Judgments about Thought Experiments

Alexander Geddes

Thought experiments invite us to evaluate philosophical theses by making judgments about hypothetical cases. When the judgments and the theses conflict, it is often the latter that are rejected. But what is the nature of the judgments such that they are able to play this role? I answer this question by arguing that typical judgments about thought experiments are in fact judgments of *normal counterfactual sufficiency*.

I begin by focussing on Anna-Sara Malmgren's defence of the claim that typical judgments about thought experiments are mere possibility judgments. This view is shown to fail for two closely related reasons: it cannot account for the incorrectness of certain misjudgments, and it cannot account for the inconsistency of certain pairs of conflicting judgments. This prompts a reconsideration of Timothy Williamson's alternative proposal, according to which typical judgments about thought experiments are counterfactual in nature. I show that taking such judgments to concern what would *normally* hold in instances of the relevant hypothetical scenarios avoids the objections that have been pressed against this kind of view. I then consider some other potential objections, but argue that they provide no grounds for doubt.

Thought experiments are widespread in philosophy. Hypothetical situations are described and judgments are made about them, often on the basis of mere reflection. When these judgments appear to conflict with philosophical theories, it is common to respond by rejecting or amending the theories rather than by revising the judgments. Indeed, this is one of the main ways in which philosophers evaluate their views. But how are we to make sense of this process? In particular, what is the nature of these judgments such that they are able to play this role?[1]

Anna-Sara Malmgren (2011) responds to this question by arguing that intuitive judgments about thought experiments are metaphysical possibility judgments whose contents straightforwardly contradict the modal implications of the theses

---

[1]There is disagreement about the psychological and evidential bases of such judgments. In particular, some take there to be certain psychological states or episodes ('intuitions') the occurrence of which we typically rely on for evidence when making judgments about hypothetical cases (e.g. Brown 2011; Goldman 2007; Nagel 2012; Pust 2000). Others are sceptical (e.g. Cappelen 2013; Deutsch 2015; Williamson 2007b). The discussion below remains neutral on this issue. Note that the possibility of classifying certain judgments as intuitive (on some basis or other) remains even if one agrees with the sceptics. Accordingly, any reference to intuitive judgments below should be understood in a non-theoretically loaded way.

they lead us to reject.[2] In doing so, she aims to defend the view that such judgments are justified *a priori*, if at all, from some arguments of Timothy Williamson's (2005; 2007a; 2007b). In what follows, I show that Malmgren's proposal fails and develop a variant of Williamson's own proposal, according to which judgments about thought experiments are counterfactual in nature.

First, I outline Malmgren's route to her view, identify two problems that arise concerning misjudgments, and rule out some possible responses. (§§1–2) I then make the case that these problems result from a failure to recognise that when engaging with thought experiments we are primarily concerned with what the relevant hypothetical scenarios suffice for. (§3) This brings into sharper focus what is right in Williamson's proposal, and I show how Malmgren's criticisms of his view can be avoided by appealing to the notion of *normal counterfactual sufficiency*.[3] (§4) I then consider two objections to this view, in response to which it is fleshed out in greater detail and provided with some extra support. (§5) I end with some brief remarks concerning the bearing that this has on the apriority of such judgments. (§6)

1 PROPOSALS

Following Malmgren, let us focus on the following variation on one of the original Gettier (1963) cases:

> … Smith believes that Jones owns a Ford, on the basis of seeing Jones drive a Ford to work and remembering that Jones always drove a Ford to work in the past. From this, Smith infers that someone in his office owns a Ford. … [F]urthermore … someone in Smith's office does own a Ford—but it is not Jones, it is Brown. (Jones's Ford was stolen and Jones now drives a rented Ford.) (Malmgren 2011, p. 272)

Call this 'the Gettier Case'.[4] There is a particular judgment about this case that many

---

[2]There are (at least) two kinds of thought experiment, only one of which is the focus of the present paper. One kind presents a hypothetical case and invites the thought experimenter to make a judgment about whether or not the case is possible. (Think zombies and inverted spectra.) Such judgments are manifestly metaphysical possibility judgments. Another kind, however, presents a hypothetical case, *presupposes* that the thought experimenter accepts its possibility, and invites him or her to make a judgment about what else holds in the case. (Think Gettier cases, trolley problems and brain transplants.) It is the latter with which we are presently concerned.

[3]Gardiner (2015) has independently developed some related objections to (Malmgren 2011), and likewise suggests invoking normalcy in an account of judgments about thought experiments. Her paper came to my attention too late to include any detailed discussion of the differences between our approaches. Here I simply note that she neither offers precisely the same diagnoses of where other views go wrong as are offered below, nor develops her positive proposal specifically as a version of the counterfactual view.

[4]I use 'case', 'scenario' and 'situation' interchangeably throughout to speak of ways for some things to be, or to stand to each other: a kind of general object that I take to be specified by descriptions such as the one just given when they are put to use in thought experiments. And I use 'the Gettier Case' solely as a label for the case that involves all and only those features made explicit in the description above. (Note that, as is standard, 'thought experiment' is in fact used ambiguously: sometimes to refer to procedures that make use of cases, sometimes to refer to cases themselves. In what follows, I rely

make if prompted. A common way of expressing the judgment in question is via the following sentence:

(S) Smith has a justified true belief, but does not know, that someone in his office owns a Ford. (p. 272)

Call the judgment thereby expressed 'the Gettier Judgment'.[5]

As Malmgren points out, it would be a mistake to think of this as a judgment to the effect that some particular individual, picked out by 'Smith', has the property of justifiably and truly believing, but not knowing, that someone in his office owns a Ford. There is no such individual, and expressing the judgment in this way need involve no confusion on that point. This presents us with an instance of what she calls 'the content problem' (p. 267). Roughly speaking, the problem is that we typically voice judgments about hypothetical situations in just the same ways as we would if those situations were real, and yet the relevant judgments do not plausibly have contents of the same kind. So what *is* the content of the Gettier Judgment?

Any acceptable answer to this question must allow us to make sense of the role that this judgment plays in our philosophical theorising, namely, leading us to reject the thesis that knowledge is justified true belief. Presumably, it does so in virtue of having a modal significance that conflicts with the claim that, necessarily, someone knows a proposition p if and only if he or she has a justified true belief that p. One straightforward way to satisfy this requirement, then, is simply to identify the Gettier Judgment's content with a suitable modal proposition, and indeed this is precisely what Malmgren does.[6]

The three main proposals she considers are:

(□) Necessarily, anyone who stands to a proposition p as in the Gettier Case has a justified true belief that p but does not know that p.

(□→) If someone were to stand to a proposition p as in the Gettier Case, then he would have a justified true belief that p but not know that p.

(◊) It is possible that someone stands to a proposition p as in the Gettier Case and that he has a justified true belief that p but does not know that p.[7]

---

[5]An anonymous referee suggested that one might reject the Gettier Judgment on the basis that, in this case, Smith's stated evidence fails to justify his belief that Jones drives a Ford. But one may alter the case by enriching his evidence somewhat and, *mutatis mutandis*, all of the points below will stand. (That is, as long as one does not stipulate that his belief is justified; see §5.1.)

[6]One might doubt whether modal propositions of the kinds identified below could plausibly be the contents of judgments about thought experiments. One might also doubt whether we use sentences such as (S), in the context of thought experiments, to give voice to genuine judgments at all. (Cf. Yablo 2001 and Eagle 2007.) In either case, one can interpret most of what follows as addressing the following, less theoretically loaded question: in performing the relevant mental act given voice to via (S), what modal commitment is incurred?

[7]See pp. 274, 277 and 281. I have substituted 'the Gettier Case' for 'the Gettier case (as described)'. I have also substituted 'he' for 'she', on the basis that the original case description and (S) both employed the phrase 'his office'.

on context to determine its sense on each occasion of use.)

These can be thought of, respectively, as ways of making a little more precise the rough but straightforward suggestions that the Gettier Judgment concerns what *must* hold, what *would* hold, and what *could* hold, in an instance of the Gettier Case.

Now, the problem for ($\Box$), pointed out by Williamson (2005, pp. 6–7; 2007b, p. 185) and endorsed by Malmgren (2011, pp. 275–277), is that the Gettier Case description is rather limited in detail. As a result, it can be consistently elaborated so as to reverse the impression that Smith's belief is justified. For example, it is consistent with the case description above that Smith has been told that Jones is renting, but that it momentarily slips his mind when he infers what he does. Someone could therefore stand to a proposition as in the Gettier Case and yet not have a justified true belief without knowledge. This entails that ($\Box$) is false. But the Gettier Judgment is correct. So ($\Box$) cannot be its content.[8]

($\Box\!\!\rightarrow$) is Williamson's alternative. Against it, Malmgren builds on the previous argument as follows (pp. 278–281). If someone *were* to be related to a proposition as in the Gettier Case while lacking justified-true-belief-without-knowledge, then ($\Box\!\!\rightarrow$) would be false. But the Gettier Judgment would still be correct. That is to say, our semantic intuitions seem to tell us that such 'deviant' instances of the case have no bearing on the correctness of the judgment, no matter what their proximity in modal space. So ($\Box\!\!\rightarrow$) cannot be the Gettier Judgment's content either.[9]

Malmgren's solution is to retreat to ($\Diamond$). This avoids the problems above, as neither the possibility nor the actuality of apparently 'deviant' instances of the Gettier Case can suffice to falsify ($\Diamond$). She goes on to address a number of objections, to which her responses are largely convincing (pp. 281–306). Taking the Gettier Judgment to be paradigmatic, then, she concludes that all of our typical judgments about thought experiments are metaphysical possibility judgments of roughly the same form.[10]

There are, however, serious problems for this view that Malmgren fails to consider, and to which I now turn.

