Safety, Closure, and Extended Methods[1]

Simon Goldstein and John Hawthorne

Abstract
Recent research has identified a tension between the Safety principle
that knowledge is belief without risk of error, and the Closure principle
that knowledge is preserved by competent deduction. Timothy
Williamson reconciles Safety and Closure by proposing that when an
agent deduces a conclusion from some premises, the agent's method for
believing the conclusion includes their method for believing each
premise. We argue that this theory is untenable because it implies
problematically easy epistemic access to one's methods. Several
possible solutions are explored and rejected.

A rich tradition argues that knowledge is belief that is safe from error. Here is a fairly straightforward articulation of the safety approach, one that will serve as a helpful starting point:

Basic Safety S knows p if and only if S could not easily have falsely believed p using their actual method of belief formation.[2]

One current controversy is whether Basic Safety -- or some close variant -- is compatible with the principle that knowledge is closed under competent deduction.

Closure If S knows some premises and competently deduces a conclusion from them, then the agent knows the conclusion.[3]

---

[2] For defenses of principles closely related to Basic Safety, see among others: Sosa 1999, Williamson 2000, Manley 2007, Williamson 2009a, Pritchard 2005, and Pritchard 2012. Williamson 2009a, Beddor and Pavese 2018, and Grundmann 2018 embrace the necessity and sufficiency of safety for knowledge, while others only accept necessity. We shall later make trouble for the strategy of avoiding the challenges of this paper by denying the sufficiency of Safety, and by considering more sophisticated versions of Safety. Throughout, we appeal to a method relative version of Safety. In appealing to method relativity, we join a long tradition dating from Nozick 1981 (who used it to bolster an account of knowledge tied to certain counterfactuals). To see why a Safety theorist should deploy method relativity, consider Nozick's Grandmother case (p. 179). Grandmother visits her grandson and sees that he is healthy. Imagine that the grandson could easily have been unwell. But if he had been unwell, the family would have avoided her visit and just told her that her grandson was well anyways. Grandmother could easily have believed falsely that her grandson was healthy. But she could not have done so while looking at her grandson, and this was her actual method of forming her belief. This example shows that not every case in which an agent could easily have believed falsely produces ignorance. Basic safety controls for this, since on that view, ignorance only results when the agent believes falsely using the same method as they actually employed. Later on, we will also show that our central arguments generalize to various alternative formulations of Safety that appeal to similar methods and counterpart propositions.

[3] See Williamson 2000 and Hawthorne 2004 for similar formulations of the Closure principle. Here we operate with a conception of competent deduction in which the agent retains knowledge of the premises throughout, and forms a belief rather than some weaker attitude towards the conclusion. Some versions of our argument apply not only to multi-premise but also to single-premise closure, as we show below.

Kvanvig 2004, Murphy 2005, 2006, Alspector-Kelly 2011, and others argue that Basic Safety is incompatible with Closure. An agent can believe a premise using a reliable method and deduce a conclusion, even though they easily could have believed the premise using an unreliable method while deducing the same conclusion. In such a case, these authors argue that Basic Safety predicts the agent knows the premise without knowing the conclusion.

To preserve Closure, Williamson 2009a offers a sophisticated conception of belief forming methods:

Extended Methods If S competently deduces a conclusion from some premises, then S's method for believing the conclusion includes as an essential part S's method for believing each premise.[4]

This paper argues against Extended Methods, by showing that Extended Methods and Safety imply implausibly easy paths to knowledge of one's methods. Section 1 reviews Safety and Closure. Section 2 presents our argument. Section 3 explores responses to our argument. Section 4 relates the discussion to the choice between internalist and externalist individuation of methods. Section 5 concludes. (In an appendix, we consider an alternative to Extended Methods defended by Das and Salow 2018 and Schulz 2020.) We shall in a few places indicate possible escape routes for one who wishes to combine Closure and Safety. But none of them are pain-free and it is instructive to see what kinds of costs one may have to pay to secure that combination.

1. Safety and Closure

To see why Basic Safety threatens Closure, consider:

Field Sheba knows there is a sheep in the field, having seen it. She infers that there is a sheep or a cow in the field. The sheep almost escaped right before Sheba saw it, and the field contains no cows. If the sheep had escaped, someone would have prevented Sheba from looking at the field and simply told her that there was a sheep in it.[5]

In Field, Sheba knows there is a sheep by perception. But she could have easily believed the same thing falsely by testimony. Using her knowledge, Sheba deduces there is a sheep or a cow in the field. Plausibly, Sheba thereby knows that there is a sheep or cow in the field. But she could have easily deduced this conclusion when falsely believing the premise, and so she could have easily falsely believed the conclusion.

Whether Sheba knows the conclusion depends on her method for believing it. One natural conception of her method is that Sheba formed a belief that there is a sheep or cow by deducing it from the believed premise that there is a sheep. She would count as employing that method whenever she believes there is a sheep and deduces that there is a sheep or cow. Sheba therefore would fail to know that there is a sheep or a cow, since she could have believed this falsely using the same method of belief.

---

[4] See Goldman 1979 for a defense of something like Extended Methods in the setting of reliabilism about epistemic justification, in particular regarding the status of conditionally reliable processes.
[5] See Das and Salow 2018. See Alspector-Kelly 2011 and Grundmann 2018 for a similar case, Parking Lot.

In response to this kind of challenge, Williamson 2009a, 2009b offers an extended conception of methods. An agent's method for believing a conclusion essentially includes their method for believing the premises:

Extended Methods If S competently deduces a conclusion from some premises, then S's method for believing the conclusion includes as an essential part S's method for believing each premise.[6]

Throughout, the notion of one method containing another as an essential part is shorthand for a modal profile. It means that the agent could not have believed the conclusion using the same method unless they also believed each premise using the same method.

Applied to Field, Extended Methods predicts that Sheba knows the conclusion that there is a sheep or a cow in the field. Sheba's method for believing that there is a sheep or a cow is not merely deducing this conclusion from the premise that there is a sheep. Rather, at a first pass her method for believing that there is a sheep or a cow is believing that there is a sheep using perception, and then deducing that there is a sheep or a cow.[7]

Extended Methods validates Closure. Suppose an agent knows a premise. Then they could not easily have believed this premise falsely using their actual method for believing the premise. Now suppose they competently deduce a conclusion from the premise. Then they could not easily have believed the conclusion falsely using their actual method for believing the conclusion. Their method for believing the conclusion includes their method for believing the premise; since this input method could not have gone wrong, neither could their belief in the conclusion.[8] [9] [10]

2. The Boomerang Argument

We now show that Extended Methods and Basic Safety imply the following striking generalization.

---

[6] Williamson 2009b, p. 326. See Das and Salow 2018, Schulz 2020, and Luper 2020, Section 2.1 for further discussion.

[7] The agent's method for believing that there is a sheep in the field is more fine grained than merely using perception. If it was just perception, the same problem could arise in a case where the agent could employ more or less reliable perceptual mechanisms. We address this point in detail in our later discussion of Easy Knowledge of Methods.

[8] Here we idealize away from cases in which risk of error is introduced into the agent's belief forming mechanism through the risk of reasoning badly. See Lasonen-Aarnio 2008 for discussion.

[9] We have used a single premise example. Extended Methods also applies when the agent reasons from multiple premises.

[10] Note that extended methods need not be externally individuated; on one way of spelling out the view, they supervene on the agent's internal mental states. In section 4 we consider in detail how our argument bears on questions involving externalist versus internalist individuation of methods.

Easy Knowledge of Methods If S believes p by using method M, then S is in a position to know that *S believes p using method M.*[11] [12]

To prove this, we suppose that an agent believes p by using method M, and show that the agent is in a position to know that they believe p using method M. In particular, there is a 'boomerang' method of reasoning available to them that is guaranteed to produce a true belief.

