The Explanationist and the Modalist

Dario Mortini [iD]

University of Glasgow, UK
Email: Dario.Mortini@glasgow.ac.uk

Abstract
Recent epistemology has witnessed a substantial opposition between two competing approaches to capturing the notion of non-accidentality in the analysis of knowledge: the explanationist and the modalist. According to the latest advocates of the former (e.g., Bogardus and Perrin 2020), S knows that *p* if and only if S believes that *p* because *p* is true. According to champions of the latter (e.g., safety and sensitivity theorists), S knows that *p* if and only if S's belief that *p* is true in a relevant set of possible worlds. Because Bogardus and Perrin's explanationism promises to deliver a plausible analysis of knowledge without any need for modal notions, it is an elegant proposal with prima facie appeal. However, such version of explanationism ultimately does not live up to its promises: in this paper, I raise some objections to their explanationist analysis of knowledge while showing that modalism is in much better shape than they think. In particular, I argue that their explanationist condition generates the wrong results in Gettier cases and fake-barn cases. I also offer and defend a novel version of modalism: I introduce a refined safety condition which is shown to successfully handle the same Gettier cases that beset Bogardus and Perrin's version of explanationism. The paper concludes with a reassessment of the explanationist's initial ambition to provide an analysis of knowledge without modal notions. The upshot will be that even if the prospects for such analysis remain dim, our money should be on the modalist, not the explanationist.

Introduction

While the task of providing an analysis of knowledge was once crucial to epistemology, in retrospect it looks more like a hopeless series of misplaced efforts. Troubles began when Gettier's paper was taken to show that even if truth, belief and justification are individually necessary, they are nonetheless jointly insufficient conditions for knowledge. This naturally prompted the question at the very heart of the Gettier problem: in addition to these three, which *further* condition is ultimately sufficient for knowledge? Epistemologists embarked on a long journey searching for such missing condition, and considered different candidates along the way. Modal conditions stood out among the most promising but equally fell prey to Gettier-style counterexamples, and, after almost three decades of unsuccessful proposals, the hope in a full-blooded

analysis slowly but steadily wore off.[1] Craig (1990) urged to focus on the *function* of the concept of knowledge rather than on its analysis, Zagzebski (1994) showed that the Gettier problem is virtually *inescapable*, and Williamson (2000: 2–5) suggested, quite radically, that knowledge simply *cannot* be analysed. As a result, many epistemologists no longer hold out hope that there is – or there even *can* be – a satisfactory analysis of knowledge, and such pessimism appears to be well-motivated.

Many, though not *every* epistemologist: some don't share the pessimism and still retain hope. Recent times have seen the rise of novel *explanationist* analyses of knowledge.[2] The explanationist begins by taking issue with a dominant figure of post-Gettier epistemology, the *modalist*. With Sosa (1999), modalists ask how knowledge must be *modally* related to what is known. They give two answers: *sensitivity* (roughly, S knows that *p* if and only if in the closest possible world where *p* is false, S does not believe that *p*) or *safety* (roughly, S knows that *p* if and only if in most or all close possible worlds S's belief that *p* is true).[3] Modalists occasionally quibble over the details, but they all agree that a modal condition such as sensitivity or safety is the key to solve the Gettier problem.[4] And at precisely this point, the explanationist demurs: instead of asking how knowledge must be modally related to what is known, she asks how knowledge must be *explanatorily* related to what is known. Since explanation is a hyperintensional notion (Nolan 2014: 157), the key to solve the Gettier problem is *not* a modal condition. According to the explanationist, knowledge doesn't require the subject's belief to be true in an appropriate set of possible worlds. More simply, it only requires the subject to believe *p* because *p* is true.

In this paper, I object to the explanationist and I side with the modalist: I shall argue that the most recent explanationist condition offered by Bogardus and Perrin (2020) remains insufficient for knowledge due to, once again, Gettier-style cases. But the significance of my contribution goes beyond a mere discussion of the intuitively correct verdicts in Gettier-style vignettes, for what is at stake is more than just that. If successful, this recent explanationist proposal would deliver what many epistemologists have longed for – an analysis of knowledge – and it would do so *without* modal conditions. Crucially, the success of explanationism also casts serious doubt on the relevance of modal notions to epistemological theorising, and, a fortiori, on the tenability of modalism in general. Accordingly, modalists have urgent enough reason to halt the explanationist advance, and such is the task I take on here: I will show that the most recent version of explanationism not only fails to provide a satisfactory analysis, but it also incurs difficulties that modalism easily avoids. Ultimately, such version of explanationism does not live up to its promises: thus, epistemologists should not do away with modal conditions on knowledge – not yet at least.

My plan is as follows. In section 1, I lay out the main tenets of explanationism and focus on the putative analysis of knowledge offered by its most recent advocates (Bogardus and Perrin 2020). In section 2, I argue that the explanationist condition is not sufficient for knowledge due to standard Gettier cases and fake-barn cases. The

---

[1]For an overview of post-Gettier epistemology, see Ichikawa and Steup (2012). For a more historically informed perspective, see Antognazza (2015). For discussion of Gettier-style counterexamples raised specifically against the sufficiency of modal conditions on knowledge, see in particular Goldberg (2015: section 1) and Grundmann (2018: section 3).

[2]For instance, Faraci (2019), Bogardus and Perrin (2020), Korman and Locke (Forthcoming *a*). Less recent advocates include Goldman (1984), Rieber (1998), Neta (2002), Jenkins (2005).

[3]Nozick (1981), Sainsbury (1997), Sosa (1999), Williamson (2000), Pritchard (2005), Kelp (2013), Beddor and Pavese (2020), and Clarke-Doane and Baras (2021).

[4]Nozick (1981: 173) and Pritchard (2012: 1, 2014: 94) offer sensitivity and safety as anti-Gettier conditions.

latter are especially troubling in virtue of a problematic closure failure. In section 3, I turn the spotlight to modalism and I defend it in two stages: I introduce a refined safety condition and I show its superiority in handling the same Gettier cases that beset Bogardus and Perrin's version of explanationism. I close by reassessing the case for their initial ambition to provide an analysis of knowledge without modal notions. The upshot will be that even if the prospects for such analysis remain dim and modalism is not problem-free, our money should be on the modalist, not the explanationist.

