



## Knowledge, justification, and (a sort of) safe belief

Daniel Whiting<sup>1</sup> 

Received: 18 March 2018 / Accepted: 9 August 2018  
© The Author(s) 2018

### Abstract

An influential proposal is that knowledge involves safe belief. A belief is safe, in the relevant sense, just in case it is true in nearby metaphysically possible worlds. In this paper, I introduce a distinct but complementary notion of safety, understood in terms of epistemically possible worlds. The main aim, in doing so, is to add to the epistemologist's tool-kit. To demonstrate the usefulness of the tool, I use it to advance and assess substantive proposals concerning knowledge and justification.

**Keywords** Knowledge · Justification · Safety · Epistemically possible worlds · Lottery paradox

### 1 Introduction

Consider the idea that there is a *safety* condition on knowledge (see Pritchard 2005: ch. 6; Sainsbury 1997; Sosa 1999; Williamson 2000, pp. 123–128). A rough way to capture it is to say that a person cannot know a proposition if it could easily have been false. Safety, so understood, is a modal notion. There are various ways to unpack it—perhaps the most prominent is in terms of metaphysically possible worlds (MPWs). Consider (where S is a subject, P is a proposition, and E is a proposition which belongs to the evidence she actually possesses):

---

The ideas in this paper have been in development for a long time. I presented earlier versions, some of which bear little resemblance to this one, at the University of Oslo, Keele University, the University of Southampton, and the University of Luxembourg. Thanks to participants at those events for helpful feedback, especially Nick Hughes and Cameron Boulton, my respondents at, respectively, Oslo and Southampton. Thanks also to anonymous referees, Sandy Goldberg, Alex Gregory, Conor McHugh, Neil Mehta, Duncan Pritchard, Kurt Sylvan, Lee Walters, and Jonathan Way for comments on the material or conversations on the issues it concerns (on multiple occasions in some cases). Finally, thanks to the British Academy (SG110243) and the Arts and Humanities Research Council (AH/K008188/1) for funding in support of this research.

---

✉ Daniel Whiting  
d.whiting@soton.ac.uk

<sup>1</sup> Department of Philosophy, University of Southampton, Southampton SO171BJ, UK

Safety<sub>M</sub> It is safe<sub>M</sub> for S to believe P on E if and only if, in all nearby MPWs in which E is true, P is true.<sup>1</sup>

For a person to believe a proposition on her evidence is for her to base that belief on that evidence. It can be safe<sub>M</sub> for a person to believe a proposition on evidence she has, even if she does not in fact believe it on that or any other evidence. The principle concerns what one might call *ex ante*, not *ex post*, safety<sub>M</sub>.

In view of Safety<sub>M</sub>, we can restate the condition on knowledge as follows:

$K \Rightarrow S_M$  S knows P on E only if it is safe<sub>M</sub> for S to believe P on E

I do not here argue for or defend  $K \Rightarrow S_M$ . Instead, I introduce an analogue of Safety<sub>M</sub>, which is, I think, new to the literature. The proposal is straightforward—replace reference to MPWs with reference to epistemically possible worlds (EPWs):

Safety<sub>E</sub> It is safe<sub>E</sub> for S to believe P on E if and only if, in all nearby EPWs in which E is true, P is true.<sup>2</sup>

The main goal is to present and explain this notion of safety<sub>E</sub>. The secondary goal is to demonstrate its significance by using the notion to formulate and explore substantive proposals about knowledge and justification.

I start by unpacking the notion of safety<sub>E</sub> (Sect. 2) and comparing it with that of safety<sub>M</sub> (Sect. 3). Next I propose that safety<sub>E</sub> is a necessary condition on knowledge (Sect. 4). I then turn to justification (Sect. 5). After some preliminaries, I suggest that, unlike safety<sub>M</sub>, safety<sub>E</sub> is a necessary condition on justification. Having done so, I explore the possibility that safety<sub>E</sub> is also sufficient for justification. Before concluding, I explain briefly how the accounts of justification in terms of safety<sub>E</sub> differs from accounts which might seem superficially similar due to Pritchard (2005), Smith (2016), and Wedgwood (2002) (Sect. 6). The aim is not to assess those accounts—a task for another occasion—but only to distinguish them from my own.

If the substantive proposals I make along the way fail to convince, the paper will succeed in its primary aim of adding the notion of safety<sub>E</sub> to the epistemologist's tool-kit. I hope to demonstrate that, by appeal to that notion, it is possible to articulate and evaluate theoretical options that are otherwise invisible. Even if one rejects those options, it is worth having them in view. There are no doubt more ways of exploiting the notion of safety<sub>E</sub> than I consider here.

<sup>1</sup> I owe this formulation to Smith (2016, pp. 20–21). There are other safety<sub>M</sub> conditions in circulation. Safety<sub>M</sub> here is evidence- or basis-relative. For cases which motivate this, see Sosa (2007, pp. 26–27). Some relativize safety<sub>M</sub> to methods or ways of forming beliefs (e.g. Greco 2016; Pritchard 2005, p. 156). Basis- and method-relative conditions need not be in competition (cf. Goldman 2009, pp. 83–85).

<sup>2</sup> Some associate MPWs with subjunctive conditionals (e.g. Lewis 1973). Likewise, some associate EPWs with indicative conditionals (e.g. Weatherson 2001). I take no stand here on whether one might restate the antecedents of Safety<sub>M</sub> or Safety<sub>E</sub> using conditionals.

## 2 Safety<sub>E</sub>

Recall:

Safety<sub>E</sub> It is safe<sub>E</sub> for S to believe P on E if and only if, in all nearby EPWs in which E is true, P is true

This raises several issues, addressing which helps in getting a purchase on the notion of safety<sub>E</sub>. First, what is a person's evidence? Following Williamson (2000, ch. 9), I assume that it is her knowledge:

$E = K$  S knows P if and only if P is E

This is an 'externalist' view—according to it, a person's evidence does not supervene upon her non-factive mental states.  $E = K$  is controversial (see McGlynn 2014, ch. 4). While I am broadly sympathetic to it, nothing in what follows hangs on that conception of evidence. I assume  $E = K$  for the sake of concreteness and for ease of presentation. If one replaces it with some other—perhaps 'internalist'—conception of evidence, one arrives at a different standard of safety<sub>E</sub>. One can then proceed to put safety<sub>E</sub>, so understood, to the uses I explore below.<sup>3</sup>

To make the same point differently, one might recognise a plurality of safety<sub>E</sub> principles, one formulated in terms of what a person knows, another formulated in terms of what she believes, another formulated in terms of what she experiences, and so on. These principles need not be in a competition and might serve different explanatory purposes. The debate as to which of those principles is equivalent to the one formulated in terms of a person's evidence is important, but orthogonal to the concerns of this paper.