To make our argument, we suppose that agent S believes p using method M. We then suppose that S reasons in the following 'boomerang' pattern. First, suppose S forms the belief that *S believes p using method M*, and deduces the conjunction q = (p and I believe p using method M). Second, suppose that S deduces from this conjunction the further conclusion that she believes p using method M. Now we argue:

1. By Extended Methods, S's method for believing q includes as an essential part that S believes p using method M.
2. By Extended Methods, S's method for believing by deduction from q that they believe p using method M includes as an essential part S's method for believing q.
3. From 1 and 2, S's method for believing by deduction from q that they believe p using method M includes as an essential part that S believes p using method M.
4. From 3, S could not easily have believed falsely that they believe p using method M using their actual method for believing that they believe p using method M by deduction from q.
5. From 4 and Basic Safety, S knows that they believe p using method M.

Summarizing, Extended Methods guarantees that there are argumentative paths available to agents which guarantee that part of their method for believing the conclusion entails that the conclusion is true. Basic Safety then guarantees that such argumentative paths produce knowledge. Thus, whenever an agent believes p using method M, they have a 'boomerang' method of deductive reasoning available to them which allows them to know that they believe p using method M. In this way, the agent is in a position to know that they believe p using method M. This concludes the argument that Extended Methods and Safety imply Easy Knowledge of

---

[11] Throughout the paper, we ask the reader to read constructions of the former 'S believes/knows that they believe p using method M' in such a way that 'method M' specifies the content of the belief. (There is an alternative reading of such constructions which takes 'p' as the content and 'using method M' to be merely a commentary on how the belief that p was formed. We could have disambiguated by suitable deployment of parentheses or italics throughout, but trust that this guidance will suffice for resolving any ambiguity in the appropriate way.)

[12] Throughout the paper, we rely on the concept of being in a position to know. Roughly, S is in a position to know p when S is 'a belief away' from knowing: if S forms a belief that p without otherwise altering their epistemic position, they will count as knowing (see Willard-Kyle 2020 for recent discussion). Our argument will imagine an agent who forms a belief about their methods, and explore whether this belief would count as knowing. Throughout, we help ourselves to the idealizing assumption that the relevant agent is also 'in a position' to form the relevant belief. That is, there are no conceptual barriers to the agent forming a belief about her method for forming beliefs. Stepping away from this idealizing assumption, our argument would establish a weaker result: that if S believes p using method M, and if S is in a position to believe that S believes p using method M, then S is in a position to know that S believes p using method M. Crucially, however, nothing in our argument requires that the agent forms a belief about her methods on the basis of good evidence or sound reasoning. Rather, we can imagine the agent forming the relevant belief about methods purely on the basis of a bizarre method, such as consulting tea leaves.

Methods.[13] [14] In the next sections, we consider possible responses to the argument: modifying Extended Methods, modifying Safety, or embracing Easy Knowledge of Methods.

But before we continue, we would like to make a few observations about the argument so far. First, this style of argument is stronger than we've said so far. Consider a dramatic strengthening of Basic Safety, which says that S knows p only if S could not possibly believe p falsely using their actual method. This principle is generally skeptical, saying that *any* possibility of error in which the same method is used blocks knowledge. But even such a skeptical principle still licenses knowledge of methods by a boomerang argument. The key is that Extended Methods implies that S couldn't possibly believe the conclusion of the boomerang pattern of reasoning without believing p using method M.

Second, note that nothing in the argument depends on any kind of modal stability regarding either S's belief that p or S's belief that S believes p using method M. Consider any world in which S believes the conclusion of the reasoning, that S believes p using method M, via the boomerang method. Extended Methods implies that in any such world, S does believe p using method M. This is compatible with there being another nearby possibility in which S falsely believes that she believes p using method M (though of course such a belief would not be formed using the boomerang method). It is also compatible with there being another nearby possibility in which they fail to believe altogether that they believe p using method M.[15]

Third, a natural first reaction to the boomerang argument worries that the agent reasons in a problematically question-begging or circular fashion, deducing the conclusion that she believes p using method M by a chain of reasoning that begins with this very premise. Perhaps there is some conception of competent deduction according to which not every classically valid inference is competent, because of circularity. In that case, neither Extended Methods nor Closure apply, since both principles are restricted to cases of competent deduction. This response fails for two reasons. First, it only targets the second step of the agent's reasoning, where they infer that they believe p using method M from the conjunction of p and the claim that they believe p using method M. But without having the agent employ any circular reasoning, the boomerang argument still establishes a weaker version of Easy Knowledge of Methods (discussed in more detail later), showing that whenever the agent knows p by using method M, they are in a position to know the conjunction: p and they believe p using method M. Second, there are related patterns of reasoning that are not similarly circular. Imagine that the agent begins the reasoning with the belief: either not p or they believe p using method M. They also believe p by using method M. Using disjunction elimination, they arrive at the belief that they believe p using method M. Again, Extended Methods implies that their method for believing that they believe p using method M includes as an essential part the fact that they believe p using method M. Again, Basic Safety implies that they come to know that they believe p using method M, no matter what the epistemic status is of the belief in the disjunction. Thus one cannot evade

---

[13] The boomerang argument also establishes that for any claims p and q, the agent can believe q using a method that includes as an essential part their method for believing p.

[14] The boomerang argument does not directly apply to agents who lack the conceptual resources necessary to form beliefs about what they believe. Rather, strictly speaking the boomerang argument only establishes that whenever an agent believes p using method M and is in a position to believe that this is so, they are also in a position to know that this is so.

[15] On the other hand, in our discussion of Counterpart Safety below we will consider subtler forms of Safety that require subtler assumptions about the modal profile of the agent's beliefs in order to generate an easy pathway to knowledge of methods.

the easy knowledge problems generated by Extended Methods merely by adding a 'no circular reasoning' constraint on competent deduction.

3. Responses

This section explores responses to the boomerang argument.

3.1 Variations on Extended Methods. One response is to defend a more sophisticated version of Extended Methods, which preserves Closure but which does not problematically imply Easy Knowledge of Methods. In particular, one might let the individuation conditions for methods depend on whether the premises are known. Only known premises transmit their methods to the conclusion; ignorant premises do not.

Weak Extended Methods If S competently deduces a conclusion from some premises p1 through pn, then S's method for believing the conclusion includes as an essential part S's method for believing pi iff S knows pi.

Weak Extended Methods validates Closure in the same way as Extended Methods, since Closure concerns cases in which every premise is known. Moreover, Weak Extended Methods avoids Easy Knowledge of Methods. One step in the boomerang argument lets the agent deduce that they believe p using method M from the conjunctive belief that p and they believe p using method M. Extended Methods implies that their method for believing the conclusion includes their method for believing the conjunction. But Weak Extended Methods does not have this consequence, since they may fail to know the conjunction, for example because they don't know p.

While Weak Extended Methods avoids Easy Knowledge of Methods, it still yields problematically easy pathways to knowledge of methods. In particular, it validates a weakening of Easy Knowledge of Methods:

Fairly Easy Knowledge of Methods If S believes p using method M and S knows p, then S is in a position to know that they believe p using method M.

The first step in the boomerang argument lets the agent deduce the conjunction p and they believe p using method M from their belief in each conjunct. Now suppose the agent actually knows p. Fairly Easy Knowledge of Methods implies that their method for believing the conjunction includes their method for believing p, which is M. So the agent safely believes and therefore knows the conjunction p and they believe p using method M. Another application of Weak Extended Methods then completes the boomerang argument as before. (The result is particularly absurd if you imagine that the agent forms their belief about methods on a wildly irrational basis, such as consulting tea leaves. Even then, the boomerang argument will give them knowledge of methods.)

In response, one might weaken Extended Methods further. On this new weakening, the failure to know any individual premise in an argument triggers a chain reaction, where none of the known premises end up essential to one's method for believing the conclusion.

Fragile Extended Methods If S competently deduces a conclusion from some premises p1 through pn, then S's method for believing the conclusion includes as an essential part S's method for believing pi iff S knows all of p1 through pn.

If an agent deduces a conclusion from a set of premises and the agent is ignorant of any one of these premises, then the agent's method for believing the conclusion fails to include their method for believing any of the premises. But if the agent does know every premise, then the agent's method for believing the conclusion includes their method for believing every premise. Fragile Extended Methods validates Closure, since Closure concerns cases in which every premise is known.

Fragile Extended Methods avoids the original boomerang argument. The boomerang argument begins with the agent deducing the conjunction p and the agent believes p using method M. But this deduction has as a premise that the agent believes p using method M. When the agent is ignorant of this premise, the agent's method for believing the conjunction will not contain as an essential part their method for believing p.