1. Explanationism outlined

Explanationist analyses of knowledge take the following simple form: S knows that p if and only if S believes that p because p is true. The truth of p must *enter prominently* – or *figure crucially* – in the explanation of why the target belief is held. Instead of modal conditions and related possible world talk, the focus is primarily on the *hyperintensional* notion of explanation.[5] Explanationist approaches may vary in the details, but they all accept some qualified version of this simple analysis.

Variation in the details is due to how the key notion of explanation is unpacked. Alan Goldman (1984: 101) and Rieber (1998) both invoke the truth of p as 'the *best* explanation' of why S believes p. However, as also noted elsewhere (Jenkins 2005: 142–3), both these analyses leave the central notion of explanation too vaguely stated. Accordingly, I shall put them aside for the purpose of this paper. Jenkins (2005: 139) refines the explanationist condition: she appeals to the truth of p being a good explanation for an *outsider* – someone not acquainted with a more specific set of facts. Despite her explanationist sympathies, Jenkins is nonetheless keen to deny the ambitious status of reductive analysis to her proposal. More modestly, she only aims at "*getting a handle* on the concept of knowledge" and "understanding what its *role* in our lives might be" (Jenkins 2005: 138; emphasis mine). Since my focus is on full-blooded analyses of knowledge and Jenkins is careful to stress that she's *not* offering one, I will have to put also Jenkins's proposal to one side.

This leaves me with one main candidate left. In what follows, the action will be *exclusively* on the most recent and fully developed explanationist analysis offered by Bogardus and Perrin (2020). My reason to do so is simple: Bogardus and Perrin (2020: 1) are explicit in stating that their explanationist condition is both necessary and *sufficient* for knowledge. Hence, theirs is a full-blooded *reductive* analysis. Since I am chiefly interested in whether explanationism provides a satisfactory reductive analysis and they are indeed offering one, I shall narrow down my focus specifically to their proposal.[6]

Bogardus and Perrin make a two-fold case for explanationism. First, they offer a knockdown counterexample aimed at *every* modalist approach. Then, they proceed to provide an alternative explanationist analysis of knowledge. Before focussing on their explanationist analysis, let's look at the knockdown counterexample to modalism. Consider the following case:

---

[5]Hyperintensional notions distinguish between necessarily equivalent contents and block substitution *salva veritate*. It is widely agreed that explanation is hyperintensional: there's a truth-conditional difference between "I am in danger because Doctor Jekyll is in the room" and "I am in danger because Mr. Hyde is in the room", even though Doctor Jekyll and Mr. Hyde are intensionally equivalent terms. For a state-of-the-art overview of hyperintensionality, see Berto and Nolan (2021).

[6]A quick but important terminological clarification: unless specified otherwise, I will henceforth use 'explanationism' and 'the explanationist' to refer to Bogardus and Perrin's *specific* version of explanationism.

ATOMIC CLOCK

> The world's most accurate clock hangs in Mia's office. The clock's accuracy is due to a clever radiation sensor. However, this radiation sensor is very sensitive and could easily malfunction if a radioactive isotope were to decay in the vicinity. 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 Mia'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." 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. (Bogardus 2014: 12; Bogardus and Perrin 2020: 4)

According to the explanationist, Mia knows the time but her belief is neither safe nor sensitive. Mia's belief is unsafe: since the isotope could have easily decayed, in the very close worlds where it does the clock stops and Mia forms the false belief that it's 8:22. Mia's belief is also insensitive: in the closest possible world where the isotope decays and the clock stops, she forms the false belief that it's 8:22. Thus, had p been false, Mia would have believed it anyway. From this, the explanationist swiftly concludes that no modal condition is necessary for knowledge.[7] Such conclusion paves the way for a different explanationist analysis, to which I now turn the attention.

Recall the general structure of explanationist analyses: S knows that p if and only if S believes that p because p is true. The truth of p must figure crucially in the explanation of why the target belief is held, but what notion of explanation is relevant here? Bogardus and Perrin lean on Strevens's 'kairetic test' for difference-making in scientific explanations (Strevens 2008). The idea is roughly this: begin with a set of potential *explanantes* for why the target belief is held, and progressively subtract ('abstract away') each of them until the explanation fails. The only *explanans* left is the one that crucially figures in the *complete* explanation of why the target belief is held. If such *explanans* is the truth of p, then the belief under consideration is in the market for knowledge.

We can see this explanationist analysis in action by focussing on ordinary cases of perceptual knowledge (this will be important also for the related explanationist treatment of fake-barn cases; more on this below). Suppose I truly believe that there's a cup of coffee in front of me. There is a set of different explanations for my belief: it might be because it appears to me that there's a cup of coffee, because of particularly favourable lighting conditions, because of my properly functioning visual system, because I am particularly craving coffee at that time of the day, and so on. If we start to run the kairetic test, we quickly notice how none of these explanations really makes a difference: we can subtract each of these candidate *explanantes*, and yet, according to Bogardus and Perrin, the explanation wouldn't be complete. For example, if the explanation offered for why I (truly!) believe that there's a cup of coffee in front of me is that it only appears that there is one, something would be missing because in fact

---

[7]In the explanationist's own words: "Such a case nicely pries apart our concept of knowledge from our concepts of safety and sensitivity, by polluting very many of the 'nearby' worlds with false beliefs while maintaining a tight enough connection between belief and its truth to allow for knowledge" (Bogardus and Perrin 2020: 4). For more counterexamples against the necessity of safety for knowledge, see Neta and Rohrbaugh (2004), Comesaña (2005) and Kelp (2009). I return to this counterexample in section 3, and I show it to be toothless against my version of modalism.

there truly *is* a cup of coffee in front of me. To reach a complete and satisfactory explanation, we must appeal to the truth of *p*: I believe that there's a cup of coffee in front of me because it's true that there's a cup of coffee in front of me. The truth of *p* crucially figures in the explanation of why the target belief is held: hence, the belief under consideration constitutes knowledge. According to the explanationist, the same procedure generalises beyond perceptual knowledge and it is meant to apply, *mutatis mutandis*, to further cases of interest to epistemologists.