---

[8]See §5.1 for discussion of some potential responses to this argument.

[9]Ichikawa and Jarvis (2009, pp. 225–226) present a similar objection.

[10]Technically, she generalises only to other *intuitive* judgments about thought experiments. But this restriction has no bearing on the issues discussed in this paper, for two reasons. First, given a plausible, non-factive characterisation of what renders a judgment intuitive, all of the judgments considered below could be made intuitively. (Malmgren herself (p. 268) identifies intuitive judgments as those 'relevantly similar to' certain paradigms, including the Gettier Judgment; in her (2013) she provides a more explicit, albeit tentative, characterisation. Neither gloss, I take it, implies that intuitive judgment is factive.) Second, the purpose of the restriction is to render the claim that such judgments are justified *a-priori*-if-at-all at least *prima facie* plausible. (The fact that judgments about thought experiments, like most judgments, can be justified on a variety of bases would render a parallel claim without the restriction to intuitive instances obviously false.) But while this restriction may have a place when it comes to a debate over justification, it has none whatsoever when it comes to a debate over content or modal significance.

2   PROBLEMS

Let us examine, alongside the Gettier Judgment, another judgment that might be made about the case. A typical way of expressing the judgment in question would be via the following sentence:

(S*)   Smith neither has a justified true belief, nor knows, that someone in his office owns a Ford.

Call this 'the Gettier Misjudgment'. As the name suggests, the judgment in question is incorrect. But it is nevertheless a judgment that some make when presented with the case. So what is its content?

   The most obvious way of applying Malmgren's proposal to the Gettier Misjudgment would be to assign to it the following:

($\Diamond$*)   It is possible that someone stands to a proposition p as in the Gettier Case and that he neither has a justified true belief that p nor knows that p.

Problems with assigning ($\Diamond$) and ($\Diamond$*) to the judgments expressed via (S) and (S*), however, arise immediately. I will identify two.

   The first problem is this. While the Gettier Judgment is correct, the Gettier Misjudgment is, as already noted, incorrect. The vast majority of epistemologists writing since Gettier have agreed on this. We are liable to think that those who utter (S*) have made some kind of mistake. But this difference in correctness is not reflected in the truth values of the candidate contents. Both ($\Diamond$) and ($\Diamond$*) are true. (Recall that the truth of ($\Diamond$*) was the ground for rejecting ($\Box$).) On the face of it, then, Malmgren's view would count both judgments as correct. And this is clearly the wrong result.

   There are, in principle, two ways in which Malmgren could respond. Either she could accept the content assignment suggested above, and attempt to explain away the apparent incorrectness of the Gettier Misjudgment. Or alternatively, she could deny that ($\Diamond$*) is the content that she would assign to it. However, both options are unacceptable, for the following reasons.

   Regarding the former, note that the semantic appearances persist despite explicit knowledge of the truth of ($\Diamond$*). Why, then, does the Gettier Misjudgment continue to strike us as incorrect? No plausible explanation seems to be forthcoming. But more significantly, any such explanation would only serve to undermine the case presented against ($\Box$) and ($\Box$→). That case relied on our semantic intuitions to tell us that the truth of ($\Diamond$*) does not render the Gettier Judgment incorrect. But our semantic intuitions also tell us that the truth of ($\Diamond$*) does not render the Gettier *Mis*judgment *correct*. And there is no good reason to trust our intuitions in the one case while doubting them in the other.

   The latter option holds that ($\Diamond$*) is not the content of the Gettier Misjudgment, even though ($\Diamond$) is the content of the Gettier Judgment. But then what is its content to be? To rescue the view, it would have to be a false metaphysical possibility claim

of some sort. But it is doubtful that any other metaphysical possibility claim in the vicinity would be as plausible a candidate as (◊*). And even if some such candidate could be found, it is likely, given the point above concerning the many ways in which the Gettier Case can be consistently elaborated, that it would turn out to be true as well.

Whatever the prospects for finding such an alternative, however, it would do nothing to address the second problem, which builds on the first as follows. It is not merely the case that the Gettier Judgment and Misjudgment differ in correctness. They are also clearly in tension. Two subjects who assert them in turn are manifesting disagreement about the Gettier Case, where such disagreement strikes us a matter of their giving voice to inconsistent judgments. Yet if the judgments expressed have metaphysical possibility claims as contents, then the subjects are not disagreeing in the manner envisaged. This is so for the simple reason that metaphysical possibility claims—claims of the form $◊p$—*cannot* be inconsistent.[11] So Malmgren's view entails that the judgments in question are not in tension after all. Again, though, this is clearly the wrong result.

It is worth pointing out that even if each of the proposals mentioned in §1 generates a conflict with some of our semantic intuitions concerning correctness, this second problem is unique to (◊). If (□) is the content of the Gettier Judgment, then the content of the Gettier Misjudgment should be the claim that, necessarily, anyone who stands to a proposition p as in the Gettier Case has neither a justified true belief that p nor knows that p. If (□→) is the content of the Gettier Judgment, then the content of the Gettier Misjudgment should be the claim that if someone were to stand to a proposition p as in the Gettier Case, then he would neither have a justified true belief that p nor know that p. Of course, strictly speaking, each of these pairs is consistent. But the Gettier Judgment and the Gettier Misjudgment both take place against the background assumption that an instance of the Gettier Case is possible. (Cf. n. 2 above.) So on either view, the vacuous truth of the relevant pair of contents is precluded, and the mutual inconsistency of the judgments is thereby ensured.

Now, Malmgren may object to all of this that her proposal was never intended to have any implications for the content of the Gettier Misjudgment. After all, her stated aim, in the service of which she presents the proposal, is to provide an account of a 'core set' of 'successful thought experiments — experiments that, other things equal, … provide the thought experimenter with … a justified and reliably true belief that the theory under evaluation is false' (p. 272). And while the Gettier Judgment provides us with a justified and reliably true belief that the justified-true-belief theory of knowledge is false, the Gettier Misjudgment does nothing of the sort. So Malmgren may wish to claim that although (S) is used to express a judgment of metaphysical possibility, the judgment expressed via (S*) could have a content of an entirely different form.

However, the fact that her proposal was intended to cover only a restricted range of judgments does not mean that it has no implications for judgments outside of

---

[11]Qualification: as long as p itself is neither a modal proposition nor a contradiction.

that range. Notwithstanding differences in correctness or usefulness, the Gettier Judgment and Misjudgment have a manifest similarity that makes it inappropriate to treat them as having contents of radically dissimilar form. Were one to do so, the explanations of how relevant utterances of (S) and of (S*) manage to express the judgments they do would have to proceed quite differently in each case. But given how syntactically similar the sentences are, these divergent explanations would be implausibly unsystematic. In other words, if there is no way to treat both (S) and (S*) as expressing metaphysical possibility judgments, neither should be so treated. Malmgren's proposal must therefore be rejected.[12]

These issues are not particular to the judgments discussed above. Nor are they particular to the way in which Malmgren fleshes out the basic idea that judgments about thought experiments are metaphysical possibility judgments. On any such view, parallel problems will arise for many pairs of seemingly conflicting judgments. We must therefore look elsewhere for a correct account of their content.

3   SUFFICIENCY

I suggested above that there was a single issue underlying the problems just outlined. It is this. When constructing descriptions of hypothetical situations, we do so with an eye to identifying cases which strike us as sufficing, non-stipulatively, for the presence of some philosophically significant feature. And when making judgments about such cases, we are typically in the business of assessing whether or not they in fact do so. Our practice of engaging in thought experiments, then, is best construed as a search for sufficiency.[13] And Malmgren's proposal faces the difficulties it does precisely because it fails to reflect this fact.

This can be brought out most clearly by considering a relatively minimal case description, such as the following impoverished version of the Gettier Case:

> Smith* believes that Jones owns a Ford. From this, Smith* infers that someone in his office owns a Ford.

Now, there are a vast number of properties which, it strikes us as obvious, it is possible for someone in Smith*'s position to instantiate. But for any such property,

---

[12]It might be suggested that the Gettier Misjudgment could have an *im*possibility claim as its content, and that this could restore incorrectness/inconsistency without departing too drastically in form from (◊). The most plausible option would presumably be something like: it is not possible that someone stands to a proposition p as in the Gettier Case and that it is not the case that he has neither a justified true belief that p nor knows that p. But the objection in the text still stands: whence the difference in form between the judgments expressed via (S) and (S*)? Where have all these negations come from? Moreover, this is logically equivalent to: necessarily, anyone who stands to a proposition p as in the Gettier Case has neither a justified true belief that p nor knows that p. What plausible explanation could there be of how or why these judgments are of such different strengths?