But Fragile Extended Methods faces some important challenges. First, it arguably overgenerates ignorance. Here are two useful cases to consider:

Movie Alice has four brothers: Alex, Billy, Charles, and Danny. She comes to know by reliable testimony that Alex, Billy, and Charles went to the movies last night. Alice then reads tea leaves and on this basis forms a belief that Danny also went to the movies. As a matter of fact, Danny did go; but he almost stayed home. Using all four premises, Alice infers that each of Alex, Billy, Charles and Danny went to the movies and subsequently infers that at least three of her brothers went to the movies. Alice could easily have also used tea leaves to believe that each of Alex, Billy, and Charles went to the movies, and the brothers could have easily failed to go to the movies (though under such circumstances the reliable informant wouldn't have said that they did).

Memory You see Jill, who has slightly under a pound of salmon on her plate, and believe that Jill has less than a pound of salmon on her plate. You don't know that Jill has less than a pound of salmon on her plate (because she easily could have had a pound or slightly more while you had a matching visual experience to the one you actually had). Further, it could easily have happened that an unreliable informant told you that Jill had less than a pound of salmon on her plate in a situation where she had two pounds or more. Now suppose you form a memory to the effect that Jill had less than a pound of salmon on her plate. Later someone asks you if Jill had less than two pounds of salmon on her plate. You call up your belief that she had less than a pound of salmon from preservative memory, though you have no episodic memory of any particular visual experience to go along with it (you have forgotten the perceptual details). On this basis, you deduce that Jill had less than two pounds of salmon. Jill could easily have had two pounds of salmon though you couldn't easily have had a perceptual experience that matches your actual experience under those circumstances.

We submit that in each case, the judgment that one knows the relevant conclusion (in the first case that at least three brothers went to the movies, in the second that Jill has less than a pound of salmon on her plate) has substantial plausibility. But Fragile Extended Methods delivers the result that the conclusion is known in neither case.[16] In Movie, since she doesn't know that Danny went to the movies, Fragile Extended Methods does not hold fixed the reliance on testimony (in the case of three of the brothers) when it evaluates the epistemic status of the belief that at least three brothers went to the movie. (By contrast, according to Fragile Extended

---

[16] Notice moreover that *both* Weak and Fragile Extended Methods seem to yield the result that the conclusion is not known in the Memory case, so that kind of case raises an additional concern about Weak Extended Methods.

Methods, if Alice had not formed this extra belief about Danny, but had formed the other beliefs in the same way, she would have known that at least three of her brothers went to the movies.) And in Memory, Fragile Extended Methods will not include the perceptual method for believing that Jill had less than a pound of salmon as part of her method for believing the conclusion that she ate less than two pounds. Since there is a close world where Jill has two pounds of salmon, and where you believe by dodgy testimony that she ate less than a pound, there is presumably a close world where you falsely conclude from the hypothesis that she ate less than a pound that she ate less than two pounds. Hence you will not count as knowing in the actual circumstances that she ate less than two pounds.

Some may deny that Fragile Extended Methods overgenerates ignorance. Here, it is instructive to notice that examples not dissimilar to Movie and Memory have cropped up in the literature on counterclosure, the principle that beliefs that essentially rely in their method of formation on unknown premises are themselves unknown.[17] Some have taken such examples as providing a strong case against counterclosure. But counterclosure, though generally regarded as more tendentious than closure, certainly has its advocates. It would thus be optimistic to expect such examples against Fragile Extended Methods to carry the day for all readers.[18] Still, there are reasons for thinking that the strategy of defending closure using Fragile Extended Methods is problematic.

First, notice that Fragile Extended Methods is a close cousin of counterclosure. Granted, the combination of Basic Safety and Fragile Extended Methods doesn't straightforwardly imply counterclosure. For example, consider a case where there is no close world where p is believed falsely but where p is inferred from an unknown premise. Here the combination of Basic Safety and Fragile Extended Methods certifies the belief as knowledge. Yet on the face of it, counterclosure denies that this is a case of knowledge. But the gap may not be so wide. After all, the safety theorist might wish to adopt an interpretation according to which a premise is inessential if the conclusion is not believed falsely at any close world by any method. In addition, one might well be tempted to refine Fragile Essential Methods so that a premise infects other premises only when it is an essential premise. In sum, the strategy of using Fragile Extended Methods to defend Closure in a Safety setting in effect amounts to the strategy of using a theory that implies a close cousin of counterclosure to defend Closure. And that is in itself quite striking: Closure is typically taken to be compelling. Counterclosure, by contrast, is much more tendentious. It is thus prima facie a little concerning if the safety theorist ends up relying on something in the vicinity of counterclosure in order to retain closure.

Our second observation is perhaps even more troubling. The point of Movie and Memory was to suggest that Fragile Extended Methods denies knowledge in too many cases. But it faces a

---

[17] See for example, Warfield 2005, Klein 2008, and Hawthorne and Rabinowitz 2017.We shall not get into the delicate issue of how the ideology of essential premises is to be cashed out. Hawthorne and Rabinowitz discuss cases of induction that intuitively secure knowledge but which include an unknown datum in combination with a smattering of known data as the inductive base. This is an inductive analogue of Movie. Both Warfield and Hawthorne and Rabinowitz discuss margin for error cases that bear a resemblance to Memory. (See for example, Warfield's 'handouts' case.)

[18] As a reader helpfully pointed out, those like us who wish to count Movie as a case of knowledge should also consider cases where the ratio of knowledge to ignorance is rather different. For example, if x had twenty brothers, believes each has gone to the movies, but knows in only three of the twenty cases, do we want to say that there is knowledge in this case that at least three went to movies if x indiscriminately relies on each of the twenty beliefs? And if not, what will our methodology be for drawing a line? We are inclined to count the twenty brother case as also providing knowledge that at least three brothers went to the movies, though we shall not defend that here.

second challenge: it may also provide too much easy knowledge of methods in other cases. Not all inference is deductive. We often use non-deductive patterns of inference to move from some beliefs to others. Considerations of simplicity and generality suggest that principles of method individuation should apply in the same way to deduction as to other kinds of inference. One would thus hope that any account of method individuation would extend to non-deductive modes of belief formation.[19] If holding fixed the method of the input of deduction depends on whether you know the input, then holding fixed the method of the input of other kinds of inferences that involve non-deductive patterns of inference should also depend on whether you know the input. Thus, in particular, if you non-deductively infer a conclusion from certain premises, the natural extension of Fragile Extended Methods says that if those premises are all known, then one's methods for believing those premises are an essential part of one's method for believing the conclusion. But when one extends Fragile Extended Methods in this kind of way, we get problematically easy pathways to knowledge.[20]

Consider a presumptuous agent, who performs a non-deductive inference whenever they believe p, immediately inferring that they believe that p by method M. They do not try to introspect first: they just go right ahead with the inference. (One version of this story involves an agent who adopts the practice of inferring from any proposition p to the conclusion that one believes p by method M. Other versions may have the agent adopting this practice for certain families of propositions.) Now suppose that the presumptuous agent happens to believe p by method M, and thereby knows p. If Fragile Extended Methods generalizes beyond deduction to the presumptuous inference, then the presumptuous agent knows they believe p using method M. Generalized to the presumptuous inference, Fragile Extended Methods would imply that when the presumptuous agent infers from p that they believe p using method M, their method for believing they believe p using method M includes as an essential part the fact that they believe p using method M.[21] Notice that the presumptuous agent gets to know even if they could easily have falsely believed that they believe p using method M. The result that the presumptuous agent knows is hard to live with. Knowledge of one's methods should not be so easy.

