Getting Gettier Straight. Thought Experiments, Deviant Realizations and Default Interpretations*

Pierre Saint-Germier

Abstract

It has been pointed out that Gettier case scenarios have deviant realizations and that deviant realizations raise a difficulty for the logical analysis of thought experiments. Grundmann and Horvath have shown that it is possible to rule out deviant realizations by suitably modifying the scenario of a Gettier-style thought experiment. They hypothesize further that the enriched scenario corresponds to the way expert epistemologists implicitly interpret the original one. However, no precise account of this implicit enrichment is offered, which makes the proposal somewhat *ad hoc*. Drawing on pragmatic theory, I argue that the content of Grundmann and Horvath's modified scenario corresponds to the default interpretation of the original scenario and that epistemological expertise is not required to access the deviance-proof interpretation. This Default Interpretation proposal offers thus a more general and independently motivated solution to the Problem of Deviant Realizations.

1 Introduction

According to a traditional understanding of the method of cases in philosophy (e.g. Bealer 1998), intuitive judgments elicited by the description of hypothetical cases in thought experiments provide premises for substantive philosophical arguments. Gettier's refutation of the justified-true-belief analysis of knowledge (JTB for short) provides the arch-example of this sort of procedure (Gettier, 1963). Consider for example the following Gettier-style scenario:

*This is a preprint of a forthcoming paper in Synthese https://doi.org/10.1007/s11229-019-02166-0. This version may differ from the published version.

(G) Suppose that 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 in the past. From this, Smith infers that someone in his office owns a Ford. Suppose furthermore that 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)

The description of this hypothetical situation is usually taken to elicit two intuitive judgments:

(1) Smith believes truly and with justification that someone in his office owns a Ford.

(2) Smith does not know that someone in his office owns a Ford.

These intuitive judgments are supposed to play a key role in the refutation of JTB. But how exactly do they contribute to the argument against JTB?

Taken literally, these judgments are not strictly speaking *true*: they are about a hypothetical person described as part of a thought experiment. So they cannot be straightforwardly identified as premises in an argument, if this argument is to be sound. Therefore, the judgments expressed by (1) and (2) must say something slightly different from what they seem to say, when taken at face value. Identifying the logical form of these intuitive judgments is known in the literature as the Content Problem (Malmgren 2011).

A natural starting point, suggested by Williamson (2007, p. 184), is to consider the following necessitation:

(NEC) $\Box\forall x\forall p(G(x,p) \rightarrow (JTB(x,p) \wedge \neg K(x,p)))$

Here the binary predicate "G" expresses a relation between a person x and a proposition p which holds whenever x is related to p in the same way that Smith is related to the proposition *that someone in his office owns a Ford* in the scenario (G); the binary predicate "JTB" expresses a relation which holds between a person x and a proposition p whenever p, x believes that p and x is justified in doing so; the binary predicate "K" expresses the relation which holds between a person x and a proposition p whenever x knows that p.

Even though it is not strictly speaking true that Smith does not know that someone in his office owns a Ford, it is presumably strictly speaking true that no one in his situation *can* know it. This is just what (NEC) says. Then, from (NEC) and the additional, independently plausible, assumption that (G) describes a *possible* case, we have the following valid argument to the conclusion that there could be true justified belief without knowledge, *contra* JTB:

	1.	$\Diamond\exists x\exists p G(x,p)$	Premise.
(**G**)	2.	$\Box\forall x\forall p(G(x,p) \rightarrow (JTB(x,p) \wedge \neg K(x,p)))$	Premise.
	3.	$\Diamond\exists x\exists p(JTB(x,p) \wedge \neg K(x,p))$	By 1, 2, and modal logic.[1]

Both premises of the argument seem true, so it looks like we have the sound refutation of JTB that we were looking for.

Now this reconstruction faces a well-known difficulty. A necessitation, like (NEC), is true just in case the consequent is true in all the possible worlds where the antecedent is true. But it is not too difficult, although it may require some imagination, to find possible worlds where the scenario (G) is literally true, but where (1) or (2) is false (Williamson 2007, Malmgren 2011). It follows that premise 2 is not true, and so cannot be part of a sound argument against JTB. This problem is known as the Problem of Deviant Realizations.

