Counterfactual Thinking and Thought Experiments

Winner of the Outstanding Graduate Paper Award at the
59th Annual Meeting of the Florida Philosophical Association

Josh Turkewitz, *Florida State University*

In the *Philosophy of Philosophy,* Timothy Williamson presents an account of modal reasoning where the capacities that produce and evaluate everyday modal knowledge epistemologically grounds more abstract counterfactual reasoning. Further, if thought experiments, a traditional source of purported philosophical knowledge, can be epistemologically grounded in this type of reasoning, then the knowledge gleaned from them can also be epistemologically grounded.

To that end, Williamson examines the logical structure of Gettier cases. Gettier cases are counterexamples to the justified true belief account of knowledge.[1]   In a Gettier case a subject obtains a justified false belief and, from this belief, justifiably infers a true proposition. The proposition is true and the subject is justified in holding it (since she justifiably inferred it from a justified belief) but does not constitute knowledge because it was inferred from a false premise. So having a belief that is both justified and true is not sufficient for knowledge. Gettier's famous 1963 paper demonstrates this by describing a series of thought experiments.

As part of Williamson's inquiry into how we gain knowledge from this type of thought experiment, he submits various ways of representing the argument underlying Gettier's thought experiment in modal and counterfactual terms. But all of these ways run afoul of the problem of deviance (i.e., that there are cases that might satisfy the descriptions given by a Gettier text but still fail to be counterexamples to the justified true belief model of knowledge). Problematically, this might mean that either it is too hard to know the truth of the premises of the arguments Williamson presents[2] or that the relevant premises might be false.[3] I argue that the Gettier-style arguments can make do with weaker premises (and a slightly weaker conclusion) that suffice to show that "necessarily, if one justifiably believes some true proposition p, then one knows p" is not true. The modified version of the argument is preferable because it is not troubled by the existence of deviant Gettier cases.

For Williamson "a thought experiment such as Gettier's embodies a straightforward valid modal argument for a modal conclusion."[4] The modal conclusion is that it is possible to have

justified true belief without knowledge. The Gettier text is a broad description of a situation in which individual knowers stand in a particular relation to supposedly propositional knowledge. Gettier cases are more specific; each case is a particular way of filling in the details that are not included in the Gettier text. The individual comes to believe justifiably some false proposition, and from that proposition they (using some reliable method, such as deductive inference) come to know some truth. Unfortunately, the individual's justified belief does not count as knowledge.   As Williamson writes:

> …if one is justified in believing something, and correctly deduces from it something else, one is justified in believing the latter proposition on that basis (deduction is a way of transmitting justification from the premises to the conclusion of an argument). Since any truth is deductively entailed by various falsehoods, one can believe a truth on the basis of having correctly deduced it from falsehood one is justified in believing, and thereby be justified in believing the deduced truth too…Nevertheless, one does not know, for one's belief in the truth, no matter how justified, is essentially based on a false lemma; one's conclusion cannot be epistemically better off than one's premises. Therefore, justified true belief is insufficient for knowledge.[5]

The principle that one's conclusion cannot be epistemically better off than one's premises, as Williamson notes, does not always hold.[6] Consider, for example, the following argument:

      (A)   All males are mortal.
      (B)   Socrates is male.
      (C)   Therefore, Socrates is mortal.

It might be that (C) is epistemically better off than (B). That is, one might be more sure that Socrates is mortal than that Socrates is male. Suppose there is some evidence that the person we refer to as 'Socrates' is actually Diotima of Mantinea, so (B) might very well be false. Nonetheless, we might still consider (C) to be epistemically unimpeachable, because no matter their biological sex, Socrates is most definitely mortal.

Fortunately, Gettier's argument does not rely on the principle that one's conclusion cannot be epistemically better off than one's premises. All that Gettier needs is for one case, one way of filling in the details of his story, to be a situation in which someone's justifiably believed (true) conclusion is not epistemically better off than her justifiably believed (false) premises. Since, in this instance, her premises do not count as knowledge (since they are false) and her conclusion is not

epistemically better off than her premises, she does not know her conclusion. Not every argument based on potentially shaky premises concludes that some person is mortal.