[13]Note that the claim here is not that, when engaging in thought experiments, we do so primarily in order to *assess* sufficiency claims, such as the claim that justified true belief suffices for knowledge. (Though this we do.) The claim is that typical judgments about hypothetical cases are *themselves* judgments about what those cases suffice for.

picked out by Φ, we are not thereby inclined to make or agree with the judgment that would be expressed via ⌜Smith⋆ is Φ⌝. This is so even when an opposing judgment also fails to strike us as correct. To see this, just consider the judgments that would be expressed via the following sentences:

> Smith⋆ has a justified [/unjustified] belief that someone in his office owns a Ford.[14]
>
> Smith⋆ has a true [/false] belief that someone in his office owns a Ford.
>
> Smith⋆ knows [/does not know] that someone in his office owns a Ford.
>
> Smith⋆ works [/does not work] in a large office.
>
> Smith⋆ is [/is not] over six feet tall.
>
> …

If the judgments in question concerned possible co-instantiation, then our inability to see any truth in them would be inexplicable. But the reason why none strikes us as correct is simply this: the hypothetical situation has not been filled out in enough detail to warrant any of them. We look to the explicit details of a case description to provide us with a reason for taking other, further features to be in some sense (and however unintentionally) present. And the impoverished Gettier case fails to do this with respect to any of the properties predicated of 'Smith⋆' above. But this is just to say that our interest is in working out what the case described suffices for.

With this in mind, we can see why Malmgren's view generated the problems it did. A genuinely possible hypothetical scenario may be consistent with the presence of either one of two incompatible features, but it cannot suffice for the presence of two incompatible features. If it suffices for the presence of one such feature, it must fail to suffice for the presence of the other. It is precisely because the Gettier Judgment and the Gettier Misjudgment both concern what the (non-impoverished) Gettier Case suffices for, then, that they are inconsistent. And it is because the case fails to suffice for Smith to lack justified-true-belief-without-knowledge that the Gettier Misjudgment is incorrect. By taking these judgments to concern mere compossibility instead, Malmgren lacked the resources necessary to account for these facts.

In what sense, though, can the Gettier Case be said to suffice for the presence of justified-true-belief-without-knowledge, given that it permits metaphysically possible instances in which this is absent? Well, it was Williamson's insight to see that metaphysical sufficiency is not the only option here, and that the Gettier Judgment might instead concern what the Gettier Case *counterfactually* suffices for. But what, then, are we to make of the criticism Malmgren presses against this view?

---

[14]This is a means of displaying two sentences: one in which the word 'justified' occurs and one in which the word 'unjustified' occurs.

The criticism, recall, was that the correctness of the Gettier Judgment should not depend on whether there happen to be actual (or modally proximal) instances of the Gettier Case in which justified-true-belief-without-knowledge is absent. The same objection is pressed, in a slightly different way, by Ichikawa (2009) and Ichikawa and Jarvis (2009). Williamson has responded to this worry both by trying to explain away the inclination to think that the Gettier Judgment does not so depend (2007b, pp. 200–204) and by arguing that the critics' proposed alternatives fare no better, as they leave the Gettier Judgment exposed to similar epistemic risks (2009, pp. 466–468).[15] Details aside, I agree with the critics that these responses are less than satisfactory. But there is another, better response that can be given on behalf of a counterfactual proposal.

4 Normal counterfactual sufficiency

We have a shared understanding of what it is for an instance of a given case to be abnormal, in this or that respect, *as* an instance of that case. This generalises into a grasp of what it is for an instance of a given case to be an abnormal one *simpliciter*. The existence of this grasp is testified to by Williamson's use of the the word 'abnormal' (2007b, pp. 200, 205) and Malmgren's use of the word 'deviant' (pp. 276–277 and *passim*) to generalise intelligibly about a range of instances of cases identified primarily as such—that is, as abnormal or deviant. But obviously we do not merely know what it *means* for an instance to be an abnormal one; we have a sense of whether or not particular instances of cases *are* abnormal, *qua* instances of those cases. Call this our *sense of normalcy*.[16] This sense puts us in a position to make judgments about what would normally hold in a given case, and a moment's reflection should reveal that this is something that we in fact do with some regularity.[17] The suggestion, then, is going to be that our judgments about thought experiments are typically judgments of just this sort—that is, they are typically judgments about what hypothetical scenarios *normally counterfactually suffice* for. Before making this suggestion a little more precise, though, there are three closely related points that need to be made.

The first is that what would be normal for a given scenario is not simply a mat-

---

[15]It is puzzling that he has responded in this fashion, given his acknowledgment elsewhere that the Gettier Judgment might not be falsified by the existence of such instances (Williamson 2007b, pp. 306–307).

[16]It is easy to see how this sense could be grounded in, or part of, our capacity for counterfactual judgment, if that capacity fits the account of it sketched by Williamson (2007b, pp. 134–178).

[17]It is clear that we make judgments not only about what a case would normally ensure, but also about what it would normally permit. Speaking simply of what is or would be normal for a case seems to be ambiguous between the two. (In the ensuring sense, it would be (say) normal for a t-shirt to have two armholes, and so abnormal for a t-shirt to have any number other than two. In the permitting sense, it would be normal for a t-shirt to be blue—but also normal for a t-shirt to be red, or to be white, or … .) In what follows, when speaking of what is or would be normal, I have only the former sense in mind.

ter of what happens to hold in the actual instances of that scenario.[18] This can be brought out in two ways. First, by pointing to the fact that our sense of what is or would be normal for a given scenario need not be undermined by the discovery that it has no actual instances. Second, by pointing to the conceivability of there being more instances of a given situation that are abnormal in some respect than ones that are normal. (Or to make the same basic point in another way: we can imagine it being the case that genuine coincidences have happened more often than not in the actual instances of some scenario.) This idea needn't require any notion of what is *natural* for a situation, in the sense of situations themselves having a nature. For it may suffice to underwrite a claim of prevalent abnormalcy that, in local modal space, instances with the (supposedly) abnormal feature are sufficiently outnumbered by instances with a (supposedly) normal feature that precludes it.[19]

The second is that we can be *wrong* about would be normal. We can fail to realise something is normal/abnormal, and we can think that something is normal/abnormal when it is not. That is to say, what would be normal for a given scenario is not simply a function of what we take to be normal for that scenario. Rather, what would be normal for a scenario is a matter of how things generally go in worlds that are like ours except insofar as they contain, or easily could have contained, instances of the relevant scenario. Now, we typically think of how things generally go in terms of the laws, norms and tendencies that we take to be in force. Keeping with this way of thinking, then, we can say that what is in fact normal for a scenario will be any feature whose absence from an instance of that scenario (in one of the worlds just mentioned) involves an exception to some relevant law, norm or tendency (that is in force across those worlds.)

The third is that, when engaging in thought experiments, our sense of whether or not an instance of a case is normal is not driven solely or even primarily by our sense of whether or not the properties that are of philosophical interest to us are present or absent in that instance. To take the Gettier Case as an example, our sense that certain instances of that case count as abnormal is not driven by our sense that those instances are ones in which justified-true-belief-without-knowledge is absent. If we consider an instance of the case in which the person in Smith's position is prone to hallucinations, for example, this strikes us as abnormal precisely because he or she is prone to hallucinations, and not primarily because he or she thereby lacks justified-true-belief-without-knowledge.

The relevance of these points will become clear as we proceed. For now, let us return to the suggestion mooted above. The thought was that we can characterise the

---

[18]Cf. Smith 2010, pp. 15–16.

[19]The application of the notion of outnumbering here may be suspect, as there may be an uncountably infinite number of instances (or worlds containing instances) that are close enough to be relevant to normalcy. I will simply assume that there is a way to make good the thought underlying the claim. (One suggestion, though: local modal space will presumably be the way required for there to be prevalent abnormalcy for some scenario when there are genuine, objective (though perhaps contingent and non-natural) tendencies which, for varying and disparate reasons, have failed to manifest themselves in a majority of the instances of that scenario in the actual world.)

content of judgments about thought experiments in terms of normal counterfactual sufficiency. In order to make the idea more precise, let us consider how to apply it to the Gettier Judgment.

There are in fact two ways in which we might seek to implement this idea. One is simply to provide a new counterfactual rendering of the content of the Gettier Judgment that makes explicit the restriction to normal instances. Another is to argue that ($\Box\!\!\to$) is acceptable as it stands, on the basis that a correct semantics for ($\Box\!\!\to$) would itself involve a restriction to normal instances. As nothing of immediate epistemological import rests on the choice, I will briefly comment on the second strategy before turning to the first.