Grundmann and Horvath (2014a, 2014b) have shown, interestingly, that there is a way to modify a standard Gettier-style thought experiment so that it becomes immune to the Problem of Deviant Realizations. They also conjecture that this modified scenario corresponds to the way expert epistemologists implicitly understand the original scenario. This conjecture is interesting for (at least) two reasons. First, it suggests an elegant solution to the Content Problem: It is enough to substitute to the predicate "G" in Argument **G**, a predicate "G*" which holds of x and p just in case x is related to p *as in the modified scenario* designed by Grundmann and Horvath, rather than in the original scenario (G). Second, if the experts have an advantage over non-experts in accessing this modified scenario, it may give some ammunition to the expertise defence (e.g. Williamson 2011) against the restrictionist critique of thought experimentation in philosophy based on the data gathered by experimental philosophers (Alexander and Weinberg 2007, Machery 2017). Grundmann and Horvath, however, offer no independent justification for the particular non-trivial implicit interpretation

[1] The argument is valid in the weakest normal modal logic **K**.

they hypothesize, which makes the whole conjecture look like an *ad hoc* hypothesis.

The goal of this paper is to provide this missing element of justification and to draw some consequences for the Content Problem and the role of expertise in thought experimentation. After diagnosing the sources of the Problem of Deviant Realizations (§2), I present Grundmann and Horvath's modified scenario (§3) and examine critically the Expert Interpretation Hypothesis they suggest (§4). Then I introduce the notion of default interpretation (§5) and argue that the default interpretation of (G) has the same content as Grundmann and Horvath's modified scenario (§6) but that epistemological expertise plays no crucial role (§7). The general solution to the Content Problem which emerges from this view is then fleshed out (§8).

2 The Problem of Deviant Realizations

The Problem of Deviant realizations arises for a certain type of *modal* solution to the Content Problem, where the intuitive judgment resulting from the Gettier thought experiment is construed as a necessitation claim, i.e. (NEC).

The difficulty stems from two independent sources. First, it follows from the logic of necessity that necessitation, i.e. the relation that holds between A and B whenever $\Box(A \rightarrow B)$, validates the principle *Strengthening the Antecedent*:

(SA) $\Box(A \rightarrow B) \vdash \Box((A \wedge C) \rightarrow B)$.[2]

In other words, if B is true in all the possible worlds where A is true, *a fortiori*, B will be true in all the possible worlds where A *and* C are true.

The second source of the problem is the inevitable incompleteness of the scenarios used in philosophical thought experiments to describe hypothetical cases. The thought experimenter cannot specify all the details of the situation, and will leave some irrelevant aspects of the situation unspecified. In the foregoing Gettier case, Smith may have dark hair, blond hair, or he may be bald. The scenario remains silent on this matter. It follows that there are multiple ways to fill in those details in order to make the hypothetical case more specific. Although most ways of filling the details of a Gettier case will not affect the intuitive judgment it elicits, it is possible to specify the details of the scenario in such a way that it elicits an opposed judgment. Here is an example offered by Williamson:

[2]This principle holds in the weakest normal modal logic **K**.

> . . . in the Gettier case, if the subject's inference to the true belief **p** from the false belief **q** bizarrely happens to trigger awkward memories or apparent memories that cast doubt on **q**, the effect may be to lose justification for **q** rather than to gain it for **p**. (2007, p. 185)

Following Williamson's recipe, let us add to the scenario (G) above, between the second and the third sentence, the following piece of text:

> Smith's inference to the belief that someone in his office owns a Ford triggers the memory that he had similar hallucinations and false memories in the past, under the influence of LSD. Given that he knows he took LSD this morning before going to work, he thereby loses justification for his belief that someone in his office owns a Ford.

Just like the original scenario elicited the intuitive judgment that Smith's inferred true belief is justified, this augmented scenario arguably elicits the contrary judgment that Smith's true belief that someone in his office owns a Ford is *not* justified. If so, Smith's true but unjustified belief will not constitute a counterexample to the JTB-analysis of knowledge.

Alternatively, one might augment the initial scenario (G) in such a way that it does not elicit the intuitive judgment that Smith fails to know that someone in his office owns a Ford. Following a suggestion by Malmgren (2011, p. 275), it is enough to add the following sentence at the end:

> Smith heard from Brown himself, who never lies, that Brown owns a Ford.

In this augmented version of (G), it seems that Smith after all knows that someone in his office owns a Ford, so that the negation of (2) is elicited, instead of (2) as in the original scenario.