The central challenge to using extended methods to secure closure was that it made knowledge of one's methods too easy. Fragile Extended Methods is, so far, the most promising way of trying to get closure by suitable individuation of methods. But it is, at best, problematic. Cases that lead many to reject counterclosure will similarly lead many to reject Fragile Extended Methods on the grounds that it overgenerates ignorance. And it is in any case rather concerning that the attractive principle of Closure should be defended using a cousin of the much more tendentious principle of counterclosure. And finally, the natural extension of Fragile Methods to non-deductive patterns of reasoning again threatens to make knowledge of methods too easy. Of course one could try to keep the approach going in the face of these concerns. Cases that convince many to reject counterclosure are not completely uncontroversial. Thus, even though counterclosure-like principles are more tendentious, they are not out of the question. And this approach does at least get us Closure. Moreover one might somehow try to delimit the extension to non-deductive patterns of reasoning -- perhaps confining the extension to certain

---

[19] Das and Salow 2018 extend method-relativized safety to non-deductive methods in exactly this kind of way. (See their discussion of a principle they call 'Generality' on p. 14.)

[20] Indeed, insofar as all agents are in a position to be presumptuous in the way that we are about to explain, the considerations that follow indicate that the combination of Basic Safety and Fragile Extended Methods may still, after all, imply Fairly Easy Knowledge of Methods.

[21] The presumptuous agent is also a challenge for Weak Extended Methods, since the agent reasons from only one premise.

canonical non-deductive inference rules -- so as to avoid new easy knowledge problems.[22] Or perhaps one might bite the bullet and allow that the presumptuous agent knows, perhaps using Maria Lasonen-Aarnio's ideology of 'unreasonable knowledge' to soften the blow.[23] But we submit that the concerns are substantial enough to warrant looking elsewhere.

3.2 Variations on Safety. A second response to the boomerang argument is to reject biconditionals like Basic Safety on the grounds that safe belief is insufficient for knowing.[24] One example of such a theory retains the original Extended Methods, denies that method-relative safety is sufficient for knowledge, and adds counterclosure as an additional necessary condition. But it should already be clear what we think about that move: First, counterclosure is much more problematic than closure. Second, this combination still yields problematically easy knowledge when extended in the natural way to non-deductive inference (recall the presumptuous agent).[25]

There are other strategies for denying the sufficiency of safety for knowledge. For example, one theory says that knowledge is safe, justified belief.[26] If Safety fails to be sufficient in this way, then the argument for Easy Knowledge of Methods is blocked. The argument establishes that an agent who believes p using method M is in a position to safely believe this is the case. But the argument does not establish that they are justified in so believing, and so does not establish that they know.

This response to the boomerang argument is insufficient since it still makes knowledge of methods too easy. Although it avoids Easy Knowledge of Methods, this approach still validates Suspiciously Easy Knowledge of Methods.

Suspiciously Easy Knowledge of Methods If S knows p and S is justified in believing that they believe p using method M, then S is in a position to know that they believe p using method M.

Take any agent who knows p using method M, and who is justified in believing that they believe p using method M. Plausibly, their justification agglomerates to give them justification in believing the conjunction: p and they believe p using method M.[27] Now imagine that the agent infers from this conjunction that they believe p using method M. After such an inference, they will be justified in believing p using method M. But Extended Methods implies that their method for believing this last claim guarantees that it is true, since their method includes believing p using method M as an essential part. The result is that such an agent is both safe and justified in believing that they believe p using method M. If safe, justified belief is sufficient for knowledge, they would thereby know that they believe p using method M. Summarizing, we

---

[22] One version of this idea would say that unless one uses a non-canonical kind of non-deductive method, then the more demanding crude safety test is to be applied. Sorting out a suitable conception of canonical methods that can divide non-deductive inference rules and patterns of reasoning into canonical and non-canonical varieties will be a tricky business.

[23] See Lasonen-Aarnio 2010.

[24] See Pritchard 2012 and Goldberg 2015 for challenges to the sufficiency of safety for knowledge.

[25] For further critical discussion of the prospects of combining counterclosure with safety see Hawthorne and Rabinowitz 2017. One methodological theme is that the more one makes knowledge look like a hodge-podge of unrelated necessary conditions, the more gerrymandered it begins to look as an epistemological kind.

[26] For discussion see Pritchard 2005.

[27] To avoid worries about justification agglomeration, we can restrict the result to cases in which the agent is maximally justified in both p and the claim that they believe p using method M.

have the result that any agent who knows p and is justified in believing that they believe p using method M is also in a position to know that they believe p using method M.

Suspiciously Easy Knowledge of Methods should be rejected. Suspiciously Easy Knowledge of Methods implies that any agent is in a position to use boomerang style reasoning to escape Gettierization with respect to their method for believing p when they know p. For example, imagine that John believes p using method M, and John hears a psychologist saying "John believes p using method M". In fact, the psychologist is talking about a different John. Here, John is justified in believing that he believes p using method M. Suspiciously Easy Knowledge of Methods then implies that he knows that he believes p using method M. But we can imagine that John could easily have failed to believe truly that he believes p using method M. Below, in our discussion of Easy Knowledge of Methods, we will see several other examples where not only Easy Knowledge of Methods but also Suspiciously Easy Knowledge of Methods fails.[28] [29]

We turn to an entirely different cluster of strategies, ones that rely not on supplementing safety with an extra necessary condition but instead on modifying the safety idea itself. We shall be looking at two kinds of departure from the versions of safety presented so far. One departure appeals to counterpart beliefs, the other replaces 'same method' by 'similar method'.

Let's look first at the counterpart approach. The versions of method-relativized safety that we are looking at have something in common: if one believes p, then a nearby case of false belief will count as an epistemic indictment of one's actual belief that p only if that false belief is a false belief that p. But perhaps this requirement should be relaxed: one might respond by appealing to more sophisticated versions of Safety. Williamson considers a version of Safety that appeals to counterpart propositions. The idea is that when an agent knows p using method M, we should also consider what other beliefs they could have formed by employing method M. More precisely:

Counterpart Safety S knows p using method M if and only if S believes p and could not easily have formed a false belief using method M.[30]

It will also be useful to formulate Counterpart Safety using the ideology of counterparts. Let's say that a belief y of S's, formed at some world w, is a counterpart of some belief x iff x is (at w) formed by the same method. Counterpart Safety says that S knows P just in case S believes p and there are no close worlds where there is a counterpart of that belief that is false. In order to

---

[28] Compare Zagzebski 1994, who provides recipes for constructing Gettier cases regarding arbitrary topics.

[29] Here, one might deny that knowledge is safe, justified belief. Rather, one might think that knowledge is subject to an extra condition besides safety that accounts for Gettierization. Perhaps this condition is violated in the above example. Perhaps this style of account can allow that beliefs about methods are ignorant when they are Gettierized. We leave to our opponent the task of developing a systematic theory of Gettierization that is independent of the safety condition, and which can explain how beliefs about methods could be Gettierized even when safe, all while preserving Closure. But regardless of details, this account will predict that in these kinds of cases, if you don't know your method, it won't be because the belief is unsafe; it will be because of some other reason. Insofar as the safety condition is supposed to explain failures of knowledge like these, it would be surprising if beliefs about methods never exhibited the relevant kind of safety failure. The case of John above seems to involve a failure of precisely this kind.

[30] One might instead try a formulation along the lines of 'could not have easily formed a false belief in a similar proposition using method M'. We shall not work with this as we find the presumed similarity metric on propositions hard to fathom (a problem intensified if one repudiates structured propositions.) If a reader feels they can make good sense of such a formulation we invite them to work through its implications for our main subject matter in light of the following discussion.

validate Closure in the presence of Counterpart Safety, Extended Methods requires a slightly different formulation.[31]

Counterpart Extended Methods If S believes p1 through pn using methods M1 through Mn and deduces q from these premises, counterparts of that conclusion will be all and only beliefs reached using the same pattern of reasoning from outputs of methods M1 through Mn.

Now consider the boomerang argument. We imagined an agent who believes p using method M, and who also forms a belief that they believe p using method M. How is this latter belief formed? Our initial argument applied regardless of the agent's initial method for forming this latter belief. Now we must be more careful. Imagine the agent believes on a whim that they believe p using method M. This whimsical method could (presumably) have easily produced the false belief that they believe p using method M'. In this sort of case, Counterpart Extended Methods avoids the boomerang argument. Imagine that the agent forms the conjunctive belief: p and they believe p using method M. Counterpart Extended Methods implies that counterparts of that conjunctive belief will be those beliefs formed by conjunction introduction from a belief formed using method M and a belief formed by whimsy. But now consider a nearby world in which they still form a belief that p using method M, but by whimsy form the belief that they believe p using method M', and then deduce the conjunction: p and they believe p using method M'. In this nearby world, the conjunctive belief that they form is false. Counterpart Extended Methods implies that their method in this nearby case is the same as their original method. So Counterpart Safety implies correctly that they fail to know the conjunction: p and they believe p using method M. Similarly, if they went on to deduce that they believe p using method M, Counterpart Safety would imply that they fail to know this claim.