According to the standard semantics for counterfactuals, ($\Box\!\!\to$) is true just in case all of the closest possible worlds in which someone is related to a proposition p as in the Gettier Case are worlds in which the relevant individual has a justified true belief that p but does not know that p.[20] If the actual (or a suitably nearby) world contains an instance of the case in which justified-true-belief-without-knowledge is absent, then according to this semantics the counterfactual is false. But this should give us pause. ($\Box\!\!\to$) is what we might call a perfectly general counterfactual. Its antecedent picks out a situation which can have multiple instances in a world, each involving distinct individuals. Indeed, it makes no reference to *any* particular objects, events or times. In and of itself, therefore, ($\Box\!\!\to$) does not direct us to any manageable range of particulars such that its truth turns solely on whether the modally local instances of the antecedent situation involving those particulars also satisfy the consequent. And so accordingly, in assessing ($\Box\!\!\to$), we make no attempt at investigating the contingent circumstances of any given particulars in order to work out whether they happen meet this condition. Now, it might well be that there is, nevertheless, some limited range of possible instances of the Gettier Case such that the worlds containing them count as the closest ones containing instances of the case. And if so, it is entirely possible that justified-true-belief-without-knowledge is absent in one of these instances. But again, in assessing ($\Box\!\!\to$), we simply do not concern ourselves with working out whether or not this is the case. That is to say, we do not concern ourselves with the vagaries of happenstance, or what could have easily resulted from the vagaries of happenstance. We rely, instead, on our general sense of how things go—on our sense of normalcy—trusting it to put us in a position to assess ($\Box\!\!\to$). This gives us a reason, albeit an inconclusive reason, to think that ($\Box\!\!\to$)

---

[20]See Lewis 1973 and Stalnaker 1968. In fact, it is unclear how to apply to standard semantics to counterfactuals such as ($\Box\!\!\to$), due (in part) to the presence of so-called 'donkey anaphora'. (Cf. Geach 1962; King and Lewis 2016.) One aspect of the problem generated for the standard semantics by such counterfactuals is that their consequents are not propositions, and so it is unclear what it takes for an antecedent world to be a consequent world. Consider the gloss just given in the text. A nearby world in which someone is related to a proposition as in the Gettier Case may well be a world in which *two* people are so related to a proposition. But then who is 'the relevant individual' on whose epistemic state the truth of the counterfactual is supposed to depend? As it stands, no coherent condition is laid down for the world to meet. What is needed is a semantics that appeals to something more fine-grained, with a structure the account can then exploit. (For discussion of how to accommodate the phenomenon, see van Rooij 2006, Williamson 2007b, pp. 305–308, and Wang 2009. Cf. Fine 2012.)

is something the truth of which does not depend on the vagaries of happenstance, and which instead depends on what would be normal for the case specified in its antecedent. However, it is beyond the scope of this paper to provide an alternative semantics for such counterfactuals.[21] Moreover, a reasonable starting point for developing such a semantics might be to seek a claim in English that is parallel to ($\square\rightarrow$) but which makes explicit the restriction to normal instances. For these reasons, I shall now turn to the first way of implementing the idea mentioned above.

What, then, is a plausible content proposal for the Gettier Judgment that would count it as concerning normal counterfactual sufficiency? The suggestion is in fact very straightforward, given what has already been said. It is this:

($\blacksquare\rightarrow$) If someone were to stand to a proposition p as in the Gettier Case, then, normally, he would have a justified true belief that p but not know that p.

There are undoubtedly significant questions concerning the precise semantics of ($\blacksquare\rightarrow$) and how it should be formalised. But I will abstain from addressing such questions here. The sentence has a manifest sense. This puts us in a position to, and makes it reasonable for us to, consider its adequacy as a characterisation of the content (or modal significance) of the Gettier Judgment. For present purposes, then, our existing understanding will suffice.

First, it is clear that ($\blacksquare\rightarrow$) is both true and inconsistent with the claim that, necessarily, someone has a justified true belief that p if and only if he or she knows that p. (Granted, of course, the background assumption that an instance of the Gettier Case is possible.) It can therefore account for the role that the Gettier Judgment plays in our theorising just as well as ($\square$), ($\square\rightarrow$) and ($\lozenge$) can.[22]

Second, the present proposal avoids the problems that Malmgren's view faced. It is false that if someone were related to a proposition p as in the Gettier Case, then,

---

[21]There are a number of accounts in the literature which seek to give a semantics for 'would' counterfactuals according to which they allow for modally local exceptions. To the best of my knowledge, such proposals typically take the antecedents and consequents of counterfactuals to be propositions, and so are of no immediate help in the present context given the point made in the preceding footnote. Moreover, they are typically intended to apply to *all* 'would' counterfactuals, not simply what I have called general ones. But it is worth briefly abstracting away from these points to ask whether any might serve as a useful model. Schulz (2014) provides a semantics according to which the truth of a 'would'-counterfactual depends on whether the consequent holds at a single, arbitrarily selected (via an ε operator), close antecedent world. However, if I understand the proposal correctly, this alone would leave the Gettier Judgment at risk of being incorrect for the wrong reasons, as bad luck could see Schulz's ε operator arbitrarily selecting a world containing an abnormal instance of the Gettier Case, rendering ($\square\rightarrow$) false. Better models for present purposes can perhaps be found in (Leitgeb 2012a,b), which provides a probabilistic semantics for counterfactuals, and (Smith 2007), which provides a semantics for *ceteris paribus* conditionals that explicitly invokes a notion of normalcy. Detailed consideration of these proposals, though, is beyond the scope of this paper.

[22]Regarding its truth, it is perhaps worth reiterating that normalcy is case-relative. That is, there is a distinction to be drawn between what would be normal for an instance of believing *simpliciter*, and what would be normal for an instance of believing-as-in-the-Gettier-Case. So however abnormal it might be for instances of believing (or instances of justifiably and truly believing, or instances of not knowing) to be instances of justified-true-belief-without-knowledge, this in no way precludes it from being normal for instances of the Gettier Case to involve justified-true-belief-without-knowledge.

normally, he would neither have a justified true belief that *p* nor know that *p*. So if this is the content of the Gettier Misjudgment, then that judgment is incorrect. Moreover, whatever turns out to be normal for the Gettier Case, it is clear that the claim just mentioned is inconsistent with (■→). So the Gettier Judgment and Misjudgment will be in tension in the manner required.

Third, and most significantly, it is not subject to the criticism Malmgren, Ichikawa and Jarvis press against Williamson's view. According to this criticism, if (□→) is the content of the Gettier Judgment, then we can be right about how things would normally be in an instance of the Gettier Case, make the judgment on that basis, and nevertheless have our judgment turn out incorrect due to the existence of an actual abnormal instance. If (■→) is the content of the Gettier Judgment, however, then this is not possible. As was noted above, the distribution of features over actual instances of a case does not settle what would be normal for that case. What normalcy turns on is the distribution of features over instances in local modal space. So a single instance of the Gettier Case in which justified-true-belief-without-knowledge is absent, whether actual or modally proximal, is not enough to falsify (■→). As long as instances of the Gettier Case containing justified-true-belief-without-knowledge are the norm, then the Gettier Judgment is correct.

It is true, of course, that this leaves the Gettier Judgment exposed to *some* of the epistemic risks associated with contingency. After all, we can conceive of worlds in which it fails to be normal for instances of the Gettier Case to include justified-true-belief-without-knowledge. Consider, for example, a world in which the norms for car ownership are quite different from our own. At this world, the vast majority of car drivers rent their cars, and this is public knowledge. In such a world, seeing someone in your office drive a Ford to work and remembering that they always drove a Ford to work provides no real justification for believing that they own a Ford. So at this world, (■→) is false. There are, however, two key points to make about the contingency of (■→) and its associated epistemic risks.

The first is that, whereas it seems to us perfectly plausible that there might actually be, in our world, an instance of the Gettier Case in which justified-true-belief-without-knowledge is absent, it surely does not seem plausible to us that we inhabit a world in which (to keep with the example) the norms for car ownership are in fact vastly different from what we take them to be.[23] So (■→) is not contingent in a way that should cause us to doubt that we know it.[24] The second is that, if we imagine a subject in the renting world being presented with the Gettier Case, it is far from obvious that he or she would be correct to make the Gettier Judgment. For surely it would be natural for such a subject to make (what we call) the Gettier Misjudgment. And it seems, in these alternate circumstances, that this would be exactly the

---

[23]It is perhaps worth noting here that claims of normal counterfactual sufficiency may exhibit a significant degree of context or occasion sensitivity, not only in ways standardly associated with counterfactuals, but also, presumably, in ways associated with the adverb 'normally'.

[24]The present proposal therefore also avoids a criticism pressed against the counterfactual proposal by Ichikawa (2009, pp. 439–442), according to which (□→) is not plausibly known, due to the epistemic possibility of there being a modally local violation.

right judgment to make. In considering of this case, therefore, we find extra support for the claim that judgments about thought experiments concern normal counterfactual sufficiency: as the facts about what would be normal for a given case vary, so too the correctness of judgments about thought experiments based on that case seem to vary.[25]

The forgoing makes the case for taking (■→) to be the content of the Gettier Judgment, and so for taking typical judgments about thought experiments to be judgments of normal counterfactual sufficiency. But a question remains. Why did neither Williamson nor his critics seriously consider this option?