Let us consider this phenomenon in more general terms. For any philosophical thought experiment, we can identify the following couple $\langle S, J\rangle$ where S is a description of a hypothetical case and J is a set of intuitive judgments that S is intended to elicit. Let us say that a possible world w is a *realizer* of S iff S is true in w. A possible world w is said to be an *intended realizer* of S iff S is a realizer of S and all the members of J are true in w. A possible world w is a *deviant realizer* of S iff w is a realizer of S and some member of J is not true in w. In particular, if j is a member of J, then a realizer w of S is a *j-deviant* realizer iff j is false in w. Thus

any possible world where the Williamson-style augmented scenario is true, is a *jtb*-deviant realizer of (G), where *jtb* stands for the intuitive judgment (1) and any possible world where the Malmgren-style augmented scenario is true is a k-deviant realizer of (G), where k stands for the intuitive judgment (2).

We are now in a position to see that (NEC) cannot be a premise in a sound argument against JTB. For the two kinds of deviant scenarios considered above differ from (G) only by the conjunction of some further sentences. So any realizer w of (G) is a realizer of these augmented scenarios. By (SA), it follows that the intuitive judgments (1) and (2) are true in all deviant realizers of (G). Yet in the *jtb*-deviant realizers identified by Williamson (1) is false and in the k-deviant realizers identified by Malmgren (2) is false. This is absurd, unless the augmented scenarios describe impossible cases. But even though the deviant realization of (G) described by Williamson is highly *improbable*, it is not for that reason metaphysically *impossible*, even less logically inconsistent. Thus (NEC) cannot be a true premise in a sound refutation of JTB.

A number of strategies have been proposed in the literature to neutralize such deviant realizations. One class of strategies focuses on the first source of the problem, namely the fact that necessitation satisfies (SA). The idea is to replace necessitation by another relation for which the counterpart of (SA) fails. Williamson (2007) and Häggqvist (2009, 2017) prefer to use a counterfactual instead of a necessitation relation, Gardiner (2015) and Geddes (2018) rely on normalcy counterfactuals, while Malmgren's (2011) proposal is to use a possibility operator. These proposals all have attractive features but also face important challenges, a full review of which exceeds the scope of this paper.

Another type of strategy, on which I would like to focus here, consists in reconsidering instead the other source of the problem, namely the incompleteness of the scenario. One might be tempted to think that the only way to block deviant realizations is to specify the details of the scenario in such a way that it ends up being true in only one possible world, or at least to rule out explicitly all the possible deviant realizations. Both options may seem humanly impossible to carry out (Williamson 2007, p. 185). However, Grundmann and Horvath (2014a, 2014b) have shown that Gettier cases *can* be described, and plausibly be understood, in such a way that no deviant realizations can arise. This opens up a distinct approach to the Content Problem, according to which the difficulty has nothing to do with the logic of necessitation, but rather with the way deviant realizations can be ruled

out by an appropriate understanding of the thought experiment scenario.[3]

3 Deviance-proof Gettier scenarios

Let us consider Grundmann and Horvath's proposal in more detail. They offer a simple and efficient technique to rule out deviant realizations from a Gettier scenario such as (G), by applying a small number of modifications to it. As we have seen, one may distinguish two distinct classes of deviant realizers: the *jtb*-deviant possible worlds where Smith's true belief turns out to be unjustified and the *k*-deviant possible worlds where Smith turns out to know the truth he believes with justification, for reasons not explicitly mentioned in the scenario.

In order to rule out deviant realizers of the first type, Grundmann and Horvath recommend that the modified scenario *explicitly say* that Smith believes with justification the true proposition that someone in his office owns a Ford:

(GH_1) Suppose that Smith [***justifiedly***] 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 Smith infers, [***to the justified belief***] that someone in his office owns a Ford. Suppose furthermore that 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.[4]

This move may prompt the objection that it just trivializes the thought experiment. For it just *stipulates* the content of the intuitive judgment (1) instead of *eliciting* it. But what makes the whole scenario challenging for JTB is its capacity to elicit *at the same time* (1) and (2). And it would serve exactly the same dialectical purpose if (1) were just stipulated and (2) elicited. So trivializing the elicitation of the first intuitive judgment does

[3] Ichikawa and Jarvis's (2009) proposal, which relies on the notion of truth in fiction, also falls under this type of strategy. According to them, the principles governing the interpretation of fiction are sufficient to rule out deviant realizers. See however Williamson (2009) for some criticism.