Gettier cases are counterexamples to the claim that necessarily, if one justifiably believes some true proposition p, then one knows p. This claim can be symbolized as follows:

$$(1)\ \Box\forall x\forall p(JTB(x,p) \rightarrow K(x,p))$$

Williamson understands Gettier's thought experiment in terms of the following argument (184-186):

$$(2)\ \Diamond\exists x\exists p\ (GC(x,p))$$
$$(3^*)\ \exists x\exists p\ GC(x,p) \ \Box\!\rightarrow\ \forall x\forall p(GC(x,p) \rightarrow (JTB(x,p)\&\sim K(x,p)))$$
$$(4)\ \therefore\ \Diamond\exists x\exists p\ (JTB(x,p)\&\sim K(x,p))$$

In English:

(2) "Someone could stand in the relation described by the Gettier story to some proposition" (184).

(3*) "If there were an instance of the Gettier case, then it would be an instance of justified true belief without knowledge" (186).

(4) Therefore, "someone could have justified true belief in some proposition without knowledge" (184).

Since (4) and (1) are inconsistent and the argument is valid, if (2) and (3*) are true then (1) is false. Williamson argues that we use the cognitive capacities that ground abduction, our predictive faculties and our counterfactual and modal reasoning, to establish the truth of (2) and (3*), and thus we come to know the falsity of (1). The Gettier story, on Williamson's account, provides support for both the minor premise (2)[7] and the major premise (3*).

Williamson also considers:

$$(1)\ \Box\forall x\forall p\ (GC(x,p) \rightarrow\ (JTB(x,p)\ \&\ \sim K(x,p)))$$

However, he rejects this formulation because modal space may include deviant versions of the Gettier story that falsify the above conditional.[8] These versions might not occur to us when we

are assessing a purported Gettier counterexample and might be ruled out by the details of a particular Gettier story, but GC(x,p) above is a general schema that may or may not rule out such deviant instances. It might be that there is a possible world that satisfies GC (e.g., a case in which someone reliably infers some true proposition from a justifiably believed false proposition) but falsifies (3). Consider the following deviant Gettier case: Bob is driving past a field in Wisconsin. Out of the window Bob sees a lifelike sheep statue. Bob comes to believe that there is a sheep in the field and infers that there are sheep in Wisconsin. Bob's inferred belief is true and he is justified in holding it, but (in ordinary circumstances) it does not count as knowledge, because it was inferred from a false (though justified) belief.[9]

Now imagine that the circumstances are not ordinary. Suppose that Bob knew that the countryside was dotted with sheep statues, but he also knew that it was a U.S. custom to keep at least one sheep within 100 meters of every sheep statue. Bob might now know that there are sheep in Wisconsin (even though he thinks he saw one and in fact did not). In these circumstances, (3) would be false in virtue of having a true antecedent and a false consequent. Bob stands in the relation described by the Gettier text to the belief that there are sheep in Wisconsin and, yet, he knows that there are sheep in Wisconsin.

Williamson writes, "instead of asking whether justified true belief without knowledge is a necessary consequence of the Gettier case, one might more naturally ask whether, if there *were* an instance of the Gettier case, it *would* be an instance of justified true belief without knowledge."[10] For this reason, he prefers (3*) because "in very rough terms, it requires justified true belief without knowledge only in the closest realizations of the Gettier case, not in all possible realizations."[11]

However, even this weaker formulation of the major premise is problematic. As Williamson admits, "one might worry that the counterfactual claim overstates the Gettier intuition."[12] If the only individuals who realized the Gettier story in the actual world were deviant such that they did not have a justified true belief or did have knowledge, or if the actual world contained no instances of the Gettier story and the nearest worlds containing an instance of the Gettier case were deviant, then (3*) would be false. That is, if the deviant case of Bob and the sheep was the only actual Gettier case, then the case of Bob and the sheep would be the closest realization of a Gettier case. Similarly, it might be that the worlds containing Bob and the sheep (and no other Gettier cases) were closest to the actual world, such that all of the closest realizations of Gettier cases were deviant. In these worlds (3*) is false and, thus, Gettier's argument is unsound. Yet, one might think the success or failure of the Gettier argument should not depend on one's location in modal space. I return to this suggestion below.