A second question Safety<sub>E</sub> raises is: what is an EPW? It is epistemically possible for a person that p if and only if she is not in a position to know a priori that not-p. An EPW, then, is one which a person cannot rule out a priori, or one which is a priori coherent (see Chalmers 2011, p. 63). While there is no MPW in which water is not H<sub>2</sub>O, there is an EPW, so understood, in which water is not H<sub>2</sub>O.<sup>4</sup>

Which worlds are epistemically possible, in this sense, might vary from person to person, depending on their cognitive capacities and opportunities to exercise them. One person might be able to rule out a priori a world while another person is unable to do so. I return to this later.

Third, what is nearness relative to? When making judgements about safety<sub>M</sub>, the MPWs are ordered according to how close they are to the world the subject occupies. When making judgements about safety<sub>E</sub>, in contrast, EPWs are ordered according to

<sup>3</sup> I make two further assumptions about evidence. First, evidence is propositional. This is considerably less controversial than  $E = K$ . Defenders include Dougherty (2011), Littlejohn (2012, ch. 3), Neta (2008), and Williamson (2000, ch. 9). If one rejects that assumption in favour of the view that evidence consists of mental states (cf. Mitova 2015), for example, one can reformulate the safety principles in terms of when the content of the mental state is true in nearby worlds. Second, I assume a Fregean view of propositions, hence, evidence. To accommodate a Russellian view, one might reformulate Safety<sub>E</sub>, very roughly, as follows: It is safe<sub>E</sub> for S to believe P on E if and only if, in all nearby EPWs in which a proposition with the same cognitive significance (or mode of presentation) as E is true, P is true.

<sup>4</sup> There are, of course, other conceptions of EPWs. I do not challenge them. They might serve certain explanatory purposes, but they do not serve present purposes.

how close they are to a person's total evidence.<sup>5</sup> A person's total evidence is not maximal. So, it does not determine an EPW. Instead, it determines an epistemic *possibility* (cf. Humberstone 1981; Rumfitt 2015, ch. 6).<sup>6</sup> So, the EPWs are ordered according to how close they are to an epistemic possibility, rather than an EPW.

Nearness, then, is not a relation that holds only between worlds; it is a relation that can also hold between a (non-maximal) possibility and a world (or maximal possibility). Of course, a possibility determines a set of worlds, but it does not follow that the nearness relation really (or only) holds between a world in that set and another world.

A provisional rationale for relativizing nearness to the epistemic possibility that a person's evidence determines is that what  $\text{safety}_E$  is supposed to measure is not whether, relative to a world, a belief could easily be false but whether, relative to a subject's perspective, a belief could easily be false. Unlike a world, a subject's perspective is partial or incomplete. The real support for relativizing nearness in this way is shown by the work which the notion of  $\text{safety}_E$ , so understood, does.

Fourth, what is the measure of nearness?<sup>7</sup> I suggest that we understand nearness of EPWs in a way familiar from discussions of nearness of MPWs, namely, in terms of *similarity*.<sup>8</sup> EPW<sub>1</sub> is nearer for a person than EPW<sub>2</sub> if and only if EPW<sub>1</sub> is more similar in relevant respects to her total evidence than EPW<sub>2</sub>. If EPW<sub>1</sub> and EPW<sub>2</sub> are both consistent with the epistemic possibility that the subject's total evidence determines, EPW<sub>1</sub> might still be nearer—more similar in relevant respects—than EPW<sub>2</sub>. By way of analogy, suppose that an author writes the opening chapter of a novel, shows it to her publisher, and secures a contract. The author then produces two finished novels, each containing a version of the original chapter. The publisher might sensibly ask which of the two novels is closer to—more similar to or in keeping with—the original chapter. Indeed, she might ask this even if the original chapter remains unchanged in both novels.

A lesson from the literature on conditionals is that one must treat the relevant notion of similarity with caution. Consider: Were Donald to press the button, there would be a nuclear holocaust.<sup>9</sup> Suppose that this is true. According to Lewis's influential account

<sup>5</sup> The evidence on which it is  $\text{safe}_E$  for a person to believe— $E$  as it figures in  $\text{Safety}_E$ —need not be her total evidence; it might only be part of it.

<sup>6</sup> Given  $E = K$ , that possibility is also a metaphysical possibility. More fully: Given  $E = K$ ,  $S$ 's total evidence is true in the actual world. So, the possibility her total evidence determines obtains in the actual world. The actual world is metaphysically possible. So, the possibility that her total evidence determines is metaphysically possible. Again, for the purposes of this paper,  $E = K$  is an optional commitment.

<sup>7</sup> Baumann objects that 'safety theorists or, more generally, epistemologists who propose a modal condition for knowledge usually don't even raise the question of what determines closeness of possible worlds' (2008, p. 26). Bogardus suggests that, 'if we would like to know whether a given belief was formed safely [...], we should not consult our ordinary intuitions about similarity or nearness relations among worlds or cases. Rather, our first order of business should be to consult our intuitions about the truth values for [...] subjunctive conditionals' (2014, p. 294). Whatever the merits of this strategy for evaluating claims about  $\text{safety}_M$ , I cannot simply co-opt it for evaluating claims about  $\text{safety}_E$  (cf. fn2).

<sup>8</sup> An alternative proposal is that EPW<sub>1</sub> is nearer for a person than EPW<sub>2</sub> just in case EPW<sub>1</sub> is consistent with fewer revisions to her evidence than EPW<sub>2</sub> (see Weatherston 2001, p. 248). This might be suitable for certain explanatory projects but not for present purposes. The ordering of EPWs it results in is too coarse-grained—all EPWs compatible with a person's evidence will be equally near.

<sup>9</sup> I adapt this example from one Lewis (1979) discusses.

(1973), the conditional is true just in case, in nearby MPWs in which Donald presses the button, there is a nuclear holocaust. But—assuming that Donald does not in fact press the button—a MPW in which there is a nuclear holocaust is *dissimilar* to the actual world in many respects. This might suggest that no such world is nearby, hence, that the conditional is false. In response, Lewis (1979) suggests that principles along the following lines govern the measure of similarity (in order of importance): avoid widespread law-violations; keep large regions of the world the same; avoid local law-violations. By this measure, MPWs in which Donald presses the button and there is a nuclear holocaust count as sufficiently similar to the actual world, hence, as nearby, and so the conditional comes out as true.

In what follows, I understand talk of similarity, hence, nearness, in this familiar and well-established way.<sup>10</sup> To adapt the example, suppose that Hilary knows that Donald pressed the button. This, together with her background evidence, does not rule out the possibility that the missiles will fail to launch, that North Korea will not retaliate, and so on. But, relative to her evidence about what is happening and how things generally work, an EPW in which there will be a nuclear holocaust is more similar in the relevant sense, hence, nearer, than an EPW in which there will be no nuclear holocaust. Hence, it is  $\text{safe}_E$  for Hilary to believe that there will be a nuclear holocaust on the basis that Donald pressed the button.<sup>11</sup>

The final question is: what counts as nearby? That is, what determines whether a world is, not only near, but near enough? A plausible, though optional, answer to that question is: context. This might be the context of the object of assessment or the context of the assessor. I leave this open.