[4] The modifications proposed by Grundmann and Horvath are inserted between brackets, in italic characters. As there will be successive layers of modifications, ***the latest modifications are in italic boldface characters***, whereas *those belonging to a previous layer are just in italic.*

nothing to diminish the dialectical force of the thought experiment, as long as the other intuitive judgment is not similarly stipulated in the scenario. One may still worry that this stipulative labelling creates some need on the author's part to add an argument justifying the legitimacy of this labelling and the coherence of the case so enriched. For if readers just intuit for themselves that Smith's beliefs are justified, it counts towards the epistemic credentials of the final verdict. To this worry, it might be responded that the overall coherence of the case described by (GH_1) does not seem more problematic than the coherence of the case described by (GH), so that this additional demand, in this particular case, should not be difficult to meet. However, the fact that this demand is easily met in this case does not mean that it is always easily met. So even if the stipulative move is acceptable with this Gettier scenario, it should not be expected to be always available.[5]

The next task is to add sufficient explicit information in order to rule out k-deviant possible worlds. The original Gettier scenario describes a process of belief formation which violates a plausible necessary condition for knowledge. It seems, prima facie, that a true belief inferred from a *false* belief should not count as knowledge, however justified it might be.[6] There are two reasons why such a description is not sufficient to rule out k-deviant possible worlds, however. First, it does not rule out the possibility that this very process of belief formation has some other epistemic features, not explicitly mentioned in (G), which are sufficient to turn the resulting belief into knowledge. For example, Smith may believe that someone in his office owns a Ford on the basis of one false belief, e.g. that Jones owns a Ford, plus a number of other true and independently justified beliefs. Possibilities of this sort are ruled out by the following further modification of the scenario:

(GH_2) Suppose that Smith believes [*justifiedly*] 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***

[5] I am indebted to an anonymous reviewer for pressing this worry.

[6] I should add that this necessary condition is not universally accepted, as some epistemologists hold the view that knowledge, in some cases, can be obtained by inference from false beliefs (Warfield, 2005; Fitelson, 2010). However, this view remains controversial (Ball and Blome-Tillmann, 2014; Montminy, 2014; Schnee, 2015). Also, the kind of counterexamples that motivate the view follow a pattern that is not instantiated here. I am indebted to an anonymous reviewer for drawing my attention to this issue.

for that belief at t]. Suppose furthermore that 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.

By giving an *exhaustive* description of the epistemically relevant features of the process of belief formation, this scenario rules out the possibility that this very belief turns out to be knowledge for reasons unmentioned in the original scenario. It might be worried however that the phrasing of this modification really achieves this. For example, (GH_2) says that the logical inference from "Jones owns a Ford" and "John is in Smith's office" to "Someone in Smith's office owns a Ford" provides Smith's "only" justification for that belief at t. However, there are for sure a number of additional background beliefs that play a role in supporting Smith's conclusion, such as beliefs about the relationship between driving a car and owning it or beliefs about the stability of car ownership over time. So (GH_2) may not provide an *exhaustive* description of the epistemically relevant features of the case and therefore misdescribes the case. One way to circumvent this problem would be to provide a list of these additional background beliefs, but it seems hard to find a *complete* list of the relevant background beliefs. Another way would be to use a weaker formulation and say that Smith's logical inference provides the *decisive* piece of justification for his belief that someone in his office owns a Ford. This is compatible with some background beliefs playing some justificatory role but suffices to rule out k-deviant realizers.[7] For the sake of discussion, I will keep Grundmann and Horvath's formulation of (GH_2) and assume it is sufficient to rule out this first sort of k-deviant realizers.

There is however a second way Smith may know the true proposition that someone in his office owns a Ford, compatibly with all the information explicitly provided by (G): Smith may have antecedently acquired this knowledge by a wholly distinct source, for example by the reliable testimony of Brown himself. In order to rule out possibilities of this kind, it is sufficient to stipulate that *if* Smith knows the Gettier proposition, he knows it only in virtue of an epistemic source mentioned in the scenario. Thus the following scenario has no k-deviant realizers:

(GH_3) Suppose that Smith believes [*justifiedly*] 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.

[7] I am indebted to an anonymous referee for drawing my attention to this issue.

> 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*]. Suppose furthermore that 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***].

Thus with (GH_3), we do have an example of a Gettier-style scenario, sufficiently rich to rule out both *jtb*-deviant and *k*-deviant realizers. Consequently, the Problem of Deviant Realizations cannot arise. However, Grundmann and Horvath's modifications do not, by themselves, block the Problem of Deviant Realizations *as it affects the original scenario (G)*. If this modified scenario is to offer any help in solving the Content problem for (G), something more should be said about the relation between (G) and (GH_3). In fact, Grundmann and Horvath (2014a) at the end of their paper sketch a proposal in that direction.