Jonathan Ichikawa objects further.[13] We cannot know, he argues, whether or not we are in such a deviant position in modal space, so we cannot know the counterfactual (3*). Ichikawa's

worry seems to be that since the Gettier text is so broad and can be satisfied in so many deviant ways, it might be difficult (if not impossible) to know whether the (modally) nearest scenario that satisfies the Gettier text is deviant or not. If it is, then (3*) is false. So it is difficult (if not impossible) to know that (3*) is true. However, it is not difficult to know what Ichikawa calls the 'Gettier intuition:' that Gettier cases are cases of justified true belief without knowledge. For Ichikawa (3*) does not adequately capture the Gettier intuition because it is difficult to know if (3*) is true, but it is not difficult to know that the Gettier intuition is true.

I do not think this is a very telling objection. We might, for example, know of (or create) non-deviant Gettier cases in the actual world. If we do, then we know (without much difficulty) that we are not situated oddly in modal space with respect to (3*). Williamson's first worry, that the counterfactual itself might be false because there are deviant Gettier cases, is more pressing than the epistemic concern Ichikawa raises because, if Gettier is correct, (1) is false in all possible worlds. If the formalization of his argument is not able to demonstrate this then the formalization needs revision. For Williamson, we deal with these deviant scenarios on a case-by-case basis. When we identify deviant ways of realizing the Gettier text we modify the text to exclude those cases.

Ichikawa's solution is to offer a different analysis of the Gettier cases. He interprets the Gettier text as providing a general schema for the creation of Gettier cases. When considering such a text, a philosopher fills in the details in some determinate way (creating their own imagined, private instance of a Gettier case – a Gettier story) such that this imagined story is (without fail) an example of justified true belief without knowledge. For Ichikawa, something similar to (3) is preferable, and the problem of deviance is not handled by the logical form of the argument, it is instead ruled out by the highly detailed relation described by the open sentence 'GC(x,p).'

Both of these solutions have merits and demerits. Williamson's solution might still be vulnerable to Ichikawa's complaint that it is too hard to know the counterfactual (3*). Furthermore, (3*) still requires that a genuine instance of a Gettier story occur somewhere nearby in modal space. But this requirement is too strong. Any possible instance of justified true belief without knowledge suffices as a counterexample to (1). Regarding Ichikawa's proposal, as Williamson notes, the more explicit we become in describing a given Gettier case the less confident we might become in asserting (2).[14]

Ichikawa considers an objection raised by Williamson that his view entails that Gettier arguments are private because "everyone engages with a Gettier case via her own private scenario in her own mind."[15] Ichikawa does not find this particularly problematic:

[the] view is not that subjects should fill out thought experiments in whatever way they like; we have particular conventions, grounded in our practices with fictions, which govern how to move

from a weaker description to a stronger scenario in the intended way… Perhaps there is something less than ideal about this state of affairs, but, in my view, this added moving part is just a part of philosophy.[16]

However, I think Ichikawa is overly dismissive of this objection. Consider the following Gettier story: A dark-haired woman is running in the rain, hurriedly glancing down at her soaked silver wristwatch. The hands read 8:28. The engine of her bus, waiting at the station, coughs to life as she stumbles around the corner. Catching her balance and cursing the downpour, she lunges through the open set of double-doors. They close behind her, barely missing the hem of her coat. She sags into the worn plastic seat and sighs in relief, clutching her trusty wristwatch, which lets her know that she could catch the 8:30 bus if she hurried. Her wristwatch, shorted out by yesterday's rain, still reads 8:28. Unbeknownst to her, it is now 8:35 – the driver was having trouble with his engine and left the last stop five minutes behind schedule.