With these points in play, it's now worth pausing to appreciate the ambitions of Bogardus and Perrin's proposal. Equipped *only* with the notion of explanation and Strevens' kairetic test, they set out to cover *all* the relevant instances of knowledge: perceptual, inductive, deductive, moral, and mathematical (Bogardus and Perrin 2020: section 3). Moreover, given cases like ATOMIC CLOCK, they also reject the key modalist assumption, namely any relevance of modal conditions on knowledge. They are quite explicit on this; for sake of vividness, consider the following passage:

> We believe modalism misses something crucial about the nature of knowledge: *the connection between a believer and the truth can't be fully captured in modal terms, because it's an explanatory connection.* Modalism, then, can have no long-term success as a research project. It's standing in its own grave. So, it's time to look elsewhere for the analysis of knowledge. (Bogardus and Perrin 2020: 7, emphasis mine)

This is the main ambition of Bogardus and Perrin's explanationism: to provide a *successful* reductive analysis of knowledge *without* modal notions. Crucially, for their analysis to be successful, the explanationist condition must suffice for knowledge and thus be immune from Gettier-style counterexamples. Yet, as I argue in the next sections, the explanationist analysis has no such immunity and falls prey to Gettier-style cases. Modalism, on the contrary, does not stand in its own grave: it is not undermined by counterexamples such as ATOMIC CLOCK and it easily accommodates the same Gettier cases that will be shown to beset explanationism. Defending the superiority of modalism will occupy section 3; now I want to focus on two familiar vignettes featuring stopped clocks and fake barns.

2.1. Explanationism exposed: stopped clocks

Consider the following run-of-the-mill Gettier case:

STOPPED CLOCK

> Russell takes a competent reading from a clock that he knows to be reliable and has no reason to think is currently not working. Based on this reading, Russell forms the belief that it is 8:22 pm. What's more, the belief is true: it is indeed 8:22 pm. There is, however, a twist to the story: the clock is broken and the reason Russell's belief is true is that the clock happened to stop working exactly 24 hours ago.[8]

Intuitively, Russell has a justified true belief that it's 8:22, but he doesn't know it. Can the explanationist accommodate this verdict? In this initial case, the answer is 'yes'.

---

[8]Following Pritchard (2012), this is a *standard* Gettier case featuring *intervening* luck: the agent's belief is initially false, but it then hits upon the truth courtesy of a stroke of luck. Differently, *non-standard* fake-barn cases display *environmental* luck: the agent's belief is true, but it could have easily been false given the environment where it was formed.

Russell believes that it's 8:22 pm because the clock reads 8:22 pm. However, the clock reads 8:22 not because it *is* 8:22. Rather, it reads 8:22 because it *was* 8:22 exactly 24 hours before (Bogardus and Perrin 2020: 17). Russell doesn't believe that *p* because *p* is true: the bi-conditional of the explanationist analysis is not satisfied, and, accordingly, the explanationist condition is not met. Since the truth of *p* doesn't figure crucially in the explanation of why the target belief is held, the target belief does not to constitute knowledge. So far, so good: explanationism can capture absence of knowledge in this familiar Gettier-style case and it secures the right result.

But now consider the following slightly different version of STOPPED CLOCK:

DEFECTIVE STOPPED CLOCK

Russell takes a competent reading from a clock that he knows to be reliable and has no reason to think is currently not working. Based on this reading, he forms the belief that it's 8:22 pm. What is more, it is 8:22 pm and the clock correctly reads 8:22 pm. There is, however, a twist to the story: *in virtue of an undetected manufacturing defect, the clock is designed to stop at exactly 8:22 pm, which is also when Russell happens to look at it*. It's 8:22 pm, the clock stops at 8:22 pm, and Russell truly believes that it's 8:22 pm.

Here too, Russell may have a justified true belief that it's 8:22 pm, but he doesn't know it. The reason why Russell fails to know is the uncontroversial assumption that one can't know the time from a stopped clock *regardless* of when exactly the clock happens to stop. Can the explanationist secure the right result also in this case? Unfortunately, now the answer is 'no'. Russell believes that it's 8.22 pm because the clock reads 8:22 pm. Unlike the previous version of this Gettier case, the clock reads 8:22 pm because it is 8:22 pm. Nevertheless, the clock *also* stops because it is 8:22 pm, and one hardly comes to know the time by consulting a stopped clock. Russell believes that *p* because *p* is true: the truth of *p* crucially figures in the explanation of why the target belief is held. However, the same fact also explains why the clock stops: given the undetected manufacturing defect, the fact that it's actually 8:22 pm is also the reason why the clock stops. Russell believes that *p* because *p* is true, but, *pace* the explanationist, he intuitively fails to know.

The explanationist analysis fails to pass the sufficiency test with more refined versions of standard Gettier cases. The explanationist condition is met: Russell believes that it's 8:22 pm because the clock reads 8:22 pm, the clock reads 8:22 pm because it is 8:22 pm and it's true that it is 8:22 pm. Yet, knowledge is intuitively absent. After all, you can't come to know the time from a stopped clock.[9] Even granting that the explanationist condition is necessary, DEFECTIVE STOPPED CLOCK shows that it remains

---

[9]Or can you? Explanationists may bite the bullet and hold that the agent knows in DEFECTIVE STOPPED CLOCK. I anticipate the following problems with this response. First, linguistic evidence tells against utterances like "I took a reading from a stopped clock, and I came to know the time". Upon reflection, this knowledge attribution sounds implausible. Second, the case features the double stroke of luck typical of Gettier-style vignettes: one bad (the clock stops) and one good (the clock stops at an only accidentally correct time). If knowledge is absent in the original version, it should be absent in this structurally similar version too. Perhaps the explanationist could treat the case as an unusual instance of testimony: Russell gains knowledge from an unreliable source as in other alleged cases of unreliable testimonial knowledge (Hawthorne 2004: 68; Goldberg 2005). While I take this to be the most promising response, I think it rests on a too inclusive (and hence distorted) conception of testimony: the case doesn't involve a single speech act, arguably a key defining feature of testimonial knowledge. For reasons of space, I will not pursue this issue further.

insufficient for knowledge. Standard Gettier cases like this call into question the adequacy of explanationism. As I proceed to show next, non-standard Gettier cases like fake-barn cases do the same and raise additional difficulties for the explanationist condition.

2.2. Explanationism exposed once more: fake barns

Consider another Gettier-style case, offered by Alvin Goldman (1976) but credited to Carl Ginet:

FAKE BARN

Barney, a reliable barn spotter, is driving through the countryside. He looks out of the window, sees a barn on the hill and comes to believe that he is looking at a barn. Whilst Barney's belief is true, unbeknownst to him, the structure he is looking at is the only real barn in an area filled with fake barns that are indistinguishable from real barns.