4 The expert-interpretation hypothesis

Grundmann and Horvath's suggestion is that the content of (GH_3) just is the content that expert epistemologists spontaneously and implicitly assign to (G) when they imagine the hypothetical case:

> ... improved case descriptions turn out to be not only possible but also relatively simple. For this reason, something along the lines of [GH_3] seems like a good reconstruction of how the relevant *experts* implicitly interpret the original case description [G] when they perform the thought experiment. Arguably, no professional epistemologist would be seriously tempted to interpret [G] in any of the deviant ways that are left open by this description. Moreover, the way in which we amended the initial description [G] should be readily available and transparent to every professional epistemologist. It therefore seems quite plausible that professional epistemologists tacitly interpret [G] in a

> way that roughly corresponds to our improved description [G]. (2014a, p. 531-532)

Two things deserve to be noticed. First, Grundmann and Horvath claim that the interpretation, by professional epistemologists, of (G) as (GH_3) is prima facie plausible. They argue for this claim by pointing out that no professional epistemologist will spontaneously interpret (G) in any of the deviant ways mentioned above. This is probably true but beside the point. The relevant alternative is between interpreting (G) in a way which is *general* enough to leave those possibilities *open* and interpreting (G) in a way which is *specific* enough to rule them out. In order to interpret (G) as (G) *simpliciter* rather than (GH_3), professional epistemologists would not have to interpret (G) specifically in any of the deviant ways, but only to have an interpretation of the scenario which is consistent with them, just like a non deviant interpretation can be consistent with Smith being blond-haired, dark-haired or bald. Once this difference is appreciated, the considerations offered by Grundmann and Horvath in favor of the plausibility of their view are far from conclusive. A stronger way to support the view that professional epistemologists interpret (G) as (GH_3) would be to provide a *derivation* of that interpretation, based on an independent theory of interpretation. The need for such a derivation is all the more pressing that (GH_3) is a very specific interpretation which resulted from a non trivial step-by-step elaboration. What reason do we have to believe that expert epistemologist will end up understanding the scenario in precisely *that* way? Grundmann and Horvath offer none.

Second, Grundmann and Horvath interestingly restrict their proposal to a certain class of interpreters, namely *expert* epistemologists. If the main reason for finding the view plausible is that no professional epistemologist would seriously interpret (G) in any of the deviant ways, then, arguably, this reason equally applies to less sophisticated subjects. For it takes the imagination and the logical sophistication of an expert philosopher to come up with such bizarre or obviously unintended realizations of the scenarios. Thus, assuming the plausibility of the view restricted to professional epistemologists, one may wonder if the (GH_3)-interpretation will be natural only to experts, or also to naive subjects.

There may be reasons to favor the view that only experts will retrieve the (GH_3)-interpretation, but they would have to be different from the ones offered in this passage. One way to support the interpretation hypothesis would be to show that some piece of background knowledge or some specific ability possessed only by expert epistemologists is necessary to retrieve the

(GH_3)-interpretation. But in order to adjudicate this point, one also needs to specify the kind of process by which expert epistemologists reach the (GH_3)-interpretation. And on this crucial point, Grundmann and Horvath unfortunately remain silent. Without an account of this process, the expert interpretation hypothesis seems irreducibly *ad hoc*, and thus of little help to solve the Content Problem.

5 Default interpretations

If it is plausible at all that readers of philosophical thought experiments spontaneously enrich the scenarios offered to them, these enrichments should be accountable by pragmatic theories designed to explain how utterances can communicate more than their literal content.

For example, in most contexts, if someone utters the sentence:

(3) Alice finished writing a paper and went to the pub

they communicate, and are understood by their audience to communicate, the information that

(4) Alice finished writing a paper and *afterwards* went to the pub.

What the sentence (3) says is just that there is a time t in the past such that Alice finishes writing a paper at t and a time t', also in the past, such that Alice goes to the pub at t' but does not by itself say anything specific about the ordering of t and t'. Yet, in most contexts, the utterance of (3) manages to convey a more specific information about the ordering of t and t', namely that t is before t'.