### 3 Compare and contrast

No doubt there are ways in which one might try to make the notion of  $\text{safe}_E$  more precise but I hope to have said enough to give it substance and to afford a grip on it. I now explore how it is like and unlike the notion of  $\text{safe}_M$  before putting it to work in articulating substantive proposals.

First, like  $\text{safe}_M$  (cf. Williamson 2000, ch. 5),  $\text{safe}_E$  does not iterate. It might be  $\text{safe}_E$  for S to believe P on E but not  $\text{safe}_E$  for S to believe on E that it is  $\text{safe}_E$  for S to believe P on E. Possibilities that are remote when assessing first-order beliefs might be close when assessing higher-order beliefs.

If it is  $\text{safe}_E$  to believe P, P is true in the nearby EPWs.<sup>12</sup> So, if it is  $\text{safe}_E$  to believe that it is  $\text{safe}_E$  to believe P, it is  $\text{safe}_E$  to believe P in the nearby EPWs. If it is  $\text{safe}_E$  to

<sup>10</sup> Of course, there are widely-appreciated difficulties here. Which similarities are relevant? What determines this? How are they weighted? These are not difficulties peculiar to the views I advance here—they arise for any theorist appealing to the idea of worlds standing in relations of similarity.

<sup>11</sup> The detour via conditionals is one way to introduce the measure of nearness. However, it is not the only way. One might eschew talk of similarity altogether and suggest that we have an independent and suitably firm grip on the idea of a close possibility. Alternatively, one might suggest that judgements as to whether an EPW is nearby are informed by judgements as to what a person is justified in believing. Later I introduce the view that a justified belief is a  $\text{safe}_E$  belief. So, this suggestion would make that view circular. But, to make a point I return to, circularity need not be vicious. (Compare Williamson 2000, p. 100).

<sup>12</sup> For simplicity, I omit reference here to evidence.

believe  $P$  in the nearby EPWs,  $P$  is true in the EPWs nearby the nearby EPWs. But, if  $P$  is true in the nearby EPWs, it does not follow that  $P$  is true in the EPWs nearby the nearby EPWs. So, if it is  $\text{safe}_E$  to believe  $P$ , it does not follow that it is  $\text{safe}_E$  to believe that it is  $\text{safe}_E$  to believe  $P$ . Compare: if the nearby shops have milk, it does not follow that the shops nearby the nearby shops have milk.

A second feature that  $\text{safety}_E$  shares with  $\text{safety}_M$  is that it obeys closure principles such as:

If it is  $\text{safe}_E$  for  $S$  to believe  $P$  on  $E$ , and it is  $\text{safe}_E$  for  $S$  to believe  $Q$  on  $E$ , and  $P$  and  $Q$  a priori entail  $R$ , then it is  $\text{safe}_E$  for  $S$  to believe  $R$  on  $E$ .<sup>13</sup>

If it is  $\text{safe}_E$  to believe  $P$  on  $E$ ,  $P$  is true in the nearby EPWs in which  $E$  is true. If it is  $\text{safe}_E$  to believe  $Q$  on  $E$ ,  $Q$  is true in the nearby EPWs in which  $E$  is true. So,  $P$  and  $Q$  are true in the nearby EPWs in which  $E$  is true.  $P$  and  $Q$  a priori entail  $R$ . So,  $R$  is true in the nearby EPWs in which  $E$  is true. So, it is  $\text{safe}_E$  to believe  $R$  on  $E$ .

An important respect in which  $\text{safety}_M$  and  $\text{safety}_E$  differ is that, while  $\text{safety}_M$  is factive, given a factive conception of evidence,  $\text{safety}_E$  is non-factive. When it is  $\text{safe}_M$  to believe a proposition, that proposition is true; but, when it is  $\text{safe}_E$  to believe a proposition, that proposition need not be true.

If it is  $\text{safe}_M$  for  $S$  to believe  $P$  on  $E$ ,  $P$  is true in the nearby MPWs in which  $E$  is true. Given  $E = K$ ,  $E$  is true in  $S$ 's world. And  $S$ 's world is trivially nearby, since it is more similar to itself than any other MPW. So, if it is  $\text{safe}_M$  for  $S$  to believe  $P$  on  $E$ ,  $P$  is true (cf. Smith 2016, p. 106).

If it is  $\text{safe}_E$  for  $S$  to believe  $P$  on  $E$ ,  $P$  is true in the nearby EPWs in which  $E$  is true. Given  $E = K$ ,  $E$  is true in  $S$ 's world. However,  $S$ 's world need not be nearby relative to  $S$ 's total evidence.  $S$ 's world, a maximal epistemic possibility, is consistent with the non-maximal epistemic possibility her total evidence determines, given  $E = K$ . Nonetheless,  $S$ 's world might be less similar to that non-maximal possibility in relevant respects than other EPWs. In particular, and recalling the Lewisian measures, large regions of  $S$ 's world might differ from that region of it her evidence concerns, or the principles which hold according to  $S$ 's evidence might not hold for  $S$ 's world as a whole. As a result, when it is  $\text{safe}_E$  for  $S$  to believe  $P$  on  $E$ ,  $P$  might not be true. As one might put it, the world from God's total perspective might turn out to be quite different to that part of it which falls within a person's partial perspective.

An earlier example illustrates this point. If Hilary believes that there will be a holocaust on the basis that Donald pressed the button, her belief is  $\text{safe}_E$ . However, if the missiles fail to launch due to some computer malfunction, or if there is no retaliation thanks to a shift in North Korean strategy of which Hilary is ignorant, her belief is false.

## 4 Knowledge

So far, I have introduced the notion of  $\text{safety}_E$  and noted some of its features. I now put it to work in articulating substantive proposals. Recall  $K \Rightarrow S_M$ . I take no stand

<sup>13</sup> This is a synchronic principle (cf. Dretske 2005; Pritchard 2005, p. 27; Vogel 1990). Some formulate diachronic closure principles. Synchronic and diachronic principles need not compete.

here on whether it is true. Instead, I suggest that  $\text{safety}_E$  is a necessary condition on knowledge:

$$K \Rightarrow S_E \quad S \text{ knows } P \text{ on } E \text{ only if it is } \text{safe}_E \text{ for } S \text{ to believe } P \text{ on } E$$

The claim that, if it is not  $\text{safe}_E$  for a person to believe a proposition on evidence she possesses, she cannot know it on that evidence is non-trivial. For one thing, it has explanatory significance. To see this, consider *lottery propositions*. Suppose that Miyuki knows that she holds a ticket in a lottery, there are one thousand tickets, and a winning ticket has been drawn. It is highly probable on Miyuki's evidence that the ticket lost. Moreover, it is true that her ticket lost. However, a widespread view is that Miyuki cannot know that her ticket lost and, more generally, that subjects cannot know lottery propositions.