On the part of the critics, the answer to this question is clear. Malmgren, Ichikawa and Jarvis were all concerned to defend the apriority of the Gettier Judgment. Accordingly, their fundamental problem with (□→) as a content proposal was that, in opening up the judgment to epistemic risks associated with contingency, it threatened the judgment's supposed apriority. On that score, (■→) is clearly no bet-

---

[25]One might object that what explains this difference in correctness is not the difference in norms for car ownership, but rather the difference in the evidence that the subjects in question have (or are naturally assumed to have) concerning those norms. As a reason to accept this alternative explanation, one might suggest that it is also plausible that, were a subject in the renting world to have misleading evidence suggesting that the norms for car ownership there are more or less as they in fact are in our world, they would be correct to make the judgment that they would express via (S). Now, note that the correctness of a modal judgment about a single, given hypothetical case is not sensitive to the evidence of the subject making it. (If some particular feature must (or would, or normally would, or could) be present in instances of a given scenario, this is so *simpliciter*, not relative to a body of evidence.) So if the foregoing accurately represents the pattern of correctness for the renting-world judgments in question, then they must strictly speaking concern different hypothetical cases. And the explanation for this difference, I take it, will be that the subjects in question are interpreting the case description in such a way that they take Smith to have the same evidence concerning the norms of car ownership that they do. The claim, then, is that in making a judgment in response to the Gettier Case description, one's judgment really concerns whichever enriched case results from one's evidence-dependent interpretation.

There are four brief points I want to make in response to this. First, the objection implies that if someone in the actual world responds to the Gettier Case description by making a judgment that they would express by (S*), then, as long as they are appropriately deluded about the norms for car ownership due to misleading evidence, we are obliged to regard their judgment as correct. It also implies that, having made the Gettier Judgment, each of us is obliged to regard certain further judgments expressible by sentences of the form 'Smith has evidence that *p*' as trivially true, despite *p* concerning factors that are not explicitly mentioned in the case description. But it is not at all clear that these are obligations we are (or should be) willing to recognise. Second, and relatedly, the kind of sensitivity to evidence envisaged would suggest that individuals with different relevant evidence are likely to be making judgments about distinct cases despite being involved in a shared discourse. But there is no good reason to think that differences in evidence between subjects will frustrate public discussion of and disagreement about thought experiments in this way. (Cf. Williamson 2009, p. 468.) These points suggest that the alternative explanation being proposed is questionable in its own right. But third, even if it is correct, it would not directly call the view I am proposing into question. It would, instead, simply undercut some extra support claimed for the view. And fourth, merely taking Smith to possess certain pieces of evidence that we ourselves possess would do nothing to avoid any of the problems facing the other views under consideration, as outlined in §1–3 and expanded below in §5.1. To the extent that the arguments presented against those views are correct, then, they would carry over to any non-counterfactual view that adopts this alternative explanation. (I thank an anonymous referee for raising the point considered in this note.)

ter, and so the critics had no reason to think of it. But their argument against Williamson's proposal was not that it left the Gettier Judgment exposed to epistemic risks associated with contingency *per se*; it was that it left the Gettier Judgment exposed to *implausible* epistemic risks. And on that score, (■→) cannot be faulted.

On Williamson's part, there are two reasonable suggestions. The first possibility is that he saw the problems his critics presented as borne of a desire to preserve the apriority of the Gettier Judgment—a goal to which he is not sympathetic—and so failed to take them seriously. The second is that, because his aim was to formalise the Gettier Judgment so as to make explicit (and justify) the precise role it plays in rejecting the justified-true-belief theory of knowledge, the extra difficulties one would encounter in trying to formalise (■→) made it less appealing. However, neither of these points justifies misrepresenting the content of the judgment, and it remains implausible that the judgment's correctness depends on whether or not there just so happens to be a nearby, abnormal instance of the case in which justified-true-belief-without-knowledge is absent.[26]

5   TWO OBJECTIONS

I have argued that an adequate account of the Gettier Judgment must recognise that it concerns what the Gettier Case in some sense suffices for, and that taking (■→) to be its content meets this condition while avoiding problems faced by other accounts. Before concluding that we should accept this proposal, however, two potential objections must be considered. The first, addressed in §5.1, concerns whether (□) might also be amended in such a way as to avoid the problem identified in §1. The second, addressed in §5.2, concerns how well the account fares when generalised to judgments about other hypothetical scenarios, and in particular to judgments about cases that are modally remote or nomologically impossible.

5.1   AMENDING (□)

Call any view according to which typical judgments about thought experiments have strict universal implications as contents—that is, any view according to which such judgments concern what *must* hold in instances of the relevant cases—a necessity

---

[26]In fact, Williamson does, in a two-sentence footnote, briefly moot the possibility of taking the Gettier Judgment to concern normal instances of the Gettier Case, but dismisses it on the basis that 'the relevant notion of normality is an epistemological one that violates the supposed neutrality of the initial description of the case' (2007b, p. 204, n. 22). However, the only prior mention of neutrality in his discussion appears much earlier, in his characterisation of a case description as being presented 'in neutral terms, without prejudice to the target analysis' (p. 184)—that is, in this context, without prejudice as to whether or not Smith has a justified true belief, but doesn't know, that someone in his office owns a Ford. And on the present understanding of the Gettier Judgment, neither the case description nor its interpretation prejudges the target analysis in the relevant sense. That is, the case about which the judgment is being made perfectly well permits both instances in which the person in Smith's position has justified-true-belief-without-knowledge and instances in which he does not. So there is no sense in which the neutrality of the initial description of the case is violated.

view. We can think of necessity views in general as committed to the claim that the Gettier Judgment has the content:

($\Box^*$) Necessarily, anyone who stands to a proposition p as in Δ has a justified true belief that p but does not know that p,

where 'Δ' serves to pick out a particular hypothetical case: the case about which the Gettier Judgment is being made.

Taking Δ to be the Gettier Case, as seems most natural, ($\Box^*$) becomes ($\Box$). But this we have already rejected. The problem, recall, was that the Gettier Case has possible instances which lack justified-true-belief-without-knowledge. This entails that ($\Box$) is false, and so not plausibly the content of the Gettier Judgment, given its correctness.

A obvious response for one who wants to defend a necessity view, then, is to suggest that the Gettier Judgment does not really concern what must hold in instances of the case explicitly described. Rather, the response goes, it concerns what must hold in a more determinate version of that case, one richer in detail, and richer in such a way as to necessitate the presence of justified-true-belief-without-knowledge in its instances. The central question for any such view is then: what case *does* the judgment concern? What is Δ? If a plausible answer can be found that renders ($\Box^*$) non-vacuously true, then we may have no reason to prefer a counterfactual view in the first place.

Anticipating a manoeuvre of this sort, Williamson claims that any attempt to emend the description of the Gettier Case so as to rule out potential confounds is bound to fail, as '[a]ny humanly compiled list of such interfering factors is likely to be incomplete' (2007b, p. 185). Recently, however, this has been disputed by Grundmann and Horvath (2014a), who provide the following description of a more determinate version of the Gettier Case:

> Smith justifiedly [*sic*] believes that Jones owns a Ford, on the basis of seeing Jones drive a Ford to work and remembering that Jones always drove a Ford in the past. From this belief alone (plus the justified background belief that Jones is in his office), Smith logically infers, at time t, to the justified belief that someone in his office owns a Ford, which provides his only justification for that belief at t. In fact, someone in Smith's office does own a Ford, so that Smith's latter belief is true—but it is not Jones, it is Brown, and so Smith's initial belief was false. (Jones sold his car and now drives a rented Ford.) Also, if Smith knows at t that someone in his office owns a Ford, then he knows this at t only in virtue of the facts described.[27]
>
> (Grundmann and Horvath 2014a, p. 530)

Does this case description provide for a plausible defence of a necessity view?

For it to do so, at least three things need to be the case. First, ($\Box^*$) must be non-vacuously true when Δ is taken to be the case just described. Second, it must

---

[27]This incorporates the authors' proposed fix for an erratum noted in (Grundmann and Horvath 2014b, p. 536) and removes the underlining present in the original.

be plausible that Δ, the case about which the Gettier Judgment is really made, *is* the case just described. And third, the proposal must be plausibly generalisable to other judgments about thought experiments.

Let us grant, for the sake of argument, that no possible instance of this case could lack justified-true-belief-without-knowledge, and that this is so for appropriate, non-vacuous reasons. Then, if we take Δ to be the case described above, (□*) will be true, and the first condition will have been satisfied. Even then, however, the other conditions are not. It is not plausible that Δ is the case above; and even if we take it to be, the resulting proposal does not plausibly generalise to cover other judgments about thought experiments. I shall bring out each point in turn.