This well-known phenomenon can be seen as an instance of a wider class of phenomena sometimes referred to by the umbrella term of "default interpretations".[8] Like other pragmatic phenomena, default interpretations enrich the literal content of what is said and can be cancelled if further explicit content is added. Their distinguishing mark is that they go through *normally*, unless a specific feature of the context cancels them, as opposed to pragmatic inferences that occur *only if* a specific feature is present in

[8] These phenomena have been extensively studied by linguists under a varieties of conceptualizations: generalized conversational implicatures (Grice 1989a, Levinson 2000), implicitures (1994), explicatures (Sperber and Wilson 1995; Wilson and Sperber 2012), free pragmatic enrichments (Recanati 2010; Pagin 2014). See Jaszczolt (2018) for a survey.

the context.[9] For this reason, generalizations about default interpretations, according to some theorists at least (e.g. Levinson 2000) occur at the level of utterance-types, rather than utterance-tokens. These features seem to be shared by the spontaneous enrichments of thought experiment scenarios. They are cancellable, as the phenomenon of deviant realizations clearly show, and they do not seem to rely on specific features of the context of utterance, so that the level of utterance-types seems to be the appropriate level to study and explain them.

Levinson (2000) has offered a theory of default interpretations that one may take as a useful starting point to get a firmer grip on the explanation and prediction of these enrichments.[10] Default interpretations are there treated as generalized conversational implicatures and are generated by three *heuristics*:

- The Q-heuristic (*quality*): "What isn't said isn't"
- The I-heuristic (*informativeness*): "What is simply described is stereotypically exemplified"
- The M-heuristic (*manner*): "What is said in an abnormal way isn't normal"

Levinson's heuristics are subpersonal mechanisms evolved to make the computation of pragmatic inferences in real-time conversation more tractable and efficient (Levinson 2000, pp. 27-35). I will focus here on the Q- and I-heuristics, which are the most relevant for present purposes.[11]

Levinson's heuristics operate both on the speaker's and on the hearer's side. Let us start with the I-heuristic:

I-Principle

[9] In Gricean and neo-Gricean pragmatics, this contrast is what defines the difference between generalized and particularized conversational implicatures (Grice 1989a, p. 39; Horn 2004, p. 4-5). I take it that this contrast is still significant even if one thinks that default interpretations are not implicatures (but rather explicatures or implicitures).

[10] Some aspects of Levinson's theory are controversial among pragmaticists. I take it that these problems do not affect the general point made in the paper which should hold, *mutatis mutandis*, within alternative pragmatic theories. This issue is addressed in the next section.

[11] Since the Gettier scenario (G), and the scenarios of thought experiments more generally, do not usually employ marked or abnormal *expressions* (even though some thought experiments may admittedly describe highly abnormal *situations*), the M-heuristic will not be at work here.

> *Speaker's maxim: the maxim of Minimization.* "Say as little as necessary"; that is, produce the minimal linguistic information sufficient to achieve your communicational ends ....
>
> *Recipient's corollary: the Enrichment Rule.* Amplify the informational content of the speaker's utterance, by finding the most specific interpretation, up to what you judge to be the speaker's m-intended point, unless the speaker has broken the maxim of Minimization by using a marked or prolix expression. (Levinson 2000, p. 114)

By "m-intended" point, Levinson refers here to the kind of reflexive intention involved in speaker-meaning, according to Grice: an utterer U m-intends to produce effect r on an addressee A with an utterance x just in case "U utters x intending A (1) to produce r; (2) to think U intends A to produce r; (3) to think U intends the fulfillment of (1) to be based on the fulfillment of (2)" (1989b, p. 105). The effect r is typically to form a belief, when the utterer wants to convey to the addressee some factual information about the world. In the case of thought experiments, the m-intended effect is rather to imagine a particular situation. The reflexive character of this intention is crucial: U not only intends A to imagine a particular situation s that U has in mind, but also intends A to recognize this very intention and to imagine s as a consequence of this recognition. The role of the heuristics is then to ensure that A ends up imagining the situation intended by U in a cognitively tractable way.

Let me now unpack some important notions involved in the characterization of the I-principle. First, the preferred interpretation A' of an utterance A is more *informative*, in the sense that the possible worlds it rules out form a superset of the possible worlds ruled out by A. Second, this interpretation A' is more *specific*, which means that, in addition to being more informative, A' preserves the relations between the entities referred to in A: each term or relation in A' has a denotation that is a subset of the denotation of the corresponding expression in A. This means that a more specific interpretation A' will not achieve more informativeness by adding further unrelated objects or relations in the represented situation, but rather provide more informative descriptions of the objects and relations which are already described in A.