$K \Rightarrow S_E$  accords with and explains this verdict. It is not  $\text{safe}_E$  for Miyuki to believe on her evidence that her ticket lost. In many of the nearby EPWs in which her evidence about the lottery is true, her ticket lost. However, in some of those EPWs, her ticket won.<sup>14</sup>

One might wonder whether, in ruling out knowledge of lottery propositions,  $K \Rightarrow S_E$  rules out ordinary inductive knowledge. It does not. Consider a revised version of an earlier example. Hilary believes that there will be a nuclear holocaust on the basis that Donald pressed the button. Her evidence does not entail the truth of her belief but, so long as EPWs in which, say, North Korea does not retaliate are suitably remote relative to her total evidence, that is, sufficiently dissimilar to the possibility it determines,  $K \Rightarrow S_E$  allows Hilary to know that there will be a holocaust.

One might ask whether we need  $K \Rightarrow S_E$  to explain why subjects cannot know lottery propositions. After all, one might think,  $K \Rightarrow S_M$  alone delivers this result.<sup>15</sup> However, while  $K \Rightarrow S_M$  predicts that in many cases subjects cannot know lottery cases, it does not do so in all cases. The key point here is that whether a belief is  $\text{safe}_M$  depends on matters that can lie beyond a person's ken.

<sup>14</sup> The claim that subjects cannot know lottery propositions, alongside closure principles for knowledge, raises a puzzle (see Hawthorne 2004; Vogel 1990). A full discussion of the thorny issues surrounding it is beyond the scope of this paper. However, I offer some brief remarks. To avoid distraction from the main narrative, I relegate them to this note.

Many propositions which one might take to be knowable entail lottery propositions. For example, one might think that Miyuki can know that she cannot afford a holiday. But, given closure, if she cannot know that her ticket lost, she cannot know this. To make matters worse, many ordinary propositions which one might take to be knowable, such as that I will watch a film later, entail lottery-like propositions, such as that I will not have a fatal heart-attack (which, given the statistics concerning heart-attacks among people in my circumstances, is highly likely but not certain). Given closure, if I cannot know the lottery-like proposition, I cannot know the proposition which entails it. Scepticism looms.

This puzzle is not unique to proponents of  $\text{safety}_E$  or, for that matter,  $\text{safety}_M$ ; it arises for any view according to which it is not possible to know lottery propositions. In brief, the response I favour echoes the response in Luper (2016), Mills (2012) and Smith (2016, pp. 54–58). I allow that, since Miyuki cannot know that her ticket lost, she cannot know that she cannot afford a holiday. But I deny that this generalises in a worrying fashion, since the allegedly lottery-like propositions are not relevantly similar to lottery propositions. In the above example, the evidence that I will not have a heart-attack is that I am healthy, have low blood pressure, enjoy a low-fat diet, show no signs of a heart disease, exercise regularly, and so on. Relative to this evidence, it is  $\text{safe}_E$  to believe that I will not have a fatal heart-attack, hence, that I will watch a film tonight.

<sup>15</sup> For the point that, in standard lottery cases, it is not  $\text{safe}_M$  to believe the relevant proposition, see Pritchard (2005, pp. 162–163).

Suppose that, unbeknownst to Miyuki, the lottery was rigged. Moreover, it is no accident that the lottery was rigged; there are powerful forces at work ensuring that Miyuki could never win the lottery. In this case, it is true, not only in the actual world that Miyuki's ticket lost, but in all nearby MPWs. So, it is consistent with  $K \Rightarrow S_M$  that Miyuki knows that her ticket lost.<sup>16</sup>

$K \Rightarrow S_E$  promises to explain why the subject lacks knowledge in this case.<sup>17</sup> Miyuki is unaware of the forces at work. So, the nearby EPWs in which her evidence is true—that is, the EPWs suitably similar in relevant respects to what she knows to be the case concerning the lottery—include EPWs in which her ticket won. So, it is not  $\text{safe}_E$  for her to believe that her ticket lost. So, given  $K \Rightarrow S_E$ , she cannot know this. To use the unofficial gloss, while Miyuki's ticket could not easily have won, it remains the case that, relative to her perspective, it could easily have won.

A similar line of thought shows that certain cases offered as counterexamples to  $K \Rightarrow S_M$  are not counterexamples to  $K \Rightarrow S_E$ . Nobody suggested otherwise but, given the similarity between the modal constraints, one might assume that a problem for one is a problem for both. Consider:<sup>18</sup>

On the basis that what she is drinking is tasteless, colourless, odourless, and from a bottle labelled 'water', Lily believes that it is water. However, unbeknownst to Lily, she stands near a person who won the lottery. Had this person lost, he would have caused a commotion during which he would have replaced the contents of the bottle with a tasteless, odourless, colourless toxin so as to vent his frustration.

Plausibly, Lily can know on her evidence that she is drinking water, but  $K \Rightarrow S_M$  seems not to allow this. There are nearby MPWs in which Lily's evidence is true—that is, in which what she is drinking is colourless, from a bottle labelled 'water', and so on—but in which she is drinking a toxin, namely, those MPWs in which her volatile neighbour lost.

I take no stand on whether this is a successful counterexample to  $K \Rightarrow S_M$ . The important point is that it is not a counterexample to  $K \Rightarrow S_E$ . The consideration that threatens to make Lily's belief  $\text{unsafe}_M$ —that her neighbour holds a lottery ticket and would have replaced the drink had it lost—is not part of her evidence. While there are EPWs in which she is drinking toxin—she cannot rule this out *a priori*—those EPWs are remote relative to Lily's total evidence. So, it is  $\text{safe}_E$  for Lily to believe that she is drinking water.

One might revise the case so that Lily knows that her neighbour would have replaced her drink with toxin had the ticket lost. In that case, Lily's belief that she is drinking water remains  $\text{safe}_E$ . Relative to her total evidence, which now includes knowledge

<sup>16</sup> Compare: 'counterfactual robustness can be ensured through a combination of features that is entirely fortuitous' (Lackey 2006, p. 289; cf. Roush 2005, pp. 122–123). The lottery case I present differs from the kinds of 'defeat' cases Lasonen-Aarnio (2010) discusses.

<sup>17</sup> Of course, it is not the only candidate explanation. If knowledge entails justification, then one might appeal to the accounts of justification I consider in Sect. 6 to explain why subjects lack knowledge in lottery cases of this sort.

<sup>18</sup> The example is due to Neta and Rohrbaugh (2004). For similar examples, see Baumann (2008), Bogardus (2014), Comesaña (2005), Kelp (2009) and Yamada (2011). No doubt there are further putative counterexamples in the literature. It is neither feasible nor desirable to discuss them all.

about her neighbour, an EPW in which she is drinking toxin is still remote. So,  $K \Rightarrow S_E$  allows Lily to know that she is drinking water. To use the unofficial gloss, while it is the case that, given Lily's perspective, she could easily have not been drinking water, it is not the case that, given her perspective, she could easily not be drinking water.