There are multiple reasons to be sceptical of the claim that the Gettier Judgment concerns the case above. To begin, note that we need some explanation of how those who make the judgment arrive at this more determinate case.[28] The authors suggest that, when presented with the (original) Gettier Case description, we 'implicitly interpret' it in this more determinate way. It is unclear what exactly this implicit interpretation is supposed to involve. But it is, at any rate, such that the enriched description above 'should be readily available and transparent to every professional epistemologist' (pp. 531–532).[29] This is, however, simply not the case. It takes time and intellectual effort to work out how to emend the initial description in this way. And while somewhat *ad hominem*, it is surely philosophically relevant to observe here that, given the significant erratum noted in (Grundmann and Horvath 2014b), this way of enriching the description was apparently neither readily available nor transparent even to the authors themselves.

Next, note that the truth of (□*), when Δ is taken to be the case above, is supposed to be ensured by the authors' stipulations concerning the epistemic structure of the case. This gives rise to two issues.

The first is that the Gettier Judgment is naturally taken to be epistemologically substantive in two ways, both in its commitment to the absence of knowledge and in its commitment to the presence of justification. But this proposal renders the verdict that Smith's belief is justified entirely trivial.[30] Insofar as we do not take our judgments concerning justification when presented with the original Gettier Case description to be trivial, then, we have reason to think that we are not interpreting

---

[28] The demand here is for an explanation of how it is that one comes to have in mind the case on which one reflects, and about which one makes a judgment, given that its details go beyond what is provided by the initial description. This is distinct from a demand for an explanation of how it is that one arrives (perhaps justifiably or knowledgably) at a judgment about it. Given any content proposal, there must ultimately be some account of the latter sort. But only given the present kind of proposal is an account of the former sort required.

[29] One suspects that many professional epistemologists would be surprised to learn what this seems to imply, namely that Williamson and Malmgren are not to be counted amongst their number.

[30] The authors do in fact note this point (p. 529, n. 8). As evidence that Gettier himself intended his original cases not to leave the presence of justification an 'open question', they cite the fact that he made substantive claims about the nature of justification that were intended to 'incline us to attribute justified beliefs to Smith.' But surely this is just evidence that he took the judgment that Smith's belief is justified to be *correct*? Besides, if it was intended to be stipulated, why did he not just stipulate it?

it in the way the authors claim.

The second concerns views of justification according to which the case above would be non-instantiable. The point here is not that such views might be correct, and that (□*) might be vacuously true—we have granted for the sake of argument that it is not. Rather, the point is that, on the present account, we fail to make sense of those who hold these non-standard views of justification. Now, the authors themselves make mention of those who deny the possibility of justified false belief (p. 529, n. 9; cf. Sutton 2007 and Littlejohn 2012). But they do not recognise that such philosophers would surely dispute the Gettier Judgment while urging on us the Gettier Misjudgment. The present proposal implies, however, that either they are misinterpreting the case due to a lack of expertise, or that they are, according to their own theories, making a simple mistake in failing to regard both judgments as merely vacuously true. The problem only becomes more serious once it is noticed that certain other views about justification also render the case impossible. For example, a coherentist about justification (such as BonJour 1985) might well deny that a person's 'only justification' for believing something could be a single other belief of theirs. Yet such a philosopher might nevertheless demur from the Gettier Misjudgment, make the Gettier Judgment, and rule out the thesis that knowledge is justified true belief on its basis. And once again, we would be forced to attribute to them either misinterpretation or mistake, when neither seems appropriate.

Even if Δ were the case above, though, our other problem would remain. This is that there is simply no good way to generalise the proposal under consideration to other judgments about thought experiments. Now, the authors themselves voice some scepticism about whether their proposal generalises (p. 531). But we concluded in §2 that one should not take judgments about thought experiments to have contents of inexplicably different strengths. And this conclusion, I want to suggest, should be taken to apply not only to distinct judgments about a single thought experiment, but also to judgments about different thought experiments. That is to say, we should only assign (□*) to the Gettier Judgment if a necessity view is defensible in general. So if the present proposal is to provide the blueprint for a defence of a necessity view, we must suppose that in every case in which we make correct, non-trivial judgments about thought experiments, the relevant case descriptions are being implicitly interpreted in ways that necessarily but non-trivially ensure the presence of the features that we attribute to them in our judgments.[31] Yet it already seems a leap of faith to suppose that more determinate case descriptions of this kind will always be generable. To suppose that they are always in some sense present to

---

[31]It would be a mistake to object to this on the basis that no general recipe has been provided for enriching descriptions so as to determine exactly which cases (expert) judgments concern. The generalised claim is simply that, when making judgments about thought experiments that are in good order, we skilfully interpret the cases in more determinate ways so as to make the relevant strict implications non-vacuously true. There is no reason to think that the plausibility of such a view depends on the possibility of spelling out in advance an exact procedure which this skilful interpretation employs. (Though we might well wonder whether we should expect all of the expert judgments prompted by a given description to concern the same case. Cf. the second point raised in n. 25 above.)

our minds when making the judgments we do is simply incredible. The only option that remains for one who wants to generalise this proposal is to conclude that, when engaging in thought experiments, we are guilty of making many rash and incorrect judgments of manifestly inappropriate strength. But this is a broadly sceptical result, one which we should not accept unless it is genuinely forced on us. And the availability of the proposal outlined in §4 shows that it is not.

Note, then, that (■→) does not suffer from any of these problems. It allows us to take as the case in question the one corresponding to a literal and unembellished reading of the case description provided back in §1. So the relevant case will be transparent and readily available not only to all experts, but to anyone who understands the language. Judgments both about justification and about knowledge will be substantive. And philosophers with the various views of justification canvassed above can arrive at the same judgment about what would normally hold in instances of the case in question, or at different judgments about what would normally hold in instances of one and the same case, without being guilty of any implausible misinterpretations or mistakes.

Now, the proposal just considered exemplifies only one strategy for defending a necessity view.[32] Before moving on, I want to briefly consider another. The strategy I have in mind is one which invokes *relevance*.[33] Showing what is wrong with this way of defending a necessity view will provide us with a diagnosis, made available by (■→), of why such views continue to strike some as appealing, despite their inadequacy.

The motivating thought here is simply that, when assessing whether or not a hypothetical case has certain features, we take the case description to include everything relevant to whether or not those features are present. But there are various ways to apply this rough thought in the present context. One way is to stay at the same level of generality, and claim that we take the Gettier Case description to include everything relevant to whether or not Smith has justified true belief without knowledge. Another is to make the claim a little weaker, but more specific to the case at hand, claiming instead that we take the Gettier Case description to include all of the evidence Smith has that is relevant to whether or not someone in his office owns a Ford. We can think of both of these claims as entailing that Δ can be described by the Gettier Case description concatenated with an extra sentence making explicit whichever relevance restriction we want to go with.[34] Either way, the hope is that this will ensure that no possible instance of the resulting case can exhibit features rendering Smith's belief either unjustified or knowledge.

---

[32]Ichikawa and Jarvis (2009) pursue another strategy that appeals to idea that we read case descriptions in accordance with the norms for reading fiction. I will not discuss their view here; for criticism, see Malmgren 2011, pp. 303–306 and Williamson 2009, pp. 465–469.

[33]'This possibility was raised by an anonymous referee. I have also encountered it numerous times in conversation.

[34]In the former case, this might be the sentence: 'There is nothing else relevant to whether or not Smith has a justified true belief without knowledge.' In the latter case, it might be: 'Smith has no other relevant evidence concerning whether or not someone in his office owns a Ford.'

The problem with the former, more general way of implementing the idea is simply that *any* instance of the Gettier Case will have features that are not explicitly mentioned in its description, but which are relevant to whether or not the person in Smith's position has justified-true-belief-without-knowledge. To see this, consider the following feature that an instance of the case might have: the person in Smith's position having been told that the person in Jones's is renting. Now, the *presence* of this feature is clearly relevant to whether or not the person in Smith's position has justified-true-belief-without-knowledge. But it is not explicitly mentioned in the original description of the Gettier Case. Therefore, an instance of the original case with this feature is not an instance of the case that includes the relevance restriction. So far, one may think, so good. But then, by the same token, the *absence* of this feature—that is, the person in Smith's position *not* having been told that the person in Jones's position is renting—must also be relevant to whether or not the person in Smith's position has justified-true-belief-without-knowledge. Yet this, too, is not explicitly stated in the original description. Therefore, an instance of the original Gettier Case without this feature is *also* not an instance of the case that includes the relevance restriction. It follows that any instance of the enriched case must neither have nor lack this feature. In other words, it follows that there are no possible instances of the enriched case. And so a version of ($\square^*$) that takes this enriched case to be Δ will only be vacuously true.[35]

The same basic objection can also be pressed one level up, as it were, in the following way. I claimed above that the appeal to relevance is driven by the thought that the case description can be read as including everything relevant to whether or not Smith has justified-true-belief-without-knowledge. But one thing that is relevant to whether or not Smith has justified-true-belief-without-knowledge is whether or not he has been told that Jones is renting. And the case description simply does not state whether or not Smith has been told that Jones is renting. It therefore straightforwardly fails to include everything relevant to whether or not Smith has justified-true-belief-without-knowledge, and cannot be read as doing so, *contra* the thought driving the appeal.