In particular, the I-heuristic prompts the hearer to:

(Ia.) assume the richest temporal, causal and referential connection between described situations or events, consistent with what is taken

for granted,[12]

(Ib.) assume that stereotypical relations obtain between referents or events, unless this is inconsistent with (a),

(Ic.) avoid interpretations that multiply entities referred to (assume referential parsimony). (Levinson 2000, p. 114-115)

The first submaxim can be seen as a drive towards *coherence*; the second one as a drive towards *stereotypicality*, the third one as a drive towards *parsimony*.[13]

These factors in the determination of a default interpretation are also recognized by other theorists. For example, Pagin's Enrichment Theory (2014) explains free pragmatic enrichments (a subclass of Levinson's) as *coherence*-raisers, acknowledges *Resemblance* as coherence-relation and introduces a *No New Entity* Principle to account for the *simplicity* of the available enrichments.

The drive towards coherence thus explains conjunction buttressing, i.e. (3)-(4) above, as well as the phenomenon of *bridging*:

(5) a. John unpacked the picnic. The beer was warm.

b. $\rightsquigarrow$ The beer was part of the picnic.[14]

The drive towards *stereotypicality* explains enrichments such as:

(6) a. John said "Hello" to the nurse and then John smiled.

b. $\rightsquigarrow$ John said "Hello" to the female nurse and then John smiled.

and the drive towards parsimony explains the preference for *local co-reference*:

[12]This phrasing should not be taken to imply that the I-heuristics ends up selecting a *unique* and *maximally specific* situation, where all the details would be fleshed out. Rather, the I-heuristic prompts the hearer to maximize such connections in her interpretation of the scenario, which means that if several more specific interpretations are available, the hearer will be prompted to choose an interpretation that maximizes these connections. This does not mean that the interpretation will correspond to a unique maximally specific situation but only to a relatively more constrained interpretation.

[13]See Fischer and Engelhardt (2017) for a useful review of recent psychological work on stereotypes and their relevance for the study of philosophical intuitions.

[14]The wavy arrow symbol is used to express a relation of pragmatic inference between an utterance (on the line above) and a sentence (on the right-hand side) expressing an aspect of the preferred interpretation of that utterance.

(7) a. John came in and he sat down.

b. $\rightsquigarrow$ John_1 came in and he_1 sat down.

Let us now turn to the Q-heuristic ("What isn't said isn't"). It also divides into a speaker's maxim and a hearer's maxim:

> *Q-Principle:*
>
> *Speaker's maxim*: Do not provide a statement that is informationally weaker than your knowledge of the world allows, unless providing an informationally stronger statement would contravene the I-principle. ...
>
> *Recipient's corollary*: Take it that the speaker made the strongest statement consistent with what he knows. (Levinson 2000, p. 76)

Levinson is specifically interested in the phenomena of scalar implicatures, i.e.

(8) a. Did the students come to the lecture?

b. Some came.

c. $\rightsquigarrow$ the speaker of (8b) knows that not all the students came.

but his Q-heuristic can arguably be put to work to explain other related phenomena such as exhaustive interpretations of answers, i.e.:

(9) a. What did you have for breakfast?

b. A croissant and a cup of coffee.

c. $\rightsquigarrow$ A croissant and a cup of coffee, and nothing else.

The speaker of (9b) presumably knows what she had for breakfast, and if she had had more than just a croissant and a cup of coffee, she could and would have said so—unless some specific aspect of the context would prevent her from doing so. So the hearer's default interpretation is that she did not have more than a croissant and a cup of coffee.

We have now a clearer picture of the principles governing default interpretations. Assuming that this general approach is valid, we are in a position to determine how a Gettier scenario like (G) will be interpreted by default.

6 Default Gettier cases

In this section, I argue for the claim that (GH_3) corresponds to the default interpretation of (G). In other words, readers of (G) interpret it *by default* as having the same content as (GH_3). If we characterize content in terms of possible worlds, then the claim is that the content of the default interpretation of (G) is the set of possible worlds where (GH_3) is literally true. Is this plausible?

Since I am interested in the content of (GH_3), rather than the exact way this content is expressed, what I need to show is that the *jtb*-deviant realizers and the *k*-deviant realizers excluded by (GH_3) are also excluded from the interpretation of (G) favored by the I-heuristic and Q-heuristics.