What can the explanationist say about FAKE BARN? Bogardus and Perrin (2020: 16) tentatively suggest a no-knowledge verdict.[10] However, their treatment of the case is ambiguous between two different *de re* and *de dicto* readings: once disambiguated, their diagnosis turns out problematic because of an especially troublesome closure failure. Let me take each point in turn and start with the details of their diagnosis of FAKE BARN.

Bogardus and Perrin begin by considering an ordinary case of seeing a barn, and treat it on a par with equally ordinary cases of perceptual knowledge. In such ordinary cases, this is the relevant explanation: "Barney believes that there's a barn on the hill because it looks like there's a barn on the hill and it's true that there's a barn on the hill". The truth of *p* (that there is a barn on the hill) crucially figures in a *complete* explanation of why Barney believes that there is a barn on the hill; accordingly, in this ordinary case, Barney knows. However, in FAKE BARN, a different explanation enters the picture. It goes as follows: "Barney believes that there's a barn on the hill because it looks like there's a barn on the hill. However, it looks like there's a barn on the hill *not* because it's true that there's a barn on the hill *but* only because he is in an environment where most objects *look like* barns". Otherwise put, Barney believes that there's a barn on the hill because he finds himself in Fake Barn Country. To back up such a potentially surprising diagnosis, Bogardus and Perrin offer the following remarks, worth quoting in full:

Even if your eyes happen to fall upon a real barn in a forest of fakes, we might begin to think it false that it looks like there's a barn before you because there is a barn before you. As the barn facades proliferate, a rival explanation looms into view: *that it looks like there's a barn before you because you're in a region full of structures that look like barns*. In other words, it becomes plausible to say that you believe there's a barn before you because it looks like there's a barn before

---

[10]For the purpose of this paper, I shall grant to Bogardus and Perrin that *some* no-knowledge verdict in FAKE BARN is correct. However, I also note that both epistemologists (Lycan 2006; Sosa 2010: 472–3; Turri 2016; Schellenberg 2018: 211) and experimental philosophers (Colaço *et al.* 2014) have adduced arguments to challenge the correctness of the no-knowledge verdict in FAKE BARN. Accordingly, I care to emphasise that, unlike DEFECTIVE STOPPED CLOCK, the ensuing objections do not clearly apply to other versions of explanationism that provide a different diagnosis of fake-barn cases. I am grateful to an anonymous reviewer for pushing me to clarify this point.

you, *and it looks like there's a barn before you because you're in Fake Barn Country.*
(Bogardus and Perrin 2020: 16; emphasis mine)

This is how Bogardus and Perrin aim to capture the no-knowledge verdict in FAKE BARN: unlike ordinary instances of perceptual knowledge, in this case the truth of *p* doesn't figure crucially in the explanation of why the agent believes that there's a barn. The agent doesn't believe that *p* because *p* is true: rather, they believe that *p* because *p* appears to be true. According to Bogardus and Perrin, in FAKE BARN we ought to explain Barney's true belief by appealing to the fact that objects look like barns, not the fact that they are real barns. Hence, no knowledge.[11]

It's worth noting that Bogardus and Perrin's explanationist diagnosis seems prima facie hard to make sense of. Let's zero in on the alleged 'rival explanation' of the belief under consideration: why exactly should we explain Barney's *true* belief by appealing to the fact that 'he's in a region full of structures that look like barns'? The explanation sounds ad hoc, and it doesn't go far enough. Suppose that Barney is back in real barn country, where barns look like barns and are actual barns. Should we also explain Barney's true belief by appeal to the fact that he's in a region full of structures that look like barns? If the explanation doesn't stop at appearances in the good case of real barns, it's unclear why it should stop at appearances in the 'bad' case of seeing the only real barn in the mists of papier-mâché copies. After all, in both cases the agent believes that there's a barn because it looks like there's a barn and it's true that there's a barn. The two appearance-based explanations are equally incomplete: Bogardus and Perrin's explanationism seems to lack the resources to provide a principled difference between the two cases. The main motivation for the no-knowledge verdict starts to lose plausibility.

But even granting that their explanationist diagnosis is correct (as I shall charitably grant), it applies at best to *one* version of the case. For sake of clarity, it is worth distinguishing between two versions of FAKE BARN – *de re* and *de dicto* respectively.[12] In the *de re* version, Barney comes to believe that *that object* on the hill is a barn. Barney's demonstrative belief is *distinctively* about that object, *that* particular barn on the hill. In the *de dicto* version, Barney comes to believe, more simply, that there is *a* barn on the hill. The indefinite article is crucial. Barney's *de dicto* belief is not demonstrative: it is not about *that* particular barn, it is just the belief that there is *a* barn on the hill.

With this distinction in play, let's revisit Bogardus and Perrin's explanationist diagnosis. Consider the *de re* version of FAKE BARN. In this case, the explanationist has to grant knowledge to *de re* barn beliefs: in fact, they are structurally identical to ordinary cases of knowledgeable perceptual beliefs. Barney truly believes that that very object on the hill is a barn because that very object looks like a barn and it's true that that very object is a barn: Barney forms a demonstrative true belief because *that* very object looks like and is a barn, not because *other* objects look like barns. Thus, the explanationist is committed to presence of knowledge *de re*, just like they predict presence of knowledge in the good case of seeing a barn and in general instances of perceptual knowledge. In general, explanationists do not have much of a choice here: since it would carry over to ordinary perceptual cases, a no-knowledge verdict in the *de re* version of FAKE BARN would open the door to a significant (and hence disturbing) sceptical threat.

---

[11]Another early advocate of explanationism, Alan Goldman (1984: 44), also agrees: "In such cases … the proper explanation for the belief appeals to the broader context of the perceiver's being in the vicinity of all these look-alike objects, any of which would produce the belief in question." Bogardus and Perrin acknowledge Alan Goldman's point in a footnote (Bogardus and Perrin 2020: 16, footnote 23).

[12]For discussion of *de re* barn beliefs, see Brown (2000), Pryor (2004: 71), Hawthorne (2004: 56, footnote 17), Hiller and Neta (2007: 312), and Bernecker (2020: 5107).