The details of the above story appear to be consistent with one another, so there is no *prima facie* reason to think the scenario is impossible.[17] However, each of us will render the story slightly differently. The scene might take place in the morning or the evening or the woman might have black hair. The bus could appear as a streetcar, a cable car, or a school bus. The rain may be cold or warm, seasonal or unseasonal. She might be in a city or a suburb, and so on.

What is significant about these private aspects of the story is that they are irrelevant to the important features of the case – the features that make it a Gettier case (e.g., her wristwatch being a reliable source of beliefs about the current time, her looking at it and coming to believe a false proposition (that it was 8:28), her inferring a true proposition from this belief (that she could catch her bus if she rushed)). It is important that these features of the case be are not private, as they are the philosophically interesting features.

Ichikawa might think that any successful Gettier story would be sufficiently clear about these details to keep misunderstandings to a minimum. In contrast, instead of specifying the story at hand to rule out deviant cases it would be better to include only the salient features of the Gettier story in the minor premise, while letting readers fill in the details as they may.

Ideally, the Gettier argument would accept a very general schema for GC(x,p) – such as the Gettier text – and still generate a counterexample to (1). To this end, I propose the following:

$$\text{(2) } \lozenge\exists x\exists p \, (GC(x,p))$$

$$\text{(3') } \square\forall x\forall p \, (GC(x,p) \supset \lozenge \, (JTB(x,p) \,\&\, {\sim}K(x,p)))$$

$$\text{(4'):. } \lozenge\exists x\exists p \, (\lozenge(JTB(x,p) \,\&\, {\sim}K(x,p)))^{18}$$

The major premise (3') says that necessarily, for any person and true proposition p, if that person stands in the relation described by the Gettier text to that proposition then it is possible that she justifiably believe p and does not know p. (3') is entailed by (3) and (3*), so a philosopher who accepts either (3) or (3*) should accept (3').

The Gettier text is a blueprint that can be filled out in various ways, each of which is a Gettier case. In some of these cases the subject has justified true belief but not knowledge. The general blueprint implies that (at least) one of these ways is possible. This is just what (3') says – if I am related to a true proposition in the way described by the Gettier text, then it is possible that I justifiably believe that proposition without knowing it. If there are no consistent, non-deviant Gettier stories, then the Gettier text straightforwardly fails to generate a counterexample to (1). Denying (3') is tantamount to denying the success of the Gettier text.

(4') states that it is possible that there exist a person x and a true proposition p such that it is possible that x be justified in believing p without knowing p. (4') is weaker than (4), but – given the modal system S4 – (4') is inconsistent with (1).[19] (4') is a bit less wieldy than (4), but the inconvenience is bearable.[20]

Importantly, (3') is not falsified by the existence, actual or possible, of deviant Gettier stories. Suppose that we are situated in modal space such that the only instance of a Gettier story in any nearby world is deviant. So long as, given the possibility of the instantiation of the Gettier text, there is some possible world in which a subject has a justified true belief without knowledge (3') is true. Thus, this formulation of the argument avoids the concerns raised by Ichikawa and Williamson.

Not every case that satisfies the Gettier text will be an example of justified true belief without knowledge. It is not important that each case (or that every nearby case in modal space) be a counterexample to (1). All that is required is that a counterexample to (1) exists. Such is the nature of counterexamples – one is all you need.

Objection. Ichikawa could argue that my account collapses into his. That is, Ichikawa might insist that we verify (3') by imagining a very specific Gettier story, one that rules out many deviant cases. Since we might judge this story is possible and it is determinately a case of justified true belief without knowledge, we know that (3') is true.