## 5 Justification

Before exploring links between safety<sub>E</sub> and justification, some preliminaries. First, I focus on outright justification, as opposed to degrees of justification. What it takes to be justified to some degree in believing a proposition, and what the relationship is between that status and being justified outright, are interesting questions but not my present concern.

Second, I focus on *ex ante* justification—what a person is justified in believing—rather than *ex post* justification—what a person justifiably believes. A standard thought is that a person justifiably believes a proposition on her evidence just in case she is justified in believing that proposition on her evidence and bases her belief on that evidence in the right kind of way. What is that way is another interesting question but not my present concern.

Third, I assume that justification is non-factive, in the sense that a person can be justified in believing a falsehood.<sup>19</sup> As it happens, for the purposes at hand a weaker assumption suffices, namely, that false beliefs can possess some positive epistemic status, specifically, the status they enjoy in Gettier (1963) scenarios. I give the label ‘justified’ to that status but the reader is welcome to use a different label (say, ‘reasonable’).<sup>20</sup>

Finally, in places I consider (a crude version of) the commonplace view that a person is justified in believing a proposition on her evidence just in case that proposition is sufficiently probable on her evidence. The aim in doing so is not to engage seriously with a competitor but to bring into relief, by way of contrast, features of a view that appeals to the notion safety<sub>E</sub>. More generally, the aim of this paper is constructive, not critical—I focus on developing positive proposals, rather than arguing against alternative views in circulation.

### 5.1 Necessary

As discussed, safety<sub>M</sub> is factive. Since justification is non-factive, safety<sub>M</sub> is not a necessary condition on justification. However, safety<sub>E</sub> is non-factive. So, it is a candidate condition on justification:

$J \Rightarrow S_E$   $S$  is justified in believing  $P$  on  $E$  only if it is safe<sub>E</sub> for  $S$  to believe  $P$  on  $E$

$J \Rightarrow S_E$  has intuitive appeal. To use the unofficial gloss, it is plausible that a person is not justified in believing a proposition when, relative to her perspective, it could easily

<sup>19</sup> For factive conceptions of justification, see Littlejohn (2012), Steglich-Petersen (2013), Sutton (2007) and Williamson (forthcoming).

<sup>20</sup> I do not identify that status with blamelessness. On the difficulties here, see Srinivasan (2015).

be false.<sup>21</sup> It also captures an intuitive link between justification and risk-avoidance. A person is not justified in believing a proposition unless the risk of making a mistake is, given her perspective, remote.<sup>22</sup>

In addition,  $J \Rightarrow S_E$  offers a resolution of the lottery paradox (from Kyburg 1961). Given her evidence, Miyuki does not know that her ticket lost. But is she justified in believing this? The assumption that she is, alongside further plausible principles, leads to an absurd conclusion. Consider:

- (i) Miyuki is justified in believing that  $\text{ticket}_1$  lost.
- (ii) If Miyuki is justified in believing that  $\text{ticket}_1$  lost, she is justified in believing that  $\text{ticket}_2$  lost, and justified in believing that  $\text{ticket}_3$  lost ... and justified in believing that  $\text{ticket}_{1000}$  lost.
- (iii) If Miyuki is justified in believing that  $\text{ticket}_1$  lost, and justified in believing that  $\text{ticket}_2$  lost, and justified in believing that  $\text{ticket}_3$  lost, ... and justified in believing that  $\text{ticket}_{1000}$  lost, Miyuki is justified in believing that  $\text{ticket}_1$  lost and  $\text{ticket}_2$  lost and  $\text{ticket}_3$  lost ... and  $\text{ticket}_{1000}$  lost.
- (iv) So, Miyuki is justified in believing that every ticket lost.

(iv) is surely false. To make matters worse, suppose that Miyuki is justified in believing that some ticket won. In that case (i) entails that Miyuki is justified in holding contradictory beliefs.

A proponent of the view that a person is justified in believing a proposition if it is sufficiently probable on her evidence is committed to (i), that is, to allowing that subjects are justified in believing lottery propositions (assuming that ‘sufficiently probable’ does not mean *certain*). The standard alternative, following Kyburg (1961), is to reject (iii), a closure principle.

I will not here rehearse the costs of this (see Smith 2016, pp. 79–83). Instead, I consider rejecting the starting assumption (i) that Miyuki is justified in believing that  $\text{ticket}_1$  lost. Importantly, one can deny this while allowing that Miyuki is justified in believing that  $\text{ticket}_1$  *probably* lost, or in having a high degree of credence in the proposition that it lost.<sup>23</sup> As importantly, to say that Miyuki is not justified in believing that  $\text{ticket}_1$  lost is not to say that she is justified in believing that it won. That belief too is unsafe.

$J \Rightarrow S_E$  explains why subjects are not justified in believing lottery propositions. As discussed above, it is not safe<sub>E</sub> for Miyuki to believe that  $\text{ticket}_1$  lost. Given  $J \Rightarrow S_E$ , she is not justified in believing this.

The view that subjects are not justified in believing lottery propositions is independently plausible and a number of philosophers accept it (e.g. Bird 2007, pp. 100–103;

<sup>21</sup> This assumes that a person’s evidence – rather than, say, her higher-order beliefs about her evidence – determines her perspective. One might reject this assumption. I do not want here to quibble over what counts as a perspective. For present purposes, I can retreat to the claim that a person is not justified in believing a proposition when, relative to her evidence, it could easily be false. This seems no less plausible.

<sup>22</sup> Pritchard (2015) connects risk and safety<sub>M</sub>.

<sup>23</sup> One might think that the claim that subjects are not justified in believing lottery propositions, alongside closure principles for justification, gives rise to a puzzle analogous to the puzzle concerning knowledge I mention in fn14. The response I endorse there carries across (*mutatis mutandis*).

Nelkin 2000; Ryan 1996; Smith 2016, ch. 3; Smithies 2012b, p. 271). The view explains why Miyuki is not justified in acting on the belief that ticket<sub>1</sub> lost, say, by throwing it away (cf. Smithies 2012b). If a person is not justified in believing a proposition, she is not justified in acting on it.<sup>24</sup> In addition, the view helps to explain why subjects cannot know lottery propositions (cf. Ryan 1996, p. 138), on the assumption that knowledge entails justification:

$$K \Rightarrow J \text{ If } S \text{ knows } P \text{ on } E, S \text{ is justified in believing } P \text{ on } E$$

One might worry that denying that lottery beliefs are justifiable commits one to an implausible ‘infallibilism’, according to which a person is not justified in believing a proposition when it is possible on her evidence that that proposition is false. It does not. Consider again Hilary, who, on the basis that Donald pressed the button, believes that there will be a nuclear holocaust. If she is to be justified in believing this,  $J \Rightarrow S_E$  does not require that her total evidence rule out EPWs in which no holocaust occurs, only that those EPWs be remote.