So far, it appears that Counterpart Safety avoids our result. Unfortunately, this response is still problematic. While the boomerang argument does not apply to whimsical agents, it does apply to another kind of agent, who uses a 'selective' method. Let's say an agent's method of forming a belief that p is selective just in case that agent could not easily have formed a belief in some other proposition by that very method. Now suppose that S believes p using method M. Suppose that S now uses method M' to form the belief that S believes something using method M. But suppose further that S uses M' 'selectively': they could not easily have used M' to produce any output except a belief that S believes something using method M. Now suppose S deduces the conjunction q = (p and S believes something using method M). Finally, suppose that they deduce from q that they believe something using method M. So this is the pattern of reasoning: They form a belief by M, a belief by M', conjoin the contents of beliefs formed by M and M' and then by conjunction elimination conclude with the content originally formed by M'. Now we can show that S knows that they believe something using method M:

1. Since M' is selective, every counterpart of S's belief that S formed a belief by method M is a belief that S formed a belief by method M.
2. By Counterpart Extended Methods, every counterpart of the conclusion is a belief that S formed a belief by method M (since that pattern of reasoning involves concluding by conjunction elimination the premise originally believed by method M', which by 1 is the belief that S formed a belief by method M.)

---

[31] See Williamson 2009b, p. 326 for a related formulation. As Williamson is aware, counterpart methods will avoid the conclusion that one knows any necessary truth that one believes, a result that is not avoided by any of the approaches we have been looking at so far. (More generally it avoids the result that one knows for free any proposition that one believes and that is true at all nearby worlds.)

3. By Counterpart Extended Methods, every counterpart of the conclusion is a belief by a chain of inference that includes a belief formed by method M.
4. So, by 2 and 3, every counterpart of S's conclusion is true.
5. So by Counterpart Safety, S knows that S believes something using method M.

Being selective is an interesting property. Note that the case does not require that S's belief about method M is stable, in the sense that S is guaranteed to have this belief throughout nearby modal space.[32] Rather, all it guarantees is that whenever S uses their actual method M', they have the belief that they believe something using method M. Note also that selectiveness does not preclude the possibility of error. S could use M' selectively even though they could very easily falsely believe that they believe something using method M (though not via a counterpart of the boomerang argument). Since selectiveness does not preclude the possibility of error, our result has bite.

It is totally implausible that using a method selectively in this way suffices for knowledge of methods. And yet Counterpart Safety and Counterpart Extended Methods imply that selectively using a method does put one in a position to have knowledge of methods. We thus still have an easy knowledge problem.

Different theories of method individuation make different predictions about what a selective method looks like exactly. But existing approaches to method individuation all imply that it should be possible to selectively use a method. The basic trade-off is between individuating methods finely and coarsely. The coarser the method, the harder it is to know anything; the finer the method, the easier it is to use a method selectively.

For example, on one popular proposal, when someone forms a belief on the basis of testimony, their method holds fixed the identity of the speaker, facts about the speaker's reliability, and the fact that the belief is formed by testimony. But now imagine that the agent's belief about their own method of belief formation came from testimony (for example, from a reliable therapist). Imagine further that the testifier could not easily have failed to assure the agent that they believe something using method M and could not easily have testified to some further proposition as well.[33] A selective scenario is particularly easy to imagine if the testimony comes from a very short book that explicitly says only that S believes something using method M and could not easily have explicitly said something else. This is one of many easy recipes that produce a selective method once a reasonable fine-grained individuation of methods is assumed.

On the other hand, a coarser approach to method individuation would say that the agent's method above is simply forming a belief by testimony. In the presence of Counterpart Safety, this coarse approach to method individuation threatens to block too much knowledge: even when you receive information from a reliable informant, the nearby possibility of speaking to an unreliable informant destroys knowledge, even if you can distinguish the two informants.[34] Nonetheless, even such a coarse approach to method individuation still in principle allows for

---

[32] Thanks to an anonymous referee for pressing us to clarify this issue.
[33] Alternatively, consider a case in which, while they could have easily testified to some other thing, the only thing that the subject is disposed to trust is testimony with respect to the proposition that they believe something using method M.
[34] Note that in the setting of non-counterpart safety, the sceptical effects of coarse methods may not be quite so widespread as beliefs in other propositions are irrelevant. Also, even in the presence of Counterpart Safety, sceptical effects can be somewhat contained by imposing a draconian constraint on which worlds count as genuinely close (though that will not help when one has multiple beliefs that are in fact produced by the same coarse method, some of which are false).

selective use of methods: simply imagine that the agent could not easily have learned anything from testimony except that they believe something using method M. In short, it is extremely challenging to contrive a conception of methods that blocks the possibility of easy knowledge from selective methods.

So much for Counterpart Safety. The second kind of variation on safety that we wish to consider appeals to similarity of methods rather than sameness of methods:

Similar Safety S knows p using method M if and only if S could not easily have falsely believed p using any method M' similar to M.[35]

Given Similar Safety, Extended Methods requires a slightly different formulation. It now says that when an agent deduces q from p, they employ a method that is similar to another episode of deduction only if that episode of deduction employs a method for believing p that is similar to the actual one:

Similar Extended Methods If S believes p1 through pn using methods M1 through Mn and deduces q from these premises using method N, then any method N' of S's deducing q from p1 through pn is similar to N only if method N' contains the facts that S believes p1 through pn using methods M1' through Mn', where M1' through Mn' are similar to M1 through Mn.

Similar Extended Methods avoids Easy Knowledge of Methods as we have stated it above, but falls prey to a close relative. Similar Extended Methods implies that agents are in a position to know a great deal about their method of belief formation:

Easy Knowledge of Similar Methods If S believes p using method M, then S is in a position to know that they believe p using a method similar to M.

Imagine now that the agent believes p using method M, and also forms a belief that they believe p using a method similar to M. Now imagine that they deduce the conjunction of these claims. Similar Extended Methods implies that any method similar to their method for believing this conjunction will be one in which they believe p using a method similar to M. Imagine now that they deduce from this conjunction that they believe p using a method similar to M. Call their method for believing this last claim M'. Similar Extended Methods implies that any method M'' is similar to M' only if M'' contains the fact that the agent believes p using a method similar to M. The result is that the agent could not easily have used a method similar to M' to falsely believe that they believe p using a method similar to M. So Similar Safety implies that the agent who reasons in such a way would know that they believe p using a method similar to M. In short, Similar Safety does not afford an escape route from easy knowledge problems.

4. Easy Knowledge of Methods and Method Individuation

Our central theme has been that extended methods yields implausibly easy routes to knowledge of methods. One vivid illustration of that theme was the observation that the combination of Basic Safety and Extended Methods yields Easy Knowledge of Methods. And elsewhere we have observed that variants on safety yield interesting variants on Easy Knowledge of Methods (including, for example, Easy Knowledge of Similar Methods). However, the significance of Easy Knowledge of Methods (and its variants) depends on how methods are individuated. Some readers might wonder whether one might learn to live with Easy Knowledge of Methods

---

[35] One finds a similarity-theoretic approach in Hawthorne 2004.

(or one of the variants we have looked at) by adopting a suitably internalist conception of methods.

One point bears emphasis at the outset. Even if one could somehow learn to live with Easy Knowledge of Methods, this would not dissolve all the easy knowledge problems raised by the boomerang method. This is because the boomerang argument not only militates in favor of Easy Knowledge of Methods, but for the yet stronger conclusion that one is always in a position to know one's methods *by boomerang style reasoning*. Even if one were somehow to convince oneself to learn to live with Easy Knowledge of Methods, one should still balk at the idea that boomerang reasoning is a legitimate method of coming to know one's method. For this reason we do not think it very promising to try to make easy knowledge problems go away by trying to find a conception of methods that secures Easy Knowledge of Methods. That said, we think it is instructive to inquire whether there is some conception of methods that would make Easy Knowledge of Methods somewhat plausible in a setting where one is committed to combining Closure with some version of the safety idea.

