

Safety and Future Dependence

BIN ZHAO

Institute of Foreign Philosophy, Peking University, Beijing, PR China

According to the safety account of knowledge, one knows that p only if one's belief in p could not easily have been false. In the literature, most objections to the safety account rely on intuition of knowledge that could be easily denied by the safety theorists. In this paper, an objection to the safety account which does not make use of such intuition is raised. It is argued that either there are instances of unsafe knowledge or the safety account has an implausible implication that one's epistemic status might depend on what happens in the future.

Keywords: Future dependence; knowledge; safety.

1. Introduction

In many Gettier cases, though the subject holds a true belief, her belief could easily have been false. This motivates the safety account of knowledge, according to which, S knows that p only if S 's belief in p is safe, that is, only if S could not easily have falsely believed p . A belief that satisfies the safety condition counts as knowledge unless it exhibits some non-modal shortcomings that would deprive it of the status of knowledge.

Like other accounts of knowledge in the literature, the safety account is susceptible to a variety of objections. It has been argued that, in some putative counterexamples, one could know some proposition though one could easily have falsely believed the proposition. We thus have instances of unsafe knowledge. However, the safety theorists are ready to deny their opponents' intuition of knowledge. Thus, such a debate often leads to a stalemate.

In this paper, an objection to the safety condition which does not rely on such intuition is raised. It is argued that, if the safety theorists accept the intuition of knowledge, then there are instances of unsafe knowledge. But, if the safety theorists deny the intuition of knowledge, then the safety account has an implausible implication that one's epis-

temic status might depend on what happens in the future. Therefore, no matter whether the safety theorists deny the intuition of knowledge or not, the safety account is problematic.

2. *The Safety Account of Knowledge*

Consider a scenario where one looks at a reliable clock under normal lighting conditions and thus forms a true belief that *it is now 12:00*. Since the clock is reliable and the lighting conditions are normal, the belief counts as knowledge. Consider another scenario where one looks at a clock that stops 12 hours earlier under normal lighting conditions and thus forms a true belief that *it is now 12:00*. The belief is true as a matter of luck and does not count as knowledge. After all, it is true because the dose of bad luck that the clock had stopped is canceled out by the good luck that the clock had stopped 12 hours ago.

Why is it the case that one knows the correct time in the first scenario but not in the second scenario though in both scenarios one has a true belief about the time? The safety account of knowledge offers a simple answer to this question. According to this account, *S* knows that *p* only if *S*'s belief in *p* is safe, that is, only if *S* could not easily have falsely believed *p*. To put it formally,

Safety: *S*'s belief in *p*, formed on belief-formation method *M*, is safe, if and only if, in all nearby possible worlds where *S* forms a belief in *p* on *M*, *p* is true.¹

This makes us consider whether *p* is true in nearby possible worlds where *S* believes that *p*. If *p* is false in some of these possible worlds, then *S*'s belief in *p* is not safe, and *S* does not know that *p*. If *p* is true in all these possible worlds, then *S*'s belief in *p* is safe, and *S* knows that

¹ The safety condition is usually relativized to the belief-formation methods to avoid putative counterexamples such as Alfano's (2009) REDWOOD, Goldman's (1976) DACHSHUND, and Nozick's (1981) GRANDMOTHER. Because the argument here does not hinge on whether the conditions should be thus relativized, I shall leave this point aside. For discussions of the individuation of belief-formation methods, see Alfano (2009), Becker (2008, 2012), and Zhao (2022a, 2022d, 2024a, 2024b, 2025, forthcoming). Proponents of the safety condition on knowledge include Ball (2016), Beddor and Pavese (2020), Dutant (2010), Grundmann (2020), Hirvelä (2019), Luper-Foy (1984), Manley (2007), Peet and Pitcovski (2018), Pritchard (2005, 2009), Sainsbury (1997), Sosa (1999a, 1999b, 2015), and Williamson (2000).

It has also been argued that the safety condition should be globalized to a set of propositions rather than the target proposition to account for why beliefs in necessary truths could still be true as a matter of luck. See Pritchard (2009), Sosa (2015), and Williamson (2000). Here is the globalized version of the safety condition:

Globalized Safety: *S*'s belief that *p*, formed on belief-formation method *M*, is safe, if and only if, in all nearby possible worlds where *S* forms a belief on method *M*, the belief is true.