The first step of the argument is to show that the heuristics favor an interpretation where Smith's belief that someone in his office owns a Ford is *justified* in addition to being true. There are two ways this may fail to be the case. First, Smith's belief that Jones owns a Ford, based on visual perception, memory and a number of background beliefs (e.g. "if Jones comes to work everyday with the same car, he owns that car") may turn out to be unjustified. Second, assuming that Smith's belief that Jones owns a Ford is justified, Smith's inference to the belief that someone in his office owns a Ford may not preserve justification. Both possibilities, however, are ruled out by the I-heuristic.

Consider first the belief that Jones owns a Ford based on visual perception, memory and background beliefs. As we saw, there are many possible instantiations of the scenario where the justification of those beliefs will be defeated. For example, it is consistent with the letter of the scenario that Smith's visual perceptions are hallucinations, if Smith is an addict, swallowing LSD everyday before work and hallucinating Jones coming to work on a Ford. But adding a defeater would go against all the driving forces of the default interpretation. First, in *normal* conditions, we are justified in believing that what we see is the case, when what we see is our colleague arriving to work in a familiar car. Similarly, in normal conditions, we are justified in believing what we remember to be the case, if what we remember is a familiar fact. In addition, the background beliefs that connect what Smith sees and remember to the belief that Jones owns a Ford and, in normal conditions reasonable beliefs to hold. The fact that the background belief that connects coming everyday to work with the same car and owning that car turns out to be false in the Gettier situation comes from an unusual feature of the situation. Unless Smith has some particular reason to doubt that this connections holds, he can be taken to be justified in holding this belief. Denying

this amounts to denying that most of the beliefs that guide our daily lives are justified. So the drive towards *stereotypicality* favors an interpretation where Smith's beliefs based on perception are justified. Second, *justification* is itself a relation which provides coherence to discourse. Pagin (2014, p. 72) takes it to be among the strongest type of coherence relation, so the drive towards *coherence* also recommends to choose an interpretation where the prima facie justification of Smith's belief remains undefeated. Third, an interpretation which adds gratuitously a defeater to Smith's belief, e.g. a daily intake of LSD, even though no defeater is explicitly mentioned, would contravene the drive towards *parsimony*. Therefore, we can conclude, from properties associated with the I-heuristic, that according to a default interpretation of the scenario, Smith's belief that Jones owns a Ford is justified.

Let us now turn to Smith's inference to the belief that someone in his office owns a Ford. Let us first observe that strictly speaking, the inference is not logically valid unless it is assumed that Smith also believes that Jones is in his office. The inference as described is at best enthymematic. However, there is nothing particularly problematic or unusual in reasoning enthymematically or attributing enthymematic reasonings to people when one describes how they arrived at some belief. When the suppressed premise is obvious, there is no difficulty in retrieving it. In fact, the I-heuristic, via the maxim of Minimization, pushes speakers to use enthymemes, rather than fully formed deductive arguments, if they can expect the missing premise to be available to the hearer. This is presumably what happens in (G): in normal circumstances, the people you see coming to work everyday are your colleagues, so a normal instantiation of the scenario is one where Jones is indeed in Smith's office. And normally, if someone is in your office, you know it, and *a fortiori*, you believe with justification that he or she is in your office. So even if this additional belief is not explicitly mentioned, it is naturally accommodated by the reader of the scenario (and the writer can expect her to accommodate it). Furthermore, this interpretation gives a much greater degree of connection between the events described, and in particular the epistemic states ascribed to Smith, which conforms to the drive towards *coherence*. Of course, there are instances of the description of this inference in (G) where Smith's inferred belief is not justified. Williamson's example of a *jtb*-deviant realizer is a case in point. However, it is clear that the belief-forming process described by Williamson is highly unusual and highly abnormal, given the background knowledge, contrary to the drive towards *stereotypicality*. In addition, Williamson's *jtb*-deviant realization of the original scenario introduces entities and relations not explicitly mentioned ("awkward memories") contrary to the drive towards *parsimony*. So

we can be confident that the I-heuristic forces an interpretation of the scenario which makes Smith's belief that someone in his office owns a Ford both true and justified and thus rules out all *jtb*-deviant realizers.

The second step of the argument is to show that *k*-deviant realizers are ruled out by default. As we saw in section 3, there are two ways Smith might turn out to *know* that someone in his office owns a Ford for reasons other than the ones explicitly described in the scenario. On the one hand, it