One open debate is whether to individuate methods externally or internally.[36] Williamson 2000 (p. 95), Pritchard 2005 (p. 153), Williamson 2009b, Brancano-Berrocal 2014, Grundmann 2018 and others defend an externalist conception of methods. By contrast, Goldman 1976, Nozick 1981, Sosa 1999, Goldman 2009, and others defend an internalist conception of methods. Let's see what Easy Knowledge of Methods might come to on each conception.

We'll begin with externalism about methods. This kind of approach helps to generate anti-skeptical conclusions within a Safety framework.But supplementing externalism about methods with Extended Methods leads to far too much knowledge of method.

Consider an example originally from Neta and Rohrbaugh 2004:

Chemical "I am participating in a psychological experiment, in which I am to report the number of flashes I recall being shown. Before being shown the stimuli, I consume a glass of liquid at the request of the experimenter. Unbeknownst to either of us, I have been randomly assigned to the control group, and the glass contains ordinary orange juice. Other experimental groups receive juice mixed with one of a variety of chemicals which hinder the functioning of memory without a detectable phenomenological difference. I am shown seven flashes and judge, truly and knowingly, that I have been shown seven flashes. Had I been a member of one of the experimental groups to which I was almost assigned, I would have been shown only six flashes but still believed that I had been shown seven flashes due to the effects of the drug. It seems that in the actual case I know that the number of flashes is seven despite the envisaged possibility of my being wrong. And yet these possibilities are as similar in other respects as they would have to be for the experiment to be well designed and properly executed." (Goldman 2009, p. 79)

Neta and Rohrbaugh 2004 say that this case is a counterexample to Safety. (For what it's worth, we think the best case for them is a version of the story where we do not know the setup and so do not know that any subgroup in the experiment is going to be having flash-perception difficulties.) Insofar as the Safety theorist agrees that the agent (at least in this version of the

---

[36] The problem of individuating methods appropriately in the context of Safety is structurally similar to the generality problem for reliabilism. See Grundmann 2018 for discussion.

case) knows he was shown seven flashes, the case is then a prima facie challenge to Safety.[37] In response, Williamson appeals to an externalist conception of methods: such cases "involve large external differences between the processes that actually generate the beliefs in question in the good cases and the processes that generate errors in the corresponding bad cases. Again, these differences turn out to constitute large enough dissimilarities between the bases of one's beliefs to allow one to know: the cases are not similar enough to exclude safety" (Williamson 2009b, p. 307).

In Chemical it is plausible that the agent knows that they have seen seven flashes (at least when they do not know the protocol).[38] But it is implausible to think that this agent is in a position to know that their glass contains ordinary orange juice.[39] But if Williamson's externalist conception of methods is correct, then the agent's method for believing that they were shown seven flashes essentially included their glass containing ordinary orange juice. For this reason, Easy Knowledge of Methods would falsely imply that the agent is in a position to know that their glass contains ordinary orange juice. The combination of externalist methods and Easy Knowledge of Methods is bizarre.

Note, indeed, that the boomerang argument applies to this case. Suppose the agent knows that there were seven flashes, and also forms the belief that there is a glass of ordinary orange juice in front of him. If they infer the conjunction of these claims, their method for believing the conjunction will by Extended Methods include as an essential part their method for believing there were seven flashes, which (since they know there were seven flashes) holds fixed that the orange juice was ordinary. Similarly, if they now infer that there is a glass of ordinary orange juice in front of them from this conjunction, their method for believing this conclusion will include as an essential part their method for believing the conjunction. The result is that the agent is in a

---

[37] For what it's worth, we agree with the judgments of Neta and Rohrbaugh 2004, although some informants (including an anonymous referee) disagree.

[38] An interesting separate challenge to externalist methods concerns cases where the protocol is known. Suppose one knows that half of the participants will imbibe dodgy juice that will impair flash perception. Here it is natural to think one doesn't know whether one is seeing seven flashes when one does. But extended methods threatens to deliver the result that one does still in that case know that one is seeing seven flashes if one takes experience at face value (since the widely individuated method still apparently secures Basic Safety). Thus internalist methods seems to get bad results when subjects are ignorant of the protocol and externalist methods seems to get bad results when the protocol is known. We shall not take up this challenge to external methods here though think it an important one. (We are especially grateful to a referee here.)

[39] Some might be tempted to allow that the agent in this case can know that the glass contains ordinary orange juice. After all, it is more normal to drink ordinary orange juice than it is to drink hallucinogenic orange juice. But imagine a variant where all participants are given some drug x, but where x causes anomalous vision only when combined with Coke but not with Pepsi. Some participants are given Coke, others Pepsi. It seems incredible to suppose someone who was ordinarily hopeless at telling the difference could know they were drinking Coke by the boomerang method.

position to know that they are drinking ordinary orange juice, since their method for believing this claim essentially includes the fact that the orange juice is ordinary.[40] [41] [42]

The combination of externalist conceptions of methods with Easy Knowledge of Methods is not defensible. But there are other conceptions of methods available. Goldman 1976, Nozick 1981 (pp. 184-5), Sosa 1999, and Goldman 2009 all endorse an internalist conception of methods, according to which an agent's method is constituted by their relevant internal experiences.[43] For example, one version of this thesis in the setting of Similar Safety says that two methods are similar to one another just in case they consist of similar relevant internal experiences.[44]

Here, an important choice point concerns time sensitivity. Suppose an agent believes p at time t. Their internal method for believing p could either consist of their relevant internal experiences at t, or could instead consist also of relevant experiences they had in the past leading up to t.

[40] For brevity, we focus on Chemical. But other cases can be used to make a similar point, including Sosa 2007's Kaleidoscope case (Sosa 2007, p. 31; Grundmann 2018). In addition, Williamson 2009b, p. 326 suggests that when agents form beliefs by memory, their present basis for believing includes their past and possibly forgotten basis; and that when hearers form beliefs by testimony, their basis includes the speaker's basis. Such conceptions of methods sits in tension with Easy Knowledge of Methods.

[41] In fact, the boomerang argument establishes that externalist approaches to method individuation sometimes lead to a weak version of the KK thesis, which is still too strong to be plausible. Say that a method M for believing p is factive iff necessarily if an agent uses M to believe p, then p is true. Some externalist conceptions of methods lead to instances of factive methods. Some externalists embrace the principle that the reliability of a speaker is essential to a hearer's method for believing (compare Williamson 2009b, p. 326, Grundmann 2018). Now imagine that a speaker tells the hearer: I am reliable. Similarly, consider the principle that the normality of lighting conditions is essential to the method of a perceiving believer (as in Sosa 2007, p. 31). Now imagine that such a believer picks up a newspaper in normal lighting conditions, and reads the sentence: lighting conditions are normal. In each case, the agent's method for believing is factive. They could not have believed the relevant claim using their actual method without that claim being true.

The boomerang argument establishes that Extended Methods leads to the validity of a weakening of KK: that whenever an agent believes p using a factive method, they are in a position to know that they know p. Suppose such an agent forms the belief that they know p. Now imagine as usual that they deduce the conjunction: p and they know p, and then deduce that they know p. Chaining together applications of Extended Methods, we reach the result that their method for believing the conclusion that they know p includes as an essential part their factive method for believing p.

Now we can show that their belief that they know p is knowledge. In order to show this, we can apply Safety, showing that at any easy possibility v where they believe they know p using their actual method, it is true at v that they know p. Consider any such v, where they believe they know p using their actual method. We can show that at v, they know p. After all, applying Safety again, in order to know p at v, it must be the case that at any easy possibility u accessible from v where the agent believes p using their world v method for believing p, p is true. This is again guaranteed by the factivity of the agent's method. At world u, the agent's method for believing p is again a factive one, guaranteeing the truth of p. So since they believe p at world u, p is true at world u. The result is that at v the agent knows p, and so at w the agent knows they know p.