Because the argument in this paper does not depend on which version of the safety condition we opt for, I shall not delve into the distinction here. For discussions of the globalized version of the safety condition, see Becker (2006), Hirvelä (2019), and Zhao (2022a, 2022b, 2022c).

p unless it exhibits some non-modal shortcomings that would deprive it of the status of knowledge. One knows the correct time in the first scenario because his belief is safe. In all nearby possible worlds where he believes that *it is now 12:00* via the clock, it is 12:00. One does not know the correct time in the second scenario because his belief is unsafe. There are some nearby possible worlds where he looks at the clock one minute earlier and believes that *it is now 12:00* via the clock, while the time is 11:59.

In a word, the safety condition is supposed to be a necessary condition on knowledge that helps to eliminate epistemic luck. In virtue of this, the safety account of knowledge does an excellent job handling cases where the subject's belief is true as a matter of luck, e.g., the Gettier cases.

3. *Does Knowledge Require Safety?*

In the literature, a variety of objections to the safety account of knowledge have been raised. The objections usually take the following structure: an example is constructed. It is then argued that the subject in the example knows some proposition though she could easily have falsely believed the proposition. We thus have an instance of unsafe knowledge (Baumann 2008; Bogardus 2014; Comesaña 2005; Neta and Rohrbaugh 2004; Yamada 2011). However, the safety theorists could push back and deny their proponents' intuition of knowledge (Pritchard 2009). In that case, the putative counterexamples to the safety account are discharged. Thus, the debate between them often leads to a stalemate.

In this section, I argue that the safety theorists' denial of their proponents' intuition of knowledge is not without some serious costs. To be specific, the safety theorists are faced with a dilemma here: if they choose not to deny the intuition of knowledge, then there are instances of unsafe knowledge. But, if they choose to deny the intuition of knowledge, then the safety account has an implausible implication that one's epistemic status might depend on what happens in the future.

Here is the example I shall consider,

Dead President. Consider . . . the situation of a generally well-informed citizen N. N. who has not yet heard the news from the theatre where Lincoln has just been assassinated. Since Lincoln is dead, he is no longer President, so N. N. no longer knows that Lincoln is President (knowing is factive). However, N. N. is in no position to know that anything is amiss. He continues reasonably to believe that Lincoln is President . . . N. N. does not know that Lincoln is President. (Williamson 2000: 23).

Let t_0, t_1, \dots, t_n be a series of times at one millisecond intervals in this case. To be specific, t_0 is the starting point when N. N. begins to believe that Lincoln is President; t_{n-1} is one millisecond before Lincoln dies; t_n is the time Lincoln dies. Let L be the proposition that Lincoln is President; B(L) the condition that N. N. believes that L; and K(L) the

condition that N. N. knows that L. It is reasonable to say that we have $K(L)$ at t_0 .

Let us consider the possible world w where Lincoln is assassinated one millisecond earlier at t_{n-1} rather than t_n . In w , N. N. falsely believes that L on the same belief-formation method as that in the actual world, i.e., his memory, at t_{n-1} . Is w a nearby possible world? I think so. There is at least no difference in the belief-formation method between the actual world and w . In both worlds, the beliefs are formed on the same belief-formation method, i.e., his memory. The most salient difference between them is that L has different truth-values: L is true at t_{n-1} in the actual world while false at t_{n-1} in w . However, this difference should not bother us. If a possible world counts as a nearby possible world only if the target proposition has the same truth-value as that in the actual world, then the safety condition would be satisfied for any true belief and thus the condition would be of no use in eliminating luckily true belief from the realm of knowledge. One might appeal to other differences between the actual world and w to exclude w from the realm of nearby possible worlds. However, for any such difference, we can tweak the example such that the actual world and w are the same or, at least, similar in these respects. For instance, if one thinks that Lincoln's being assassinated one millisecond before t_n makes w very different from the actual world where Lincoln dies at t_n , we can reformulate the case such that Lincoln was assassinated one microsecond, one nanosecond, or one picosecond before t_n in w .