Now, on to the *de dicto* version. In this case, Bogardus and Perrin's explanationist diagnosis may apply: Barney truly believes that there's a barn on the hill not because it's true that there's a barn on the hill, but because many other objects in that portion of the environment look like barns. If the explanationist aims to capture absence of knowledge in FAKE BARN, then it inevitably has to be absence of knowledge *de dicto*. Thus, to sum up, Barney knows *de re* but fails to know *de dicto*.

This is, however, a problematic result. Suppose that some version of the closure principle on knowledge is true. For instance,

> **Closure**. If one knows P and competently deduces Q from P, thereby coming to believe Q while retaining one's knowledge that P, then one comes to know that Q. (Williamson 2000; Hawthorne 2005)

The principle enjoys independent plausibility, but defending it exceeds the scope of this paper. For my purposes, I shall more simply point out that if we conjoin the *de re* and *de dicto* versions of FAKE BARN then this independently plausible principle fails. Let's fill in the details as follows: let P and Q be the *de re* and the *de dicto* barn propositions respectively. Barney knows (*de re*) that that very object on the hill is a barn, he competently deduces that since that object is a barn then there's a barn on the hill but, according to Bogardus and Perrin, he fails to know (*de dicto*) that there's a barn on the hill. Borrowing DeRose's (1995) terminology, their explanationist analysis is committed to the following 'abominable' conjunction: Barney knows that that object on the hill is a barn but he does not know that there's a barn on the hill. We shouldn't accept this result light-heartedly: how could Barney know that that object on the hill is a barn and not know that there's a barn on the hill?[13]

Let me make the same point, but from a different angle. Suppose further that some version of the knowledge norm of assertion is correct. For example,

> **KNA**. One must: assert *p* if and only if one knows *p*.

This norm has prominent advocates (e.g., Unger 1975; Williamson 2000; Sosa 2015; Simion and Kelp 2020), but defending it here would also take me too far afield. For my purposes, I shall more simply note that according to Bogardus and Perrin's version of explanationism Barney could permissibly assert the following: "I know that that object on the hill is a barn, but I don't know that there's a barn on the hill". Given Barney's knowledge *de re*, this sounds infelicitous. Moreover, if Barney asserted the *de dicto* proposition that "There's a barn on the hill", according to Bogardus and Perrin's explanationist condition he would strictly speaking violate the knowledge norm of assertion (though perhaps only blamelessly). This is also implausible: Barney seems intuitively entitled to make this assertion, even more so if we concede with Bogardus and Perrin that he knows *de re* that that object is a barn. Yet another result we should not accept light-heartedly.[14]

On these last two points, I want to tread carefully. My objection does not rest on accepting the closure principle or the knowledge norm of assertion: the problem has to do with the too simple explanationist diagnosis of very hard fake-barn cases. Even granting that Barney fails to know in FAKE BARN, this applies only to the *de dicto* and

---

[13]For similar reasons, Bogardus and Perrin's version of explanationism also struggles with Kripke's red barn example (Kripke 2011: Chapter 7).

[14]See Hawthorne (2005: 32) for a defence of closure principles of knowledge based on considerations of conversational propriety.

not to the *de re* version of the vignette. Once we conjoin these two distinct versions, Bogardus and Perrin's explanationist diagnosis turns out untenable. It may well be that the closure principle *sometimes* fails, but this does not seem an *acceptable* case of closure failure. It may well be also that the infelicity of Barney's assertion can be explained away by further pragmatic considerations, or that there is no theoretical pressure to square an analysis of knowledge with an independent linguistic norm of assertion. However, granting knowledge *de re* and denying knowledge *de dicto* remains problematic, and this is because Bogardus and Perrin's explanationist analysis doesn't provide a complete and plausible treatment of the FAKE BARN vignette. The reason may be this: fake-barn cases are notoriously difficult, and put pressure on very complex analyses of knowledge.[15] Accordingly, it should not be so surprising that Bogardus and Perrin's explanationist condition cannot properly deal with them. On this front too, their version of explanationism fails to deliver.

Let me take stock. In the last two sections, I have argued that Bogardus and Perrin's explanationist condition remains insufficient for knowledge. In DEFECTIVE STOPPED CLOCK, their condition is met but knowledge is absent. In FAKE BARN, the no-knowledge verdict applies only to the *de dicto* version of the case, and this consequently leads to a problematic closure failure. Overall, the explanationist analysis fails to provide a satisfactory diagnosis of the two main types of Gettier-style cases.[16] As I proceed to argue next, a different modalist approach handles the very same cases rather easily.

3. Modalism defended

For my defence of modalism, I propose and endorse this formulation of the safety condition:

> **Safety**. In most or all close possible worlds in which S believes that *p* via the same method of belief formation M that S uses in the actual world and S occupies the same environment E that S occupies in the actual world, *p* is true.

Two points of clarifications are in order. First, like other formulations of safety, this version allows for safety failure in case of false beliefs in different propositions than *p*. The focus is not on whether a belief-forming method M employed in an environment E yields true beliefs only in *p*, but also in propositions relevantly similar to *p*.[17] Second,

---

[15]As McGlynn (2014: 173) nicely puts it: "Barn cases seem to bring out just how *demanding* a state knowledge really is; the moral seems to be that knowing that *p* makes demands on one's external environment to a greater degree than we might have otherwise expected." McGlynn's remarks equally apply to the explanationist analysis: knowing that *p* makes demands to one's external environment to a greater degree than just believing *p* because *p* is true.

[16]To forestall misunderstanding and for sake of clarity, I shall briefly consider a line of response on behalf of the explanationist. In a short footnote, Bogardus and Perrin (2020: 17, footnote 27) claim that since warrant entails truth, under explanationism there are no justified beliefs without knowledge and thus, a fortiori, no Gettier cases. This move does not let the explanationist off the hook. Firstly, factive justification incurs what Kelp (2018: 84) helpfully dubs the 'new' Gettier problem, the problem to explain why agents lack justification in Gettier-style cases. On this point, explanationists remain suspiciously silent. Secondly, and more importantly, even granting that Gettiered agents lack justification, they nevertheless fail to know: at best, they have a *true belief* rather than a *justified true belief*. The explanationist condition remains insufficient for knowledge *regardless* of whether justification is present. Overall, this move does not help the explanationist with the cases considered in this paper.

