, A. (1976). Discrimination and perceptual knowledge. Journal of Philosophy, 73 (November): 771-791.

Goldman, A. (1986). The cognitive and social sides of epistemology. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 2: 295-311.

---

17 The more specialized the context is (e.g. archery and musical competitions, philosophical contexts), the more salient and fine-grained these differences are.

Goldman, A. (2009). Williamson on knowledge and evidence. In Greenough, P. & Pritchard, D. (eds.), Williamson on knowledge. Oxford: Oxford University Press.

Greco, J. (2005). Justification is not internal. In Steup, M. & Sosa, E. (eds.) Contemporary debates in epistemology. Malden, MA: Blackwell.

Hiller, A. & Neta, R. (2007). Safety and epistemic luck. Synthese, 158: 303-313.

Kelp, C. (2009). Knowledge and safety. Journal of Philosophical Research, 34: 21-31.

Levy, N. (2011). Hard Luck: How Luck Undermines Free Will and Moral Responsibility. Oxford: Oxford University Press.

Neta, R. & Rohrbaugh, G. (2004). Luminosity and safety of knowledge. Pacific Philosophical Quarterly, 85: 396-406.

Nozick, R. (1981). Philosophical explanations. Cambridge: Harvard University Press.

Manley, D. (2007). Safety, content, apriority, self-knowledge. The Journal of Philosophy, 104: 403-23.

Mcginn, C. (1984). The Concept of Knowledge. Midwest Studies in Philosophy, 9 (1): 529-554.

Pritchard, D. (2005). Epistemic luck. Oxford: Oxford University Press.

Rabinowitz, D. (2011). The safety condition for knowledge. Internet Encyclopedia of Philosophy. http://www.iep.utm.edu/safety-c/#SSH3aiii, last updated: August 10, 2011.

Sosa, E. (2010). How competence matters in epistemology. Philosophical Perspectives, 24 (1): 465-475.

Williamson, T. (2009). Replies to critics. In Greenough, P. & Pritchard, D. (eds.), Williamson on knowledge. Oxford: Oxford University Press.