The defects with the latter way of implementing the appeal to relevance are a little more subtle, but no less problematic. This relevance restriction, while weaker than the last, is still rather strong. It requires that Smith must have no evidence whatsoever that bears on whether or not someone in his office owns a Ford, other than seeing Jones drive a Ford to work and remembering that he always drove a Ford to work. But this implies that Smith has no evidence whatsoever concerning, among a great many other things, the norms of car ownership. For any such evidence would

---

[35]The point could have been made equally well with any epistemologically relevant feature not explicitly mentioned in the Gettier Case description, including: the person in Smith's position being a competent reasoner, having good vision, being able to easily recognise Fords, knowing the norms for car ownership, not being prone to hallucinations, having no other evidence that renders his or her belief knowledge, and so on. It would therefore be a mistake to attempt to resist this argument by pointing to the fact that the person in Smith's position not having been told that Jones is renting is in some intuitive sense a 'negative' rather than a 'positive' feature.

surely be relevant to whether or not someone in his office owns a Ford.[36] And now the problem is obvious: without any such evidence, it is hard to see how Smith's belief that Jones own a Ford could be justified. Unless a person in Smith's position were in some sense aware that driving a car to work every day typically indicates ownership, what could license his or her inference? Recall that this restriction was supposed to provide us with a candidate for Δ that would render (□*) true. It now looks like this restriction will generate a case that not only fails to metaphysically necessitate the presence of justified-true-belief-without-knowledge in its instances, but which instead necessitates its absence. It therefore cannot provide for a defence of a necessity view.

We are left with a puzzle. Why is it so tempting to appeal to relevance here, given the seriousness of the problems above? And what I want to suggest is that taking (■→) to be the content of the Gettier Judgment actually allows us to provide an answer—and as far as I can see the only plausible answer—to this question.

The thought is that those who would make these appeals to relevance must *already* have in mind a certain verdict about the case in question. That is to say, they must have already judged that Smith has a justified true belief without knowledge. For we can then understand their claim that the case description includes everything relevant as giving inchoate expression to the idea that it contains everything relevant to this judgment—that is, that it contains all that is required of a case in order for a judgment expressible in these words to be correct. And the temptation to invoke relevance when presented with certain potential further features can be understood as deriving from the fact their presence *would* have been relevant, in precisely the following sense: that, had one instead considered a modified version of the case with these features included, one would no longer have taken a judgment expressible in these words to be correct. But this implies that the Gettier Judgment cannot depend on 'interpreting' the case description in a way that is guided by any prior identification of various features as ones that would have been relevant to whether Smith has justified-true-belief-without-knowledge. Because taking certain features to be ones that would have been relevant, in this sense, itself depends on the judgment.

What all of this requires, of course, is that we have some account of the nature of the judgment in the first place such that it can play the required role in controlling our sense of whether or not further features would have been relevant. But this is just what (■→) provides.[37]

---

[36]In fact, once one begins to reflect on just how many things could have a bearing (however indirect) on whether or not someone in one's office owns a Ford, it starts to seem plausible that even this weaker relevance restriction generates a case without possible instances. For it is doubtful that there could be a subject with only two pieces of evidence. And yet only two pieces of evidence are explicitly mentioned in the Gettier Case description.

[37]It is worth noting that Malmgren considers a content proposal that has some affinities with an appeal to relevance (pp. 285–290). The proposal she considers also takes the form of a strict implication, but it appeals to the 'intended Gettier case' in its antecedent in order to restrict the range of semantically relevant instances of the Gettier Case to ones in which justified-true-belief-without-knowledge is present. She rejects the proposal on the basis that it is unclear how the restriction can be effected other than by stipulating that 'the intended Gettier case' is to involve justified-true-belief-

5.2   Generalising (■→)

The final worry I want to address concerns how well (■→) generalises to judgments about other cases. In particular, I will address the question of whether thought experiments involving modally remote or nomologically impossible cases pose a problem for the view that typical judgments about thought experiments are judgments about what would normally hold in instances of the relevant cases.

To a first approximation, and one which will serve present purposes, the view arrived at by generalising (■→) can be represented as follows. If a thought experiment involves a hypothetical case, described by $\ulcorner \Phi\ \tau_1\ \dots\ \tau_n \urcorner$, and a judgment is made about it that would be expressed by $\ulcorner \Psi\ \tau_1\ \dots\ \tau_k \urcorner$ ($k \leq n$), then that judgment has the content expressed by: $\ulcorner$If some things $x_1, \dots, x_n$ were to be Φ, then, normally, $x_1, \dots, x_k$ would be $\Psi \urcorner$.[38]

There are two potential problems that this view might be thought to face when Φ picks out a modally distant or nomologically impossible case, one semantic, the other epistemic.[39] The semantic worry is that the relevant normally-would counterfactuals will be at best universally vacuously true, if not universally false. The epistemic worry is that, even if the relevant normally-would counterfactuals can be non-vacuously true, we are simply in no position to make knowledgable judgments about what would normally hold in instances of such cases. I will address these in turn.

The semantic worry, I take it, is driven by the suspicion that there can't be anything that would count, from the perspective of our world, as normal or abnormal for a case that is so thoroughly alien. But this is a mistake. Recall one of the glosses given on normalcy in §4: what is normal for a scenario is a matter of how things generally go in worlds that are like ours except insofar as they contain, or easily could have contained, instances of the relevant scenario. Now it is clear that a world which could (easily) contain an instance of a nomologically impossible scenario would have to be quite unlike our own. But this in no way precludes there being facts about how things generally go in such a world. It therefore does nothing to preclude there being facts about what such scenarios normally counterfactually suffice for. If there are laws, norms or tendencies in force at these worlds to which the absence of a certain feature from an instance of the scenario would constitute an exception, then the scenario normally counterfactually suffices for the presence of

---

without-knowledge, thereby rendering the judgment trivial. One could, perhaps, think of the appeal to relevance as another attempt, alongside those she considers and rejects, to spell out the nature of the 'intended Gettier case' in a way that does not render the judgment trivial. (It is also worth noting that the supposed difficulty arises only because Malmgren wants the restriction to render the judgment plausibly justifiable *a priori*. Once it is accepted that the instances one intends to make a judgment about are the normal ones, however, the non-trivial viability of a proposal that restricts the range of semantically relevant instances is vindicated, albeit not in the form she envisaged.)

[38]Φ will pick out a way for n things to be; Ψ a way for k things to be; $\tau_1 \dots \tau_n$ will be (typically, non-referring) singular terms. There are some superficial differences between this and the way in which (■→) is worded, but none of substance.

[39]In what follows, I will focus on nomologically impossible cases. The same points apply to modally remote cases.

that feature.

What remains unclear, however, is whether or not nomologically impossible cases can normally counterfactually suffice for anything which they do not also metaphysically necessitate. Given what has already been said, this will turn on whether or not, among the worlds that contain (or could easily have contained) instances of a given nomologically impossible case, some will count as being more like our world than others. Now, either this is so, or it is not. If it is so, then there might well be laws, norms and tendencies in force in the worlds most like ours, and not in force in the others, to which the absence of a certain feature from an instance of the case would constitute an exception. In that case, the nomologically impossible scenario will normally counterfactually suffice for the presence of this feature despite not metaphysically necessitating it. If it is not so, then every world containing an instance of the case in question will count as being equally like our world. And in that case, there will be no laws, norms and tendencies in force across this entire range of worlds other than those which are metaphysically necessary, and the features whose absence would (*per impossibile*) be an exception to these laws, norms and tendencies are just those whose presence is metaphysically necessitated by the scenario.

What it is important to see here, though, is that, either way, the relevant normally-would counterfactuals can be non-vacuously true. In the former case this is already clear: if a case normally counterfactually suffices for something without metaphysically necessitating it, then there is a non-vacuously true normally-would counterfactual. But in either case, as long as there are non-vacuously true strict implications, then there will be corresponding, non-vacuously true normally-would counterfactuals. At the limit, then, normal counterfactual sufficiency simply collapses into metaphysical necessitation.[40] But this cannot be grounds for objecting to the claim that typical judgments about thought experiments concern normal counterfactual sufficiency. At worst, it simply implies that typical judgments about thought experiments involving nomologically impossible cases might turn out to be correct only when corresponding strict implications are true.

Finally, then, let us briefly consider the epistemic worry. We can immediately set aside this worry concerning claims of normal counterfactual sufficiency for which a parallel claim of metaphysical necessitation would be true. However it is that we are supposed to come to know such strict implications concerning modally remote or nomologically impossible cases, they can present no special problem for the present view, given that any strict implication will imply a parallel normally-would counterfactual. Any acceptable account of how we are able to knowledgably make the former would *ipso facto* be an acceptable account of how we are able to knowledgably make the latter.[41] The worry, therefore, can only concern those cases in which Φ

---

[40]In this respect, normal counterfactual sufficiency seems to be no different to counterfactual sufficiency *simpliciter*. If there is a 'would' counterfactual for which no antecedent world counts as being closer than any other, it will be true just in case a parallel strict implication is true.