[42] Some readers may wonder whether the defender of externalist methods can bite the bullet about the Chemical case. Here, one thing to notice is that the relevant belief about methods (that the agent was drinking ordinary orange juice rather than chemically altered juice) could have been formed essentially at random, by consulting tea leaves. Even in that case, the boomerang argument will establish that the agent knows that they drank ordinary orange juice.

[43] 'Relevant experiences' might perhaps be cashed out in terms of those experiences which sustain or cause a belief, but we will leave the issue wide open here.

[44] See Zhao 2020 for further discussion.

The 'presentist' version of internalism can do some of the work required by a successful account of method individuation. It explains Nozick 1981's grandmother case, in which granny knows that her grandson is healthy by perception even though they easily could have falsely believed the same via misleading testimony. Her internal experiences when undergoing perception differ significantly from the experiences they would have had if they received testimony.

The combination of Presentist internalism with Easy Knowledge of Methods is certainly more promising, though even here, Williamson 2000's anti-luminosity arguments raise a potential challenge for the thesis that whenever an agent has an experience, they are in a position to know that they are having that experience.[45] On the other hand, we saw above that those who accept Similar Safety need only accept Easy Knowledge of Similar Methods, which given presentist internalism says that when an agent believes p on the basis of present relevant internal experiences, the agent is in a position to know that they are having a present experience similar to one of those experiences. This condition does not fall prey to Williamson-style anti-luminosity arguments, because it does not posit the existence of any particular mental state which is known to occur whenever it in fact occurs. So the combination of Presentist Internalism, Similar Safety and Easy Knowledge of Similar Methods may be even more promising.

Unfortunately, while the presentist style of internalism may offer a principled way of endorsing Easy Knowledge of Similar Methods, it does not meet our challenge of reconciling Safety with Closure. The problem is that presentist internalism cannot accept any version of the Extended Method approach, and so cannot validate Closure. Closure characteristically involves episodes of deductive reasoning that are temporally extended. An agent believes p at an initial time t, and then engages in a process of competent deduction that results in the agent believing q at a later time t+. Presentist internalism implies that the agent's method for believing q can only consist of the agent's experiences at t+, and does not include experiences at earlier times and causal connections to beliefs at earlier times. But now Closure is in trouble. Suppose, for example, that at t, an agent has just come to believe q by deduction from knowledge that p but that at t, the knowledge that p does not manifest itself by any tell-tale experiences (we can imagine at t it has become phenomenologically dormant, as much of our knowledge is much of the time). Suppose S has a certain pattern of experiences at t. Given presentist internalism, there may be a close world where S does not believe p at all, where q is false, and where the agent believes q in the presence of the very same experiential pattern at t, without having deduced q from p. The agent fails to know the conclusion q, since there is a close world where their experiences at t are the same and they believe q falsely. But this case allows that the agent knows p: the fact that there is a close world where the agent does not believe p is compatible with their actual belief that p being knowledge, since the safety idea is that it is nearby error not nearby absence of belief that blocks knowledge. So holding fixed a presentist version of internalism, Closure fails.[46]

---

[45] But see Leitgeb 2002, Weatherson 2004, and Berker 2008 among others for potential replies.

[46] An anonymous referee wonders if the argument can be avoided at least in cases where at the later time of believing q, the agent retains the belief that p and the belief that p implies q, supposing that we let the presence of such beliefs be part of the agent's method. Here, it is instructive to note that even in such cases, presentists cannot achieve Closure. There are two problems. First, even if we hold fixed the facts that at the later time the agent believes p and believes that p implies q, this does not guarantee that at the earlier time the agent had the relevant phenomenal experiences. Second, such an approach to method individuation does not hold fixed the fact that the agent based her belief in q on her belief in p. It could be that at the later time the agent believes p and believes that p implies q, and yet based her belief in q on tea leaves.

In response to this case, the internalist could reject the presentist version of that view, allowing that past internal experiences of the agent can be part of their method for presently believing p. But this proposal is naturally combined with the rejection of Easy Knowledge of Methods. The present agent may not be in a position to know what they experienced in the past if, for example, they have forgotten what chain of experiences and reasoning led to their present belief.[47] Unfortunately, though, if we combine this version of internalism with Extended Methods, Easy Knowledge of Methods will again hold, for familiar reasons.

Summarizing, the only way of making an easy knowledge of methods principle plausible is by opting not merely for an internalist conception of methods, but moreover, one that counts only present aspects of experience as relevant to method at any particular time. But this conception of methods will preclude the use of methods to reconcile Closure with a safety-theoretic approach.

5. Conclusion

This paper argued that theories of Safety that validate Closure also lead to problematically easy ways of knowing one's method for belief formation. Our central argument was that the Extended Methods strategy for validating Closure, when combined with Basic Safety, implies Easy Knowledge of Methods.

In the face of these challenges, it is natural to wonder whether there is any way of validating Closure and Safety without generating implausibly easy paths to knowledge. Here, Fragile Extended Methods offers at least a proof of concept. This principle is far less susceptible to easy knowledge results, at least so long as one doesn't use analogous individuation of methods for certain non-deductive methods of belief formation. Yet we saw that Fragile Extended Methods faced some quite serious challenges. Counterpart-theoretic approaches to safety also blocked certain of the easy knowledge results. But it seemed that it did not quite rid us altogether of problematically easy knowledge of methods in certain settings. An open question for future research is what approach to method individuation can validate some combination of the safety idea and Closure without yielding untoward access to one's own methods. We have no impossibility proof to offer concerning a satisfactory reconciliation of Safety and Closure; rather, we have shown that the current resources of the literature are not up to the task. (In an appendix, we consider one other strategy in the literature for reconciling Safety and Closure, and show that it suffers similar problems to certain of the Extended Methods ideas.)

Appendix. Epistemic Methods

We have argued that Extended Methods strategy for validating Closure generates problematically easy ways of knowing one's methods. This raises the question of whether any alternative conception of methods can validate Closure without leading to such problematic kinds of access. In the existing literature, only one alternative to Extended Methods ensures the validity of Closure. Das and Salow 2018 and Schulz 2020 both embrace an *epistemic* conception of methods. The key idea is that when the input to a belief forming process is known, the fact that it is known is part of one's method for believing the conclusion. Here is first pass:

---

[47] Indeed Nozick 1981 explicitly rejects Easy Knowledge of Methods despite accepting internalism: "A person can use a method (in my sense) without proceeding methodologically, and without knowledge or awareness of what method he is using" (p. 184-5). See Becker 2012 for further reasons to doubt Easy Knowledge of Methods even in the presence of internalism about methods.

Epistemic Methods If S infer q from p1 through pn, then S's method for believing q includes as an essential part the fact that S knows pi iff S knows pi.

It is easy to show that, as formulated, this generates easy knowledge problems. Suppose one believes the proposition that either not q or one knows q on the basis of consulting tea leaves. As it happens one not only believes but knows q. By disjunction elimination one comes to believe that one knows q. Epistemic Methods implies that one thereby comes to know that one knows q. We submit this is quite bizarre. Even if one thought -- as KK lovers do -- that whenever one knows one is in a position to know that one knows, it is completely bizarre to suppose that *this* is a pathway to knowing that one knows that one knows. In response it is natural to retreat to Fragile Epistemic Methods:

Fragile Epistemic Methods If S infer q from p1 through pn, then S's method for believing q includes as an essential part the fact that S knows pi iff S knows each of p1 … pn.[48]

It is interesting to see how the Epistemic Methods approach plays out when extended to non-deductive patterns of inference. When putting the Epistemic Methods approach to work, Das and Salow rely on a more expansive notion of inference than deduction that includes all sorts of ampliative forms of inference. In particular, they are interested in cases where an agent performs an inference according to the rule: *if p, believe you know p*. The belief forming method here is a process whose input is a representation of p, and whose output is a representation of the claim that the agent knows p. Both Epistemic Methods and Fragile Epistemic Methods yield a striking result once one extends it to the rule above. It implies that whenever one knows, one is in a position to know that one knows using that rule. Indeed this is the key observation in Das and Salow 2018 -- they use the Epistemic Methods idea to argue in this very way for a KK thesis. Since the KK thesis has its fair share of supporters, this result might in some quarters be welcomed rather than being seen as a reason to worry about Epistemic Methods.[49]

There are reasons for concern however. First, as with Fragile Extended methods, this threatens to undergenerate knowledge in some cases. Those of us who accept closure but are dubious about principles in the vicinity of Counterclosure will not like Epistemic Methods. We have seen that the non-fragile version of epistemic methods clearly overgenerates knowledge. But insofar as one thinks that we have knowledge in examples like Movie and Memory we shall have to conclude that the fragile version undergenerates knowledge.