## 5.2 Sufficient

If safety<sub>E</sub> is a necessary condition on justification, that is a significant discovery. Might it also be a sufficient condition? Consider:

$$S_E \Rightarrow J \text{ If it is safe}_E \text{ for } S \text{ to believe } P \text{ on } E, S \text{ is justified in believing } P \text{ on } E$$

$S_E \Rightarrow J$  has intuitive appeal. To use the unofficial gloss, it is plausible that a person is justified in believing a proposition when, given her perspective, it could not easily be false. Moreover, like  $J \Rightarrow S_E$ ,  $S_E \Rightarrow J$  offers a way of capturing a link between justification and risk-avoidance. A person is justified in believing a proposition if the risk of making a mistake in doing so is, relative to her perspective, remote.

$S_E \Rightarrow J$  and  $J \Rightarrow S_E$  entail:

$$J = S_E \quad S \text{ is justified in believing } P \text{ on } E \text{ if and only if it is safe}_E \text{ for } S \text{ to believe } P \text{ on } E$$

One attractive feature of this view, to return to an earlier point, is that it accords with the idea that justification is non-factive. Another is that it vindicates plausible closure principles for justification, for example:

If  $S$  is justified in believing  $P$  on  $E$ , and  $S$  is justified in believing  $Q$  on  $E$ , and  $P$  and  $Q$  a priori entail  $R$ , then  $S$  is justified in believing  $R$  on  $E$ .<sup>25</sup>

This is a consequence of the fact, discussed above, that safety<sub>E</sub> is closed under a priori entailment. In this respect, the view of justification as safe<sub>E</sub> belief differs from the view of justification as probable belief. Without supplementation, the latter is incompatible

<sup>24</sup> Proponents of such principles include Fantl and McGrath (2009), Gerken (2011), Hawthorne and Stanley (2008), Littlejohn (2012, p. 199), Neta (2009), Smith (2016, p. 84), and Smithies (2012b). For criticism of certain ways of formulating such principles, see Whiting (2017, pp. 411–417).

<sup>25</sup> Of course, there are other candidate closure principles for justification. The suggestion is not that  $J = S_E$  vindicates all contenders.

with closure principles of the above sort.  $P$  and  $Q$  entail  $P \& Q$ , but the probabilities of  $P$  and  $Q$  will typically be higher than that of their conjunction. So, while  $P$  and  $Q$  might individually meet the probability threshold for justified belief,  $P \& Q$  might fall below it.

Another welcome feature of  $S_E = J$  is that, alongside  $K \Rightarrow S_E$ , it delivers  $K \Rightarrow J$ , that is, the view that knowledge entails justification.  $K \Rightarrow J$  is a plausible principle of considerable pedigree. (That said, it is possible to accept  $S_E = J$  while rejecting  $K \Rightarrow J$  by rejecting  $K \Rightarrow S_E$ .)

$J = S_E$  has a further non-trivial consequence. If  $J = S_E$  is true, the following principle is false:

$J \Rightarrow JJ$  If  $S$  is justified in believing  $P$  on  $E$ ,  $S$  is justified in believing on  $E$  that  $S$  is justified in believing  $P$  on  $E$ .<sup>26</sup>

As discussed above,  $safety_E$  does not iterate. Given  $J = S_E$ , it follows that justification does not iterate.

Just as one might think that whether it is  $safe_E$  to believe a proposition is context-sensitive, insofar as context determines whether an EPW is near enough to count as nearby, so one might think that whether a person is justified in believing a proposition is context-sensitive (cf. Fantl and McGrath 2009). Given  $J = S_E$ , justification might inherit the context-sensitivity of  $safety_E$ . (However, a proponent of  $J = S_E$  is not committed to this—she might deny that nearness is a contextually determined.)

Given  $E = K$ ,  $J = S_E$  is an externalist view of justification, in the sense that whether a person is justified in believing a proposition does not supervene upon her non-factive mental states. Just as my aim is not to adjudicate the dispute between internalists and externalists about evidence, it is not my aim to adjudicate the corresponding dispute about justification. While I assume  $E = K$ , one might accept  $J = S_E$  and reject  $E = K$ . Plugging in a more internalist account of evidence delivers a more internalist account of justification that is nonetheless  $safety_E$ -based. Accepting  $J = S_E$  neither settles nor pre-judges the issue over which internalists and externalists disagree.

One might worry that  $J = S_E$  is circular. It accounts for justification in terms of  $safety_E$ , which is explained in terms of epistemically possible worlds, which are explained in terms of what is knowable a priori. According to  $K \Rightarrow J$ , if a person knows a proposition, she is justified in believing it. So, one might think, I try to explain justification by appeal to notions which are themselves understood in terms of justification.<sup>27</sup>

First, if  $J = S_E$  is circular in this way, it remains an interesting claim with non-trivial implications. Second, it is a well-worn point that entailment relations need not

<sup>26</sup> Proponents of  $J \Rightarrow JJ$  include Chisholm (1986) and Smithies (2012a). To say that a person is justified in believing a proposition but not justified in believing that she is justified in believing that proposition is *not* to say that that person is justified in disbelieving that she is justified in believing the proposition or in suspending judgement as to whether she is (contra Smithies). Perhaps no higher-order attitude concerning the justificatory status of the first-order belief is justified (cf. Friedman 2013).

<sup>27</sup> One might worry that  $J = S_E$  is circular in a different way. It appeals to the notion of a person's evidence. Given  $E = K$ , her evidence is her knowledge. Given  $K \Rightarrow J$ , knowledge entails justification. My response to this version of the circularity charge is the same, though recall also that  $E = K$  is optional.

be explanatory relations. That a person knows a proposition might entail that she is justified in believing it, but knowledge might be explanatorily prior, not justification.<sup>28</sup>

Necessary truths present a rather different challenge to  $S_E \Rightarrow J$ . A familiar point is that it is trivially safe<sub>M</sub> for a person to believe necessary truths.<sup>29</sup> For any proposition that is necessarily true, that proposition is true in all nearby MPWs in which a person's evidence is true for the simple reason that that proposition is true in all MPWs. In view of this, one might wonder whether that it is trivially safe<sub>E</sub> to believe necessary truths. In turn, one might wonder whether, as a result,  $S_E \Rightarrow J$  predicts that, whatever her evidence, a person is justified in believing all necessary truths, which is implausible.

In response, I consider separately a posteriori necessary truths, for example, that water is  $H_2O$ , and a priori necessary truths, for example, that  $1+1=2$  (cf. Kripke 1980).