[41]Similarly, a general scepticism about the utility or reliability of judgments about nomologically impossible cases poses no special problem for the view. So let us also set aside such scepticism, how-

picks out a nomologically impossible case, and Ψ picks out a feature that the case normally counterfactually suffices for but which it does not metaphysically necessitate. The worry, then, is: how can we make such judgments knowledgably?

To this I think there are three appropriate responses. The first is that, on the face of it, this is a demand to account for items of knowledge which other views of judgments about thought experiments would wish to deny us. So if there are such items of knowledge, then this is immediately a reason to deny these other accounts and accept the present one, albeit with a view to providing an answer to the question. The second is that it is a presupposition of the worry that some of the worlds containing nomologically impossible cases are more like ours than others. If so, then this will be plausibly be, in part, a matter of some of the laws, norms and tendencies (however abstract and general) that are in force in the actual world still being in force in some, but not all, of these worlds. What will unite the worlds most like ours is simply their being alike in this respect. Therefore we can in principle use our sense of these abstract and general aspects of how things generally go to guide us in our judgments about the nomologically impossible scenarios. Of course, this would require that we also have some sense of what minimal differences would be required of the world in order to accommodate those scenarios. And this leads naturally to the third response, which is simply this: we really have no reason to suppose that we *can* make such judgments knowledgably. After all, it might well be that, for reasons mooted above, there is nothing that would normally hold in a nomologically impossible scenario other than what that scenario metaphysically necessitates. And even if this is not so, the truths in question may simply be unknown to us. As long as we keep in mind that there is no range of judgments which this view would render problematic but a version of the necessity view would account for, then, we are provided with no reason to doubt the claim that typical judgments about thought experiments are judgments of normal counterfactual sufficiency.

6   CONCLUDING REMARKS

I have argued that there are no good objections to taking (■→) to be the content of the Gettier Judgment coming either from a modification of the necessity view or from a consideration of thought experiments involving nomologically impossible cases. I want to conclude by raising two final questions.

First, does taking (■→) to be the content of the Gettier Judgment rule out the possibility of its being justified *a priori*? And second, to what extent can our answer be generalised to other typical judgments about thought experiments, granted that these are judgments of normal counterfactual sufficiency?

Recently, Williamson (2013) has expanded his attack on the *a priori/a posteriori* distinction, arguing against its theoretical significance on the basis of there being considerable epistemological and psychological parallels between the geneses of certain judgments that are supposed to be paradigms of *a priori* and *a posteriori*

---

ever well founded it might be.

justification. Many of the parallels he identifies seem to apply to our methods for forming judgments of normal counterfactual sufficiency. If his argument is correct, then the question of whether or not typical judgments about thought experiments can be correctly classified as *a priori* is simply no longer pressing.

But suppose we find reason to reject his argument, and that an epistemologically significant and viable notion of the *a priori* can be identified. How then should we answer the questions above?

On the face of it, the answer to the first may seem to be 'yes'. The truth of (■→) depends on what is normal for the Gettier Case. As we have seen, what is normal for this case turns on, amongst many other things, the norms for car ownership in the actual world. This is a deeply contingent matter, and so it is hard to see how the way in which our sense of normalcy underwrites the judgment could fail to undermine its classification as being justified *a priori*. And, again on the face of it, the answer to the second question may then seem to be that this verdict will extend to any other judgment the correctness of which depends on some contingent law, norm or tendency in the same kind of way.

In the end, however, the answers we are after will turn on delicate questions concerning what exactly the *a priori/a posteriori* distinction is supposed to come to, questions which cannot be fruitfully pursued when abstracting away from the details of any particular account of that distinction. All that can be said at this juncture is that friends of the *a priori* who want to stake out a role for it in the epistemology of judgments about thought experiments will have to show one of two things. Either it must be shown that *a priori* justification is at least sometimes available for judgments of normal counterfactual sufficiency whose antecedents do not metaphysically necessitate their consequents. Or it must be shown that the epistemic ground of such judgments involves a significant *a priori* component, despite the judgments themselves failing to count as such.

I do not take a stand on these deep and general issues here. Suffice it to say that, for most judgments about thought experiments, we are not isolated from the epistemic risks that arise due to the contingency of normalcy. The full bearing of this fact on the apriority of these judgments is a question that must be left to another time.[42]

References

BonJour, Laurence (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

Brown, Jessica (2011). Thought Experiments, Intuitions and Philosophical Evidence. *Dialectica* 65 (4): 493–516.

Cappelen, Herman (2013). *Philosophy Without Intuitions*. Oxford: Oxford University Press.

Deutsch, Max (2015). *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method*. Cambridge, MA: MIT Press.

Eagle, Antony (2007). Telling Tales. *Proceedings of the Aristotelian Society* 107: 125–147.

Fine, Kit (2012). Counterfactuals Without Possible Worlds. *Journal of Philosophy* 109 (3): 221–246.

Gardiner, Georgi (2015). Normalcy and the Contents of Philosophical Judgements. *Inquiry* 58 (7-8): 700–740.

Geach, P. T. (1962). *Reference and Generality*. Ithaca, NY: Cornell University Press.

Gettier, Edmund (1963). Is Justified True Belief Knowledge? *Analysis* 23 (6): 121–123.

Goldman, Alvin I. (2007). Philosophical Intuitions: Their Target, Their Source, and Their Epistemic Status. *Grazer Philosophische Studien* 74 (1): 1–26.

Grundmann, Thomas and Joachim Horvath (2014a). Thought Experiments and the Problem of Deviant Realizations. *Philosophical Studies* 170 (3): 525–533.

— (2014b). Erratum To: Thought Experiments and the Problem of Deviant Realizations. *Philosophical Studies* 170 (3): 535–536.

Ichikawa, Jonathan (2009). Knowing the Intuition and Knowing the Counterfactual. *Philosophical Studies* 145 (3): 435–443.

Ichikawa, Jonathan and Benjamin Jarvis (2009). Thought-Experiment Intuitions and Truth in Fiction. *Philosophical Studies* 142 (2): 221–246.

King, Jeffrey C. and Karen S. Lewis (2016). Anaphora. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2016. Metaphysics Research Lab, Stanford University.

Leitgeb, Hannes (2012a). A Probabilistic Semantics for Counterfactuals. Part A. *Review of Symbolic Logic* 5 (1): 26–84.

— (2012b). A Probabilistic Semantics for Counterfactuals. Part B. *Review of Symbolic Logic* 5 (1): 85–121.

Lewis, David (1973). *Counterfactuals*. Oxford: Blackwell: pp. 278–281.

Littlejohn, Clayton (2012). *Justification and the Truth-Connection*. Cambridge: Cambridge University Press.

Malmgren, Anna-Sara (2011). Rationalism and the Content of Intuitive Judgements. *Mind* 120 (478): 263–327.

— (2013). Review of *Philosophy Without Intuitions* by Herman Cappelen. *Notre Dame Philosophical Reviews*: URL: http://ndpr.nd.edu/news/39362-philosophy-without-intuitions/.

Nagel, Jennifer (2012). Intuitions and Experiments: A Defense of the Case Method in Epistemology. *Philosophy and Phenomenological Research* 85 (3): 495–527.

Pust, Joel (2000). *Intuitions as Evidence*. New York, NY: Routledge.

Schulz, Moritz (2014). Counterfactuals and Arbitrariness. *Mind* 123 (492): 1021–1055.

Smith, Martin (2007). Ceteris Paribus Conditionals and Comparative Normalcy. *Journal of Philosophical Logic* 36 (1): 97–121.

— (2010). What Else Justification Could Be. *Noûs* 44 (1): 10–31.

Stalnaker, Robert C. (1968). A Theory of Conditionals. In: *Studies in Logical Theory*. Ed. by Nicholas Rescher. Oxford: Blackwell: 98–112.

Sutton, Jonathan (2007). *Without Justification*. Cambridge, MA: MIT Press.

Van Rooij, R. (2006). Free Choice Counterfactual Donkeys. *Journal of Semantics* 23 (4): 383–402.

Wang, Y. (2009). Counterfactual Donkey Sentences: A Response to Robert van Rooij. *Journal of Semantics* 26 (3): 317–328.

Williamson, Timothy (2005). Armchair Philosophy, Metaphysical Modality and Counterfactual Thinking. *Proceedings of the Aristotelian Society* 105 (1): 1–23.

— (2007a). Philosophical Knowledge and Knowledge of Counterfactuals. *Grazer Philosophische Studien* 74 (1): 89–123.

— (2007b). *The Philosophy of Philosophy*. Oxford: Blackwell Publishing.

— (2009). Replies to Ichikawa, Martin and Weinberg. *Philosophical Studies* 145 (3): 465–476.

— (2013). How Deep is the Distinction between A Priori and A Posteriori Knowledge? In: *The A Priori in Philosophy*. Ed. by Albert Casullo and Joshua C. Thurow. Oxford: Oxford University Press.

Yablo, Stephen (2001). Go Figure: A Path Through Fictionalism. *Midwest Studies in Philosophy* 25 (1): 72–102.