To clarify, this is not to say that temporal proximity suffices for modal proximity. In cases where temporal proximity is not accompanied by violations of law or large-scale mismatch of particular facts, temporal proximity leads to modal proximity. For instance, w is a nearby possible world because the only difference between w and the actual world is that John Wilkes Booth assassinated Lincoln slightly more efficiently in w . In contrast, in cases where temporal proximity is accompanied by violations of law or large-scale mismatches of particular facts, temporal proximity does not lead to modal proximity. For instance, there will be a total Solar eclipse on September 3, 2081, at 9:07:31 (Terrestrial time, Central Europe) in the actual world. However, the possible world where the total Solar eclipse happens on September 3, 2081, at 9:07:32 (Terrestrial time, Central Europe) is not a nearby possible world because there are either violations of law or large-scale mismatches of particular facts in that possible world. In sum, we have good reasons to think that w is a nearby possible world.

Anthony Brueckner and M. Oreste Fiocco (2002) think that, in Dead President, we intuitively have $K(L)$ at t_{n-1} in the actual world. The problem for the safety theorists is that there is a nearby possible world, i.e., w , where N. N. falsely believes that L at t_{n-1} . This is a counterexample for the safety account as the subject knows some proposition though he could easily have falsely believed the proposition (Brueckner and Fiocco 2002: 288).

In sum, nothing prevents us from counting w as a nearby possible world. As a consequence, we are forced to give up the safety account. Is there a way out for the safety theorists? The safety theorists may bite the bullet and deny $K(L)$ at t_{n-1} in the actual world. If the only reason for accepting $K(L)$ at t_{n-1} in the actual world is that it seems intuitive, then the safety theorists are ready to deny such intuition as this is what they usually do regarding other putative counterexamples.

This is the stance taken by Williamson. As he writes,

[safety] need not be determined by local properties of the basis. For instance, if someone continues over some time to believe that Lincoln is President on the basis of automatic updating, without receiving further confirmation, the [safety] of the basis may depend on whether Lincoln is about to be assassinated. (Williamson 2008: 280)

That is to say, N. N.'s belief that L at t_{n-1} is not safe because there is a nearby possible world, i.e., w , where Lincoln is assassinated one millisecond earlier at t_{n-1} such that he falsely believes that L on a similar basis at t_{n-1} . In virtue of its being unsafe, N. N.'s belief that L at t_{n-1} does not count as knowledge. Contrary to the intuition of knowledge from the anti-safety theorists, we do not have $K(L)$ at t_{n-1} in the actual world. Therefore, the putative counterexample for the safety condition is dismissed.

Once again, we seem to have a stalemate here. On the one hand, the anti-safety theorists argue that, because we have $K(L)$ in at t_{n-1} , safety is not a necessary condition for knowledge. On the other hand, the safety theorists argue that, because safety is a necessary condition for knowledge, we do not have $K(L)$ in at t_{n-1} . The safety theorists' argument is merely a modus tollens of the anti-safety theorists' modus ponens. If so, then it seems that Dead President is not in a better place than other putative counterexamples in the literature to resolve the debate. Nonetheless, I don't think this is a dead-end for the discussion because the safety theorists' move is not without some serious cost. To explicate this point, let us examine the anti-safety and the safety theorists' stances in turn.

If the anti-safety theorists' story is true, then N. N. knows that L until t_n . Why doesn't N. N. know that L at t_n ? The anti-safety theorists' answer is as follows: because knowledge is factive, one cannot know something if it is not factive. That is to say, L 's being false at t_n makes it the case that $\sim K(L)$ at t_n .

In contrast, suppose that the safety theorists' story is true, that is, N. N. does not know that L at t_{n-1} . What makes it the case that N. N. does not know that L at t_{n-1} while he knows that L at t_0 ? The safety theorists' answer is as follows: because knowledge is safe, one cannot know something if one's belief is not safe. His belief at t_0 is safe but it becomes unsafe at some point before t_{n-1} .

Here comes another question: why is his belief that L at t_{n-1} unsafe? The safety theorists' answer should be as follows: because there is a nearby possible world, i.e., w , where Lincoln is assassinated one mil-

lisecond earlier at t_{n-1} such that he falsely believes that L on the same belief-formation method as that in the actual world, i.e., his memory, at t_{n-1} . w 's being close to the actual world is in virtue of N.N.'s falsely believing that L one millisecond later at t_n in the actual world. In addition, N.N.'s falsely believing that L one millisecond later at t_n in the actual world is in virtue of L's being false at t_n . That is to say, L's being false at t_n , at least partly, makes B(L) unsafe at t_{n-1} and thus $\sim K(L)$ at t_{n-1} .