$S_E \Rightarrow J$  does not suggest that, whatever her evidence, a person is justified in believing that water is  $H_2O$  or, more generally, in believing a posteriori necessary truths. While it is true that water is  $H_2O$  in all MPWs, it is not true that water is  $H_2O$  in all EPWs. There are worlds which cannot be ruled out a priori in which water is not  $H_2O$ . Suppose that Lily knows that water is a colourless, odourless, transparent liquid, but knows nothing about its chemical constitution. The nearby EPWs in which her evidence is true—that is, in which water is a colourless, odourless, and transparent liquid—might include EPWs in which water is  $H_2O$  but also EPWs in which water is not  $H_2O$ . In that case, it is not safe<sub>E</sub> for Lily to believe on her evidence that water is  $H_2O$ . Hence,  $S_E \Rightarrow J$  does not predict that Lily is justified in believing this.

In contrast,  $S_E \Rightarrow J$  does predict that, whatever her evidence, a person is justified in believing that  $1+1=2$  or, more generally, in believing a priori necessary truths. Suppose that  $1+1=2$  is a priori for Ananya. In that case, Ananya can rule out a priori worlds in which  $1+1 \neq 2$ . Such worlds are not a priori coherent for Ananya. Hence, they are not epistemically possible for Ananya. So, whatever her evidence, it is true that  $1+1=2$  in all nearby EPWs in which Ananya's evidence is true, because it is true that  $1+1=2$  in all EPWs. So, it is safe<sub>E</sub> for Ananya to believe on her evidence that  $1+1=2$ . Given  $S_E \Rightarrow J$ , whatever her evidence, Ananya is justified in believing that  $1+1=2$ .

This is the right result. If Ananya is in a position to know a priori that  $1+1=2$ , and assuming  $K \Rightarrow J$ , Ananya is justified in believing that  $1+1=2$ . Of course, it does not follow from this that Ananya justifiably believes that  $1+1=2$ . The claim is that Ananya possesses *ex ante* justification, not *ex post* justification.

One might protest. Surely there can be a priori necessary truths, for example, arcane mathematical truths such as that  $0.999\dots = 1$ , that a person is not justified in believing.

At this point, the earlier observation about the variability of epistemic possibility is important. Ananya might be able to rule out a priori worlds in which  $1+1 \neq 2$ , but she might be unable to rule out a priori worlds in which  $0.999\dots \neq 1$ . In that case, worlds in which  $0.999\dots \neq 1$  are epistemically possible for her. In that case, in turn,

<sup>28</sup> One might view knowledge as explanatorily prior to justification without viewing knowledge, with Williamson (2000), as explanatorily fundamental.

<sup>29</sup> Pritchard (2012, p. 256) notes that necessary truths trivialise certain formulations of the safety condition on knowledge.

$S_E \Rightarrow J$  does not predict that, whatever her evidence, Ananya is justified in believing that  $0.999\dots \neq 1$ . More generally,  $J = S_E$  predicts that, if a proposition is a priori for a person, she is justified in believing it, whatever her evidence. But  $J = S_E$  is neutral on which propositions are a priori for a subject. That might vary from person to person.

This is not in tension with the earlier point that justification is closed under a priori entailment. While what Ananya knows might in some sense entail that  $0.999\dots \neq 1$ , the entailment is not a priori for her (though it might be for others).

## 6 Similarities and differences

As noted at the outset, the modal characterisation of justification I propose here bears some resemblance to existing proposals. In this penultimate section, I mention three prominent accounts of justification that appeal to safety or kindred notions. Since, once again, the aim of this paper is constructive, not critical, I will not challenge those accounts but indicate briefly how they differ from the one I develop above.

Pritchard (2005, pp. 175–176) introduces a notion of *reflective* safety, which one might capture as follows:

Safety<sub>R</sub> It is safer<sub>R</sub> for S to believe P if and only if, given what S can know by reflection alone, P is true in nearby MPWs

He then suggests that one might give an account of justification as safer belief.

Safety<sub>R</sub> differs importantly from Safety<sub>E</sub>. First, Pritchard appeals to MPWs not EPWs. Second, in Safety<sub>R</sub> nearness is relative to what is accessible via reflection, whereas in Safety<sub>E</sub> nearness is relative to a person's evidence, which need not be reflectively accessible. Indeed, I argued above that an account of justification in terms of safety<sub>E</sub> need not satisfy a certain accessibility constraint, namely,  $J \Rightarrow JJ$ . Third, I suggested that safety<sub>E</sub>, hence, justification, is a condition on knowledge. According to Pritchard, safety<sub>R</sub> is a condition only on (what he calls) *reflective* knowledge. Finally, Pritchard suggests that one might appeal to Safety<sub>R</sub> to capture an internalist notion of justification. In contrast, one can use Safety<sub>E</sub> to deliver both internalist and externalist notions of justification, depending on the conception of evidence one plugs in.

According to Smith (2016), a person is justified in believing a proposition on her evidence if and only if, in *normal* worlds in which her evidence obtains, that proposition is true. Though he does not develop the point, Smith suggests in passing that normal worlds might be epistemically, as opposed to metaphysically, possible (2016, p. 115). To explain the notion of normal worlds, Smith offers an analogy with ‘simplified models of a complex phenomenon which abstract away from [...] certain factors in order to expose underlying patterns and mechanisms’ (2016, pp. 113–114). A normal world, he suggests, is an ‘idealised model *writ large*’. It is not clear how to understand this—possible worlds are complete and fully determinate whereas idealisations are incomplete and partially indeterminate. In any event,  $J = S_E$  makes no appeal to normal worlds so understood. Smith also associates normalcy with explanation: ‘normal conditions require *less explanation* than abnormal conditions do’ (2016, p. 39). On this account, a person is justified in believing a proposition on her evidence just in case a circumstance in which her evidence is true but the proposition is false demands

more explanation than one in which her evidence is true and the proposition is true. When explanation is called for is a tricky issue which  $J = S_E$  avoids, since it makes no appeal to the notion of explanation or of requiring it. This is not the place to attempt a critical assessment of Smith's account.<sup>30</sup> The point is just that it involves notions which do not figure in the account of justification I advance.

According to Wedgwood (2002), a person is justified in believing a proposition via a certain method just in case she is justified in believing that it is safe<sub>M</sub> to do so, that is, just in case she is justified in believing that following that method does not lead to false beliefs in nearby MPWs.  $J = S_E$  differs from Wedgwood's proposal in several ways. First, it does not appeal to methods of belief-formation. Second, it refers to EPWs, not MPWs. Third, as discussed above,  $J = S_E$  predicts that a person can be justified in believing a proposition but not justified in believing that so believing is safe in the relevant sense.<sup>31</sup>

## 7 Conclusion

I introduced a notion of safety understood in terms, not of metaphysical possibilities, but of epistemic possibilities. I then put that notion to work in articulating and exploring substantive epistemological proposals. I tried to present those proposals in a plausible light. However, as mentioned at the outset, the main aim is to add a new tool to the epistemologist's tool-kit, a tool which might be used to argue for or, indeed, against those proposals.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

Baumann, P. (2008). Is knowledge safe? *American Philosophical Quarterly*, 45(1), 19–30.