[17]Brown (2000), Williamson (2000) and Pritchard (2012) advocate similar versions of global-method safety. See Rabinowitz (2011), Bernecker (2020: section 3), and Hirvelä (2019) for discussion.

and *unlike* other formulations of safety, this version is relative to *both* methods and environments. Such double index betters captures the spirit of the safety condition on knowledge. Safety theorists aim to assess whether a specific enough belief-forming method would produce true beliefs in relevantly similar situations: to make such an assessment, it's crucial to focus on the method *and* the environment where the method is employed. Consider a non-epistemic case: if we aim to assess whether a car is safe, we will make sure to drive it in a suitably specified environment (say, in appropriate driving conditions). In general, judgements on safety tacitly assume environmental factors, so it's worth making these factors explicit in the formulation of the safety condition on knowledge. In light of these points, I will now show that such refined version of safety is *immune* to the ATOMIC CLOCK counterexample and handles the previous Gettier cases *better* than explanationism. With an eye on these two theoretical virtues, I shall finally reassess the score in the debate between the explanationist and the modalist.

Let's begin with the knockdown counterexample. The explanationist rejects every modalist analysis because of cases like ATOMIC CLOCK: subjects know in the actual world and yet they form a false belief in a *very similar nearby* world. The case purportedly shows that modal conditions such as safety are not necessary for knowledge. However, for the case to have teeth, the error-possibility has to obtain in a similar nearby world, and it is notoriously difficult to individuate *exactly* which worlds count as relevantly similar and sufficiently close. To reach a higher level of precision, we should *not* assess ATOMIC CLOCK by relying on the ordinary sense of 'close' or 'similar'. Instead, we should focus on the relevant subjunctive conditionals and hold fixed not only the method employed but also (and most crucially) the environment in which it is employed.[18] Thus, let's ask: were S to form beliefs via the same method M in the *same* environment E, would she continue to believe truly? With the proper focus on the relevant worlds, the answer is 'yes'. In ATOMIC CLOCK, the error-possibility requires a change in the environment: the isotope *has* to decay. If we keep both method and environment fixed and we focus on the worlds where the subject reads from the atomic clock and the isotope *doesn't* decay, then the subject continues forming true beliefs. The safety condition is satisfied: accordingly, the knowledge verdict is aptly captured. Consider the following figures:

Following Lewis (1973, 1986), let's imagine possible worlds falling into concentric circles centred on the actual world (the black sphere), with "proximity serving as a [visual] metaphor for similarity" (Smith 2016: 106). Figure 1 represents the ATOMIC CLOCK counterexample. The wide dotted sphere includes the worlds where *only* the method (reading from the atomic clock) is held fixed: since the isotope decays, in these worlds the subject's belief is false. Figure 2 highlights the benefits of a double index to methods and environments. In the smaller striped sphere, we see the close similar worlds where *both* the method (reading from the clock) and the environment (the isotope does *not* decay) are held fixed: in such worlds, the subject's belief is true. Crucially, the belief turns out false only in the *further* dotted sphere, namely in the *less* similar worlds where the method is fixed but the environment changes. In the similar close worlds in the striped sphere, the subject's belief remains true: for the belief to be false, we

---

[18]A word of caution on the methodology adopted here. I invoke subjunctive conditionals because they give us a better grip on which worlds count as relevantly close. By attending to these properly formulated conditionals and focussing on their truth-values, we get a better (though not perfect) understanding of the modal profile of the belief under consideration. Crucially, this is not the *only* way to flesh out the safety condition: rather, it's just a useful heuristic to achieve clearer judgements about especially controversial cases. See Bogardus (2014: 6–9) for similar remarks and Smith (2016: 111–16) for an illuminating discussion on the modal structure of the safety condition on knowledge.
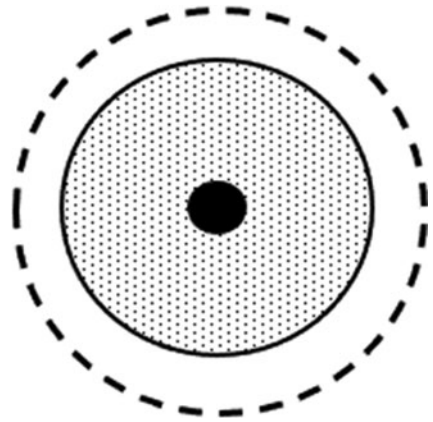
Fig 1. The dotted sphere represents the close worlds where only the method is held fixed. In these worlds, the belief under consideration is false and unsafe.
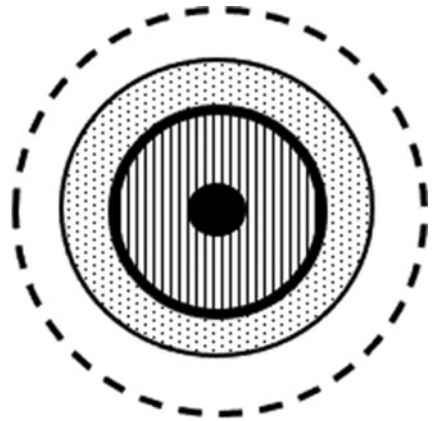

Fig 2. The striped sphere represents the close worlds where both the method and the environment are held fixed. In these worlds, the belief under consideration is true and safe.

need to travel out to the worlds in the *next* sphere, which may be similar but not similar enough to threaten a properly understood safety condition.

In ATOMIC CLOCK, the explanationist focusses on worlds that may seem intuitively close but nonetheless require a change in the environment: this change places them *further* away from the actual world. Thus, the case is not strong enough to motivate a full rejection of modalism. It's easy to come up with error-possibilities, but way harder to evaluate precisely how close they are: given a suitable index to both methods and environments, the error-possibility that the explanationist is appealing to is not after all close enough to render the target belief unsafe. Accordingly, the knockdown counterexample to modalism does not apply to the formulation of safety I am offering here. To do away with modal conditions on knowledge, the explanationist must do better.

Now, let's look at the Gettier cases that explanationism struggled with. Equipped with the refined safety condition canvassed above, the modalist accommodates the right verdict in each vignette. Consider again DEFECTIVE STOPPED CLOCK. Since the clock *actually* stops at 8:22 pm, we should focus on the close worlds where the clock *remains* stopped. Russell's belief is not safe, and hence does not amount to knowledge: in the close worlds where Russell looks at the stopped clock, he too easily forms a false belief about the time. In fact, Russell too easily forms false beliefs in distinct but relevantly

similar propositions. Suppose that, two minutes later, Russell forms the belief that two minutes have passed since 8:22 pm. Since the clock would *still* read 8:22 pm, Russell's belief in this distinct but relevantly similar proposition is false. Unlike the explanationist analysis, this refined formulation of safety correctly predicts that Russell fails to know.