Reply. We might come to know (3') by knowing that a specific, highly detailed story of justified true belief without knowledge is possible. However, we need not do this. In part, I think this is because any story, no matter how detailed and specific, will never be able to rule out every deviant instance. Regardless of the level of specificity provided, we could not know that a given story is *definitely* a non-deviant case. Nonetheless, we can know the truth of (3') without knowing

which particular story is non-deviant. We might come to know the truth of (3') by considering the Gettier story abstractly: there is a possible world in which someone justifiably comes to believe a false premise. From this premise, she infers (through a justified belief forming process) something true. But she might not have knowledge, precisely because the truth was inferred from a false belief. These are the philosophically salient details of Gettier cases; they are what let us know that (1) is false. Further details may be of some literary or philosophical interest. They might make one's writing more persuasive or relatable, or they might help better define the boundary between deviant and non-deviant Gettier cases, but these aims are not essential. By insisting that they are, Ichikawa needlessly muddies the waters.

Recall the Gettier story of the woman in the rain. In this story, she heard the bus's engine cough to life as she was approaching the stop. This might have provided evidence (independent from the evidence provided by her wristwatch) for her justified true belief that she could catch the bus if she hurried, and so her conclusion (that she could catch the bus if she hurried) might have been epistemically better off than her justified false belief (that it was 8:28). It is unimportant how one might modify the original story to remove this ambiguity.[21] Insisting on making such modifications would be failing to extract the philosophically salient features in favor of stating the case with unnecessary precision.

This precision might be required if we are trying to find the precise boundary between knowledge and justified true belief (Was the combined evidence provided by both her wristwatch and the sputtering engine enough for the woman to know that she could catch her bus if she hurried or not?). Yet, it is not required for us to know what we need to know to refute the justified true belief account of knowledge – that truths can be justifiably inferred from falsehood, and so we can be on shaky epistemic ground even when justifiably believing truths. I think (3') lends itself to a strong response to the demand for greater precision: 'despite the lack of precision one can spell out these details in non-deviant ways Since the bare fact of this possibility refutes (1) without knowing precisely which cases are non-deviant, we should move on with the discussion.' One might think I am unjustified in asserting this, but I rather think the onus is on the objector to show that (surprisingly) there are no possible non-deviant Gettier cases. The objector, I suspect, would only be able to do this by discussing the general features of Gettier cases, and finding some incoherence. They could not check every highly specific story for incoherence, as that would take many lifetimes.

Objection. Williamson could argue that my account collapses into his. That is, Williamson might claim that we know (3') by knowing the counterfactual (3*), which entails (3').

Reply. This is all well and good; if we are able know (3*). But, as Ichikawa argues, it is not clear that we are able to know (3*). Furthermore, (3*) is false in some possible worlds. As we have seen, it is not the case that were any Gettier case to obtain, it would be a case of justified true belief

without knowledge, because we could be surrounded by deviant Gettier cases in modal space. But (3') is not false in any possible world, so long as there is at least one non-deviant Gettier case out there in modal space. So we cannot always know (3') by way of (3*), because (3*) is false in some worlds, and in those worlds (3') is not false. It might be that we come to know (3') by using the cognitive capacities Williamson describes to gain all our modal knowledge, but this alone does not imply that the epistemology of (3') reduces to the epistemology of (3*) any more than it implies that the epistemology of any modal truth reduces to the epistemology of any other modal truth.

Objection. I have admitted that my argument requires S4, and someone who believes in the success of the Gettier case does not thereby commit herself to the validity of S4. Williamson raises a similar objection to a different weakening of (3*).[22]

Reply. Any formalization of an argument requires an axiomatic system, and it is not the case that everyone who believes the conclusion of that argument need accept (or even be aware of) said axiomatic system. In general, I am skeptical of claims that a particular method is the only way to come by a given bit of knowledge. A philosopher who accepts S4 might accept my formalization, and thereby avoid the problems inherent in Williamson's or Ichikawa's formalizations, by paying the price of having more substantial commitments to a particular modal logic. A philosopher unwilling to do so will incur the philosophical cost of having to address problems Williamson and Ichikawa are grappling with. Both routes might be paths to the knowledge that (1) is false.