Second, one complication for Epistemic Methods (in both the Fragile and non-Fragile varieties) concerns the fact that inference rules take time to follow. Suppose that at t the agent knows it is raining and then follows the rule: if p, believe you used to know p. On this basis they infer at t' that they used to know it is raining. Epistemic Methods says that the agent's method essentially includes the fact that they know at t that it is raining. So the agent will count as knowing that they used to know p. This result goes beyond ordinary KK reasoning, since it even applies in cases where the agent's inference takes a long enough amount of time that they may have forgotten what they initially believed.

---

[48] Das and Salow 2018 and Schulz 2020 aren't completely explicit about whether they endorse Epistemic Methods versus Fragile Epistemic Methods. We think Das and Salow 2018 is best interpreted as accepting Fragile Epistemic Methods, since you would only count as 'following a rule' if you knew every premise of the rule.
[49]For sympathetic discussions of KK, see Stalnaker 2006 and Greco 2014.

Third, Epistemic Methods faces a challenge involving knowledge of which inference rules an agent follows. Suppose an agent infers q from p using rule R and knows p. It is very natural to require not only that an agent's method for believing q includes knowing p, but also that the agent used rule R.[50] This further restriction is natural, since intuitively whether an agent knows on the basis of following a rule is independent of what they would have believed as a result of following wildly different rules. But this extra component of the agent's method leads to problematic kinds of easy knowledge of rules.

In particular, imagine an agent S on some occasion follows the rule R*: From the premise that everything is F, infer (i) that Jones is F and (ii) that one has inferred that Jones is F from a universal generalization. The rule-enriched version of Epistemic Methods then predicts that when any such agent applies the rule to a known premise, they are in a position to know via an application of this rule that they inferred that Jones is F from a universal generalization. After all, the fact that they used R* is essential to their method for believing the claim that they inferred that Jones is F from a universal generalization. This kind of path to knowledge should be resisted. We can imagine agents who follow the above inference rule, but lack any internal access to the fact that they follow such a rule, and are wildly unreliable at detecting when they follow such a rule. Imagine that 99% of the time when they believe that they have inferred something about Jones from a universal generalization, they are wrong. Such agents should not count as knowing how they came by beliefs about Jones. In short, even if the proponent of Epistemic Methods (or Fragile Epistemic Methods) welcomes its KK validating capacity, natural versions of this approach yield problematically easy access to the methods by which one forms beliefs.[51]

Summarizing, the Epistemic Methods approach is an important alternative to the Extended Methods approach, offering an alternative path to reconciling Safety with Closure. But it also faces some pressing challenges, many of which are analogous to challenges to the extended methods idea that are rehearsed in the main body of the paper.

Bibliography

Alspector-Kelly, Marc (2011). Why safety doesn't save closure. *Synthese* 183 (2):127-142.

Becker, K. (2012). "Methods and How to Individuate Them," 81-97 in K. Becker and T. Black (ed.), The Sensitivity Principle in Epistemology, Cambridge: Cambridge University Press.

Beddor, Bob & Pavese, Carlotta (2018). Modal Virtue Epistemology. *Philosophy and Phenomenological Research* 101 (1):61-79.

Berker, Selim (2008). Luminosity Regained. *Philosophers' Imprint* 8:1-22.

Broncano-Berrocal, Fernando (2014). Is Safety In Danger? *Philosophia* 42 (1):1-19.

Comesaña, Juan (2005). Unsafe Knowledge. *Synthese* 146 (3):395-404.

---

[50] See for example Das and Salow 2018, p. 14. Notice that the inclusion of rules in methods will be particularly important if one uses a 'counterpart belief' approach that allows that content of epistemically relevant beliefs to vary.

[51] As with Fragile Extended Methods, one might try restricting the non-deductive applications of the epistemic methods approach to canonical inferences. We note that insofar as one goes that route, one might well question Das and Salow's KK defense on the grounds that the inference rule they appeal to is non-canonical. Since the canonical vs non-canonical distinction is murky, we will not pursue this any further.

Das, Nilanjan & Salow, Bernhard (2018). Transparency and the KK Principle. *Noûs* 52 (1):3-23.

Goldberg, Sandy (2015). Epistemic Entitlement and Luck. *Philosophy and Phenomenological Research* 91 (2):273-302.

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

Goldman, Alvin (1979). What is Justified Belief? In George Pappas (ed.), *Justification and Knowledge*. Boston: D. Reidel. pp. 1-25.

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

Greco, Daniel (2014). Could KK Be OK? *Journal of Philosophy* 111 (4):169-197.

Grundmann, Thomas (2018). Saving safety from counterexamples. *Synthese* 197 (12):5161-5185.

Hawthorne, J. (2004). Knowledge and Lotteries. Oxford: Clarendon Press.

Hawthorne, J. and Rabinowitz, D. (2017). Knowledge and False Belief. In R. Borges, C. de Almeida, and P.D. Klein (Eds.), *Explaining Knowledge: New Essays on the Gettier Problem* (pp. 325-44). Oxford: Oxford University Press.

Klein, Peter (2008). Useful False Beliefs. In Quentin Smith (ed.), *Epistemology: New Essays*. New York: Oxford University Press. pp. 25-63.

Kvanvig, Jonathan L. (2004). Nozickian epistemology and the value of knowledge. *Philosophical Issues* 14 (1):201–218.

Lasonen-Aarnio, Maria (2008). Single premise deduction and risk. *Philosophical Studies* 141 (2):157-173.

Lasonen-Aarnio, Maria (2010). Unreasonable Knowledge. *Philosophical Perspectives* 24 (1):1-21.

Leitgeb, Hannes (2002). Timothy Williamson, knowledge and its limits. Oxford: Oxford university press, 2000. *Grazer Philosophische Studien* 65 (1):195-205.

Luper, Steven, "Epistemic Closure", *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2020/entries/closure-epistemic/>.

Manley, David (2007). Safety, content, apriority, self-knowledge. *Journal of Philosophy* 104 (8):403-423.

Murphy, Peter (2005). Closure failures for safety. *Philosophia* 33 (1-4):331-334.

Murphy, Peter (2006). A strategy for assessing closure. *Erkenntnis* 65 (3):365 - 383.

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

Nozick, R. (1981). *Philosophical explanations*. Oxford: Oxford University Press.

Pritchard, D. (2005). *Epistemic Luck*. Oxford: OUP.

Pritchard, D. (2012). Anti-Luck Virtue Epistemology. *Journal of Philosophy* 109 (3): 247-279.

Schulz, Moritz. (2020). Finding Closure for Safety. Episteme 18. 1–15.

Sosa, E. (1999). How to defeat opposition to Moore. In J. Tomberlin (Ed.), *Philosophical Perspectives 13: Epistemology* (pp. 141–154). Oxford: Blackwell.

Sosa, E. (2007). *A virtue epistemology. Apt belief and reflective knowledge.* Oxford: Oxford University Press.

Stalnaker, Robert (2006). On logics of knowledge and belief. *Philosophical Studies* 128 (1):169-199.

Warfield, Ted A. (2005). Knowledge from falsehood. *Philosophical Perspectives* 19 (1):405–416.

Weatherson, Brian (2004). Luminous margins. *Australasian Journal of Philosophy* 82 (3):373 – 383.

Willard-Kyle, Christopher (2020). Being in a Position to Know is the Norm of Assertion. *Pacific Philosophical Quarterly* 101 (2):328-352.

Williamson, Timothy. (2000). Knowledge and its limits. Oxford: OUP.

Williamson, Timothy (2009). Probability and Danger. *Amherst Lecture in Philosophy*.

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

Zagzebski, Linda (1994). The inescapability of Gettier problems. *Philosophical Quarterly* 44 (174):65-73.

Zhao, Haicheng (2020). Knowledge without safety. *Synthese* 197 (8):3261-3278.