Next, on to FAKE BARN. A properly understood safety condition delivers a no-knowledge verdict in *both* the *de re* and *de dicto* versions of the vignette. This is due, once again, to the fact that safety is best understood as globalised to a set of relevantly similar propositions. The agent lacks knowledge *de re* because they would easily form a false belief in distinct and yet relevantly similar *de dicto* and *de re* propositions. Barney fails to know that the very object he is looking at is a barn because in close possible worlds he could have easily formed the false belief that there's a barn on the hill, or that another barn-looking object is three-dimensional, inhabited by farmers, meant to stock hay, and so on (cf. Bernecker 2020: 5107). Crucially, by denying knowledge to the *de re* version of fake barn, this version of safety avoids the closure failure that the explanationist is inevitably committed to. On this count too, modalism is superior – and hence preferable – to explanationism.

Before concluding, one final and very important caveat. I do not pretend to have shown that the superiority of modalism is *definitive*. I am well aware that no formulation of safety is problem-free: modal conditions on knowledge incur serious (and perhaps unavoidable) difficulties.[19] One key aspect where explanationism does generally better than virtually any version of modalism is the case of knowledge of necessary truths which, as such, are true in every nearby world and thus trivially safe. Worse still for the sufficiency of modalist conditions on knowledge, the absence of an explanatory connection between the method employed and the truth of the belief in question suggests that beliefs in necessary truths may be only luckily true. Conversely, given the key focus on explanatory – rather than modal – connections between belief and truth, explanationists easily deal with such long-standing problems afflicting modalist analyses of knowledge.[20] For all these reasons, a full-fledged defence of modalism exceeds the scope of this paper. My aim is far more modest: all I argue is that when it comes to the Gettier-style cases considered here, my modalist condition easily delivers the intuitively correct verdicts, while Bogardus and Perrin's explanationist analysis clearly does not. This gives us defeasible but strong enough reason to prefer my modalist approach to Bogardus and Perrin's explanationist alternative.

4. Concluding remarks

I have considered two competing approaches to the task of capturing the notion of non-accidentality in the analysis of knowledge: the explanationist and the modalist. I focused mostly on the most recent version of the former, and found it wanting. Bogardus and

---

[19]Modalists face four uphill battles in the form of the following questions. First, in what sense are beliefs, methods and environments relevantly *similar*? The vagueness of the notion of similarity is a well-known source of problems for modalism. Second, how do we individuate the set of propositions that the safety condition is globalised to? Modalists usually rely on an intuitive grasp of the set of relevantly similar propositions without being too specific on this point. Third, how *fine-grained* does the individuation of the belief-forming method and environments have to be? This is a version of the generality problem, which has plagued modalist analyses of knowledge for a long time. While certainly pressing, I hasten to flag that none of these issues affect my *modest* defence of modalism: the action here is primarily on Gettier-style cases rather than on the more general problems that affect modalism.

[20]See especially Faraci (2019: 12–15), Korman and Locke (Forthcoming *b*: section 3) and Lutz (2020) for discussion of the advantages of explanationism over modalism in the case of beliefs in necessary truths. See Hirvelä (2019) for a promising modalist attempt to deal with the issue.

Perrin's explanationist approach promises to offer a full-blooded reductive analysis of knowledge without modal conditions. However, when tested against a broader range of Gettier-style cases, such an approach does not live up to its promises. Their explanationist condition is not sufficient for knowledge: in DEFECTIVE STOPPED CLOCK, the agent believes that p because p is true, and yet they lack knowledge. Moreover, their version of explanationism fails to provide a satisfactory diagnosis of fake-barn cases: by delivering a knowledge verdict *de re* and an ignorance verdict *de dicto*, Bogardus and Perrin are committed to a troublesome closure failure. I take this to be a more general indication of the inadequacy of the explanationist approach: however unpacked, the notion of explanation does not seem to be well suited to deliver a satisfactory analysis of knowledge. It is of course possible that other versions of explanationism may do better than Bogardus and Perrin's, but in light of the present discussion their specific explanationist proposal remains problematic.

What about the modalist? Far from being perfect, the safety condition canvassed above has at least two important virtues. First, it escapes the knockdown counterexample raised against *any* modalist approach: ATOMIC CLOCK poses no threat to such refined formulation of safety. Second, it has a relatively easy time in dealing with the same Gettier cases that beset explanationism: the formulation delivers the correct verdict of ignorance in each of the Gettier-style vignettes considered in this paper. Does this suggest that some version of the safety condition ultimately yields a successful analysis of knowledge? I very much doubt it, and epistemologists should not get their hopes up. However, if there exists such an analysis at all, it is more likely to rest on a modal rather than an explanationist condition – or at least so I've argued here.[21]

References

Antognazza M.R. (2015). 'The Benefit to Philosophy of the Study of its History.' *British Journal for the History of Philosophy* **23**(1), 161–84.

Beddor B. and Pavese C. (2020). 'Modal Virtue Epistemology.' *Philosophy and Phenomenological Research* **101**(1), 61–79.

Bernecker S. (2020). 'Against Global Method Safety.' *Synthese* **197**(12), 5101–16.

Berto F. and Nolan D. (2021). 'Hyperintensionality.' In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/sum2021/entries/hyperintensionality/.

Bogardus T. (2014). 'Knowledge Under Threat.' *Philosophy and Phenomenological Research* **88**(2), 289–313.

Bogardus T. and Perrin W. (2020). 'Knowledge is Believing Something Because It's True.' *Episteme*. https://doi.org/10.1017/epi.2020.18.

Brown J. (2000). 'VI-Reliabilism, Knowledge, and Mental Content.' *Proceedings of the Aristotelian Society* **100**(2), 115–35.

Clarke-Doane J. and Baras D. (2021). 'Modal Security.' *Philosophy and Phenomenological Research* **102**(1), 162–83.

Colaço D., Buckwalter W., Stich S. and Machery E. (2014). 'Epistemic Intuitions in Fake barn Thought Experiments.' *Episteme* **11**(2), 199–212.