Williamson's and Ichikawa's formulations of the Gettier argument are vulnerable to the possibility or actuality of deviant Gettier stories. To avoid these vulnerabilities, I proposed an alternative formulation of the Gettier argument; one that better captures the ways in which we reason (and the ways in which we ought to reason) about Gettier cases. I then argued that the epistemology of this formulation does not collapse into either Williamson's or Ichikawa's account of the epistemology of the relevant premises. I do not think that there is anything devastatingly suspect with these epistemologies, they both might be adequate routes to knowledge of the relevant premises in my formulation or to the knowledge that justified true belief is insufficient for knowledge, but they are not the best routes to such knowledge.

Appendix

Proof of (4') from (2) and (3'):

◊∃x∃p (GC(x,p))	Premise
□ ∀x∀p (GC(x,p) ⊃ ◊ (JTB(x,p) & ~K(x,p)))	Premise
∃x∃p (GC(x,p))	Possibility Elimination
∀x∀p (GC(x,p) ⊃ ◊ (JTB(x,p) & ~K(x,p)))	Reiteration
GC(y,q)	Existential Instantiation
GC(y,q) ⊃ ◊ (JTB(y,q) & ~K(y,q))	Universal Instantiation
◊ (JTB(y,q) & ~K(y,q))	Modus Ponens
∃x∃p (◊(JTB(x,p) & ~K(x,p)))	Existential Generalization
◊∃x∃p (◊ (JTB(x,p) & ~K(x,p)))	Possibility Introduction

QED

Proof that (4') is inconsistent with (1):

◊∃x∃p (◊ (JTB(x,p) & ~K(x,p)))	Premise
□∀x∀p(JTB(x,p) ⊃ K(x,p))	Premise for r.a.a.
□□∀x∀p(JTB(x,p) ⊃ K(x,p))	□p⊃□□p (S4 axiom)
∃x∃p (◊ (JTB(x,p) & ~K(x,p)))	Possibility Elimination
□∀x∀p(JTB(x,p) ⊃ K(x,p))	Reiteration
◊ (JTB(y,q) & ~K(y,q))	Existential Instantiation
JTB(y,q) & ~K(y,q)	Possibility Elimination
∀x∀p(JTB(x,p) ⊃ K(x,p)	Reiteration
JTB(y,q) ⊃ K(y,q)	Universal Instantiation
K(y,q) & ~K(y,q) --><--	Basic Propositional Logic
~□∀x∀p(JTB(x,p) ⊃ K(x,p))	Negation Introduction

Endnotes

---

[1] Gettier 1963, 121.

[2] Ichikawa 2009.

[3] Williamson 2007.

[4] Williamson 2007, 187.

[5] Williamson 2007, 181.

[6] Williamson 2007, 181-182.

[7] The minor premise is also supported by the existence of actual Gettier cases. However, the Gettier text might provide evidence that (2) is true regardless of whether or not there were actual Gettier cases.

[8] I take a 'deviant Gettier story' to be a case that satisfies the Gettier text but is not an example of justified true belief without knowledge.

[9] Supposing that Bob has no other evidence for the claim that there are sheep in Wisconsin.

[10] Williamson 2007, 185.

[11] Williamson 2007, 186.

[12] Williamson 2007, 200

[13] Ichikawa 2008.

[14] Williamson 2007, 189 and Williamson 2009, 468.

[15] Ichikawa 2009, 442.

[16] Ichikawa 2009, 442-443.

[17] Providing reason to think (2) is true.

[18] See appendix for proof.

[19] See appendix for proof.

[20] Given S5, (4) can be inferred from (4') straightforwardly, but this extra step is unneeded.

[21] I think one would rightly be annoyed by an interlocutor who continued to press for such details.

[22] Williamson 2007*, 202*

Bibliography

Gettier, Edmund. "Is Justified True Belief Knowledge?" *Analysis* 23.6 (1963): 121.

Ichikawa, Jonathan. "Knowing the Intuition and Knowing the Counterfactual." *Philosophical Studies* 145.3 (2009): 435-43.

Williamson, Timothy. *The Philosophy of Philosophy*. Malden: Blackwell Pub., 2007.

_____. "Replies to Ichikawa, Martin and Weinberg." *Philosophical Studies* 145.3 (2009): 465-476.