If the story unfolds somewhat differently such that Lincoln is assassinated at a later time, say t_{n+1} , in the actual world, then it turns out that N.N.'s belief at t_{n-1} is safe and thus counts as knowledge because in all nearby possible worlds where he believes that L, his belief is true. After all, given the dissimilarity between the actual world and w , w would not be among the nearby possible worlds.² That is to say, L's being true at t_n , at least partly, makes B(L) safe at t_{n-1} and thus K(L) at t_{n-1} .

In sum, the safety theorists' move commits them to the idea that what happens (e.g., L's being true or false) at a later time (e.g., t_n) makes a difference to one's epistemic status (e.g., N.N.'s knowing or not knowing that L) at an earlier time (e.g., t_{n-1}). For the sake of simplicity, this implication might be called "future dependence." I can understand that L's being false at t_n makes it the case that $\sim K(L)$ at t_{n-1} since knowledge is factive. Nonetheless, how could L's being false, at least partly, at t_n make it the case that $\sim K(L)$ at t_{n-1} , and L's being true, at least partly, makes it the case that K(L) at t_{n-1} ? After all, t_n is later than t_{n-1} ! It is no surprise for epistemic externalists to argue that one's epistemic status might depend on external factors not accessible to the subject such as the reliability of the belief-formation methods and the environment. Nonetheless, the implication that one's epistemic status might depend on what happens in the future is still something hard to swallow.

One might think that it is no surprise that what happens at a later time could make a difference to one's epistemic status. For instance, suppose that we are assessing the reliability of a belief-forming process in its early stages. There has been only one instance of the process. In this case, it seems reasonable to say that its reliability depends on the outputs of the process in the future. Regarding this response, I would say that it is true that the outputs of the process in the future would help us to know more about its reliability. Nonetheless, whether the process is reliable or not and to what extent the process is reliable has already been determined by the facts about the belief-forming process as well as the environment in its early stages. Similarly, it is one thing to say that the head-to-tail ratio in the long run could help us to know more about whether a coin is a fair one; while it is another thing to say that whether the coin is fair or not depends on the head to tail ratio, in the long run, because it is intuitive to say that whether the coin is

² You may choose α_{n+j} where $j \geq 2$ if you think that the interval of two milliseconds fails to render one's belief at t_{n-1} safe.

fair or not has already been determined by the facts about the coin as well as the environment before the coin has been tossed. Thus, it is still unclear that one's epistemic status depends on what happens in the future.³

4. Conclusion

When faced with an example such as Dead President, accepting and denying the intuition of knowledge in the example constitute two horns of a dilemma for the safety account of knowledge. If the safety theorists accept the intuition of knowledge, then the safety account is outright false. If the safety theorists deny the intuition of knowledge, then the safety account makes one's epistemic status depend on what happens in the future. In sum, the safety account is problematic.^{4, 5}

References

- Alfano, M. 2009. "Sensitivity Theory and the Individuation of Belief-Formation Methods." *Erkenntnis* 70 (2): 271–281.
- Ball, B. 2016. "Knowledge, Safety, and Questions." *Filosofia Unisinos* 17 (1): 58–62.
- Baumann, P. 2008. "Is Knowledge Safe?" *American Philosophical Quarterly* 45: 19–30.
- Becker, K. 2006. "Reliabilism and Safety." *Metaphilosophy* 37 (5): 691–704.
- . 2008. "Epistemic Luck and the Generality Problem." *Philosophical Studies* 139 (3): 353–366.
- . 2009. "Margins for Error and Sensitivity: What Nozick Might Have Said." *Acta Analytica* 24 (1): 17–31.
- . 2012. "Methods and How to Individuate Them." In K. Becker and T. Black (eds.), *The Sensitivity Principle in Epistemology*. New York: Cambridge University Press, 81–97.

³ It could be conceded that one's epistemic status depends on what happens in the future in cases of future knowledge. Given that knowledge entails truth, it is no surprise that whether one obtains future knowledge or not depends on what happens in the future. For instance, whether I know that *it rains tomorrow* depends on whether it rains tomorrow. If it rains tomorrow, then I could know that *it rains tomorrow*; if it does not rain tomorrow, then I couldn't know that *it rains tomorrow*. However, this kind of future dependence in cases of future knowledge still does not establish the kind of future dependence discussed in this paper. It is one thing to say that one's epistemic status depends on what happens in the future in cases of future knowledge; while it is another thing to say that one's epistemic status depends on what happens in the future in cases of ordinary knowledge even if the truth of the target proposition has been determined, the reliability of the belief-forming process has been determined, the fact about whether the subject obtains good evidence has been determined, the fact about whether subject is in an epistemically friendly environment has been determined, etc.