Comesaña J. (2005). 'Unsafe Knowledge.' *Synthese* **146**(3), 395–404.

Craig E. (1990). *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Oxford University Press.

DeRose K. (1995). 'Solving the Sceptical Problem.' *Philosophical Review* **104**(1), 1–52.

Faraci D. (2019). 'Groundwork for an Explanationist Account of Epistemic Coincidence.' *Philosophers' Imprint* **19**. https://philpapers.org/rec/FARGFA.

Goldberg S. (2005). 'Testimonial Knowledge Through Unsafe Testimony.' *Analysis* **65**, 302–11.

Goldberg S. (2015). 'Epistemic Entitlement and Luck.' *Philosophy and Phenomenological Research* **91**(2), 273–302.

Goldman A. (1976). 'Discrimination and Perceptual Knowledge.' *Journal of Philosophy* **73**, 771–91.

Goldman A.H. (1984). 'An Explanatory Analysis of Knowledge.' *American Philosophical Quarterly* **21**(1), 101–8.

Grundmann T. (2018). 'Saving Safety from Counterexamples.' *Synthese* **197**(12), 5161–85.

Hawthorne J. (2004). *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hawthorne J. (2005). 'The Case for Closure.' In M. Steup and E. Sosa (eds), *Contemporary Debates in Epistemology*, pp. 26–43. Oxford: Blackwell.

Hiller A. and Neta R. (2007). 'Safety and Epistemic Luck.' *Synthese* **158**(3), 303–13.

Hirvelä J. (2019). 'Global Safety: How to Deal with Necessary Truths.' *Synthese* **196**(3), 1167–86.

Ichikawa J. and Steup M. (2018). 'The Analysis of Knowledge.' In E.N. Zalta (ed.), *Stanford Encylopedia of Philosophy*. https://plato.stanford.edu/archives/sum2018/entries/knowledge-analysis/.

Jenkins C.S. (2005). 'Knowledge and Explanation.' *Canadian Journal of Philosophy* **36**(2), 137–64.

Kelp C. (2009). 'Knowledge and Safety.' *Journal of Philosophical Research* **34**, 21–31.

Kelp C. (2013). 'Knowledge: The Safe-Apt View.' *Australasian Journal of Philosophy* **91**(2), 265–78.

Kelp C. (2018). *Good Thinking: A Knowledge First Virtue Epistemology*. London: Routledge.

Korman D.Z. and Locke D. (Forthcoming *a*). 'Against Minimalist Responses to Moral Debunking Arguments.' *Oxford Studies in Metaethics*.

Korman D.Z. and Locke D. (Forthcoming *b*). 'An Explanationist Account of Genealogical Defeat.' *Philosophy and Phenomenological Research*.

Kripke S. (ed.) (2011). *Philosophical Troubles. Collected Papers Vol. I*. Oxford: Oxford University Press.

Lewis D. (1973). *Counterfactuals*. Oxford: Blackwell.

Lewis D. (1986). *On the Plurality of Worlds*. Oxford: Wiley-Blackwell.

Lycan W.G. (2006). 'On the Gettier Problem Problem.' In S. Hetherington (ed.), *Epistemology Futures*, pp. 148–68. Oxford: Oxford University Press.

Lutz M. (2020). 'The Reliability Challenge in Moral Epistemology.' In R. Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 15*. Oxford: Oxford University Press.

McGlynn A. (2014). 'Is Knowledge a Mental State?' In *Knowledge First?* London: Palgrave Macmillan.

Neta R. (2002). 'S Knows That P.' *Noûs* **36**(4), 663–81.

Neta R. and Rohrbaugh G. (2004). 'Luminosity and the Safety of Knowledge.' *Pacific Philosophical Quarterly* **85**(4), 396–406.

Nolan D. (2014). 'Hyperintensional Metaphysics.' *Philosophical Studies* **171**(1), 149–60.

Nozick R. (1981). *Philosophical Explanations*. Cambridge, MA: Harvard University Press.

Pritchard D. (2005). *Epistemic Luck*. Oxford: Oxford University Press.

Pritchard D. (2012). 'Anti-Luck Virtue Epistemology.' *Journal of Philosophy* **109**(3), 247–79.

Pritchard D. (2014). 'Anti-luck Epistemology and the Gettier Problem.' *Philosophical Studies* **172**(1), 93–111.

Pryor J. (2004). 'Comments on Sosa's "Relevant Alternatives, Contextualism Included".' *Philosophical Studies* **119**(1–2), 67–72.

Rabinowitz D. (2011). 'The Safety Condition for Knowledge.' *Internet Encyclopedia of Philosophy*. https://iep.utm.edu/safety-c/.

Rieber S. (1998). 'Skepticism and Contrastive Explanation.' *Noûs* **32**(2), 189–204.

Sainsbury R.M. (1997). 'Easy Possibilities.' *Philosophy and Phenomenological Research* **57**(4), 907–19.

Schellenberg S. (2018). *The Unity of Perception: Content, Consciousness, Evidence*. Oxford: Oxford University Press.

Simion M. and Kelp C. (2020). 'The Constitutive Norm View of Assertion.' In S. Goldberg (ed.), *The Oxford Handbook of Assertion*. Oxford University Press.

Smith M. (2016). *Between Probability and Certainty: What Justifies Belief*. Oxford: Oxford University Press.

Sosa E. (1999). 'How Must Knowledge be Modally Related to What is Known?' *Philosophical Topics* **26**(1–2), 373–84.

Sosa E. (2010). 'How Competence Matters in Epistemology.' *Philosophical Perspectives* **24**(1), 465–75.

Sosa E. (2015). *Judgment and Agency*. Oxford: Oxford University Press.

Strevens M. (2008). *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.

Turri J. (2016). 'Knowledge and Assertion in 'Gettier' Cases.' *Philosophical Psychology* **29**(5), 759–75.

Unger P. (1975). *Ignorance: The Case for Scepticism*. Oxford: Clarendon Press.
Williamson T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
Zagzebski L. (1994). 'The Inescapability of Gettier Problems.' *Philosophical Quarterly* **44**(174), 65–73.

Dario Mortini is completing his PhD within the COGITO Epistemology Research Centre at the University of Glasgow. His research lies at the intersection of epistemology and philosophy of law.