Bird, A. (2007). Justified judging. *Philosophy and Phenomenological Research*, 74(1), 81–110.

Bogardus, T. (2014). Knowledge under threat. *Philosophy and Phenomenological Research*, 88(2), 289–313.

Chalmers, D. (2011). The nature of epistemic space. In A. Egan & B. Weatherson (Eds.), *Epistemic modality* (pp. 60–107). Oxford: Oxford University Press.

Chisholm, R. (1986). The place of epistemic justification. *Philosophical Topics*, 14(1), 85–92.

Comesafá, J. (2005). Unsafe knowledge. *Synthese*, 146(3), 393–402.

Dougherty, T. (2011). In defense of propositionalism about evidence. In T. Dougherty (Ed.), *Evidentialism and its discontents* (pp. 226–232). Oxford: Oxford University Press.

Dretske, F. (2005). The case against closure. In M. Steup & E. Sosa (Eds.), *Contemporary debates in epistemology* (pp. 13–25). Oxford: Blackwell.

Fantl, J., & McGrath, M. (2009). *Knowledge in an uncertain world*. Oxford: Oxford University Press.

Friedman, J. (2013). Suspended judgement. *Philosophical Studies*, 162(2), 165–181.

Gerken, M. (2011). Warrant and action. *Synthese*, 178(3), 529–547.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23(6), 121–123.

<sup>30</sup> For this, see McGlynn (2012).

<sup>31</sup> For critical discussion of Wedgwood's account, see Glüer and Wikforss (2013) and Whiting (2013, pp. 197–201).

Glüer, K., & Wikforss, Å. (2013). Against belief normativity. In T. Chan (Ed.), *The aim of belief* (pp. 80–99). Oxford: Oxford University Press.

Goldman, A. (2009). Williamson on knowledge and evidence. In P. Greenough & D. Pritchard (Eds.), *Williamson on knowledge* (pp. 73–91). Oxford: Oxford University Press.

Greco, J. (2016). Knowledge, virtue, and safety. In M. Vargas (Ed.), *Performance epistemology* (pp. 51–61). Oxford: Oxford University Press.

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

Hawthorne, J., & Stanley, J. (2008). Knowledge and action. *Journal of Philosophy*, 105(10), 571–590.

Humberstone, I. L. (1981). From worlds to possibilities. *Journal of Philosophical Logic*, 10(3), 313–339.

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

Kripke, S. (1980). *Naming and necessity*. Oxford: Blackwell.

Kyburg, H. (1961). *Probability and the logic of rational belief*. Middletown, CT: Wesleyan University Press.

Lackey, J. (2006). Pritchard's epistemic luck. *Philosophical Quarterly*, 56(223), 284–289.

Lasonen-Aarnio, M. (2010). Unreasonable knowledge. *Philosophical Perspectives*, 24(1), 1–21.

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

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs*, 13(4), 455–476.

Littlejohn, C. (2012). *Justification and the truth-connection*. Cambridge: Cambridge University Press.

Luper, S. (2016). Epistemic closure. In E. Zalta (Ed.), *Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/spr2016/entries/closure-epistemic/>.

McGlynn, A. (2012). Justification as 'would-be' knowledge. *Episteme*, 9(4), 361–376.

McGlynn, A. (2014). *Knowledge first?*. Basingstoke: Palgrave.

Mills, E. (2012). Lotteries, quasi-lotteries, and scepticism. *Australasian Journal of Philosophy*, 90(2), 335–352.

Mitova, V. (2015). Truthy psychologism about evidence. *Philosophical Studies*, 172(4), 1105–1126.

Nelkin, D. (2000). The lottery paradox, knowledge, and rationality. *Philosophical Review*, 109(3), 373–409.

Neta, R. (2008). What evidence do you have? *British Journal for the Philosophy of Science*, 59(1), 8–19.

Neta, R. (2009). Treating something as a reason for acting. *Noûs*, 43(4), 684–699.

Neta, R., & Rohrbaugh, G. (2004). Luminosity and the safety of knowledge. *Pacific Philosophical Quarterly*, 85(4), 396–406.

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

Pritchard, D. (2012). Anti-luck virtue epistemology. *Journal of Philosophy*, 109(3), 247–279.

Pritchard, D. (2015). Risk. *Metaphilosophy*, 46(3), 436–461.

Roush, S. (2005). *Tracking truth*. Oxford: Oxford University Press.

Rumfitt, I. (2015). *The boundary stones of thought*. Oxford: Oxford University Press.

Ryan, S. (1996). The epistemic virtues of consistency. *Synthese*, 109(2), 121–141.

Sainsbury, R. M. (1997). Easy possibilities. *Philosophy and Phenomenological Research*, 57(4), 907–919.

Smith, M. (2016). *Between probability and certainty*. Oxford: Oxford University Press.

Smithies, D. (2012a). Moore's paradox and the accessibility of justification. *Philosophy and Phenomenological Research*, 85(2), 273–300.

Smithies, D. (2012b). The normative role of knowledge. *Noûs*, 46(2), 265–288.

Sosa, E. (1999). How must knowledge be modally related to what is known? *Philosophical Topics*, 26(1/2), 373–384.

Sosa, E. (2007). *A virtue epistemology*. Oxford: Oxford University Press.

Srinivasan, A. (2015). Normativity without Cartesian privilege. *Philosophical Issues*, 25(1), 273–299.

Steglich-Petersen, A. (2013). Truth as the aim of epistemic justification. In T. Chan (Ed.), *The aim of belief* (pp. 204–226). Oxford: Oxford University Press.

Sutton, J. (2007). *Without justification*. Cambridge, MA: MIT Press.

Vogel, J. (1990). Are there counterexamples to the closure principle? In M. Roth & G. Ross (Eds.), *Doubting* (pp. 13–27). Dordrecht: Kluwer.

Weatherspoon, B. (2001). Indicative and subjunctive conditionals. *Philosophical Quarterly*, 51(203), 200–216.

Wedgwood, R. (2002). The aim of belief. *Philosophical Perspectives*, 16, 267–297.

Whiting, D. (2013). Nothing but the truth: On the norms and aims of belief. In T. Chan (Ed.), *The aim of belief* (pp. 184–203). Oxford: Oxford University Press.

Whiting, D. (2017). Against second-order reasons. *Noûs*, 51(2), 398–420.

Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.

Williamson, T. (forthcoming). Justifications, excuses, and sceptical scenarios. In J. Dutant & F. Dorsch (Eds.), *The new evil demon problem*. Oxford: Oxford University Press.

Yamada, M. (2011). Getting it right by accident. *Philosophy and Phenomenological Research*, 83(1), 72–105.