⁴ Thanks to Brian T. Miller, Duncan Pritchard, and three anonymous referees for *Croatian Journal of Philosophy* for helpful comments or discussions.

⁵ Funding information: This research was supported by National Social Science Fund of China (Grant Number: 24CZX044).

- Beddor, B. and Pavese, C. 2020. "Modal Virtue Epistemology." *Philosophy and Phenomenological Research* 101 (1): 61–79.
- Bogardus, T. 2014. "Knowledge Under Threat." *Philosophy and Phenomenological Research* 88 (2): 289–313.
- Brueckner, A. and Fiocco, M. O. 2002. "Williamson's Anti-Luminosity Argument." *Philosophical Studies* 110 (3): 285–293.
- Comesaña, J. 2005. "Unsafe Knowledge." *Synthese* 146: 395–404.
- Dutant, J. 2010. "Two Notions Of Safety." *Swiss Philosophical Preprints* 87: 1–19.
- Goldman, A. 1976. "Discrimination and Perceptual Knowledge." *Journal of Philosophy* 73 (20): 771–791.
- Grundmann, T. 2020. "Saving Safety From Counterexamples." *Synthese* 197: 5161–5185.
- Hirvelä, J. 2019. "Global Safety: How to Deal with Necessary Truths." *Synthese* 196 (3): 1167–1186.
- Luper-Foy, S. 1984. "The Epistemic Predicament: Knowledge, Nozickian Tracking, and Scepticism." *Australasian Journal of Philosophy* 62 (1): 26–49.
- Manley, D. 2007. "Safety, Content, Apriority, Self-Knowledge." *Journal of Philosophy* 104 (8): 403–423.
- Neta, R. and Rohrbaugh, G. 2004. "Luminosity and the Safety of Knowledge." *Pacific Philosophical Quarterly* 85 (4): 396–406.
- Nozick, R. (1981). *Philosophical Explanations*. Oxford: Oxford University Press.
- Peet, A. and Pitcovski, E. 2018. "Normal Knowledge: Toward an Explanation-Based Theory of Knowledge." *Journal of Philosophy* 115 (3): 141–157.
- Pritchard, D. 2005. *Epistemic Luck*. Oxford: Oxford University Press.
- . 2009. "Safety-Based Epistemology: Whither Now?" *Journal of Philosophical Research* 34: 33–45.
- Sainsbury, R. M. 1997. "Easy Possibilities." *Philosophy and Phenomenological Research* 57 (4): 907–919.
- Sosa, E. 1999a. "How Must Knowledge Be Modally Related to What Is Known?" *Philosophical Topics* 26 (1&2): 373–384.
- . 1999b. "How to Defeat Opposition to Moore." *Philosophical Perspectives* 13: 141–154.
- . 2015. *Judgment and Agency*. Oxford: Oxford University Press.
- Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
- . 2008. "Why Epistemology Can't be Operationalized." In Q. Smith (ed.), *Epistemology: New Essays*. Oxford: Oxford University Press, 277–300.
- Yamada, M. 2011. "Getting It Right By Accident." *Philosophy and Phenomenological Research* 83 (1): 72–105.
- Zhao, B. 2022a. "A Dilemma for Globalized Safety." *Acta Analytica* 37: 249–261.
- . 2022b. "Epistemic Closure, Necessary Truths, and Safety." *American Philosophical Quarterly* 59 (4): 391–401.
- . 2022c. "Knowledge from Falsehood, Ignorance of Necessary Truths, and Safety." *Philosophia* 50: 833–845.

- . 2022d. “Sensitivity, Safety, and Epistemic Closure.” *International Journal of Philosophical Studies* 30 (1): 56–71.
- . 2024a. “On Relativizing the Sensitivity Condition to Belief-Formation Methods.” *American Philosophical Quarterly* 61 (2): 165–175.
- . 2024b. “On Translating the Sensitivity Condition to the Possible Worlds Idiom in Different Ways.” *American Philosophical Quarterly* 61 (1): 87–98.
- . 2025. “On Mentioning Belief-Formation Methods in Sensitivity Subjunctives.” *Ergo* 12: 232–246.
- . forthcoming. “On the Tracking Account of Inferential Knowledge.” *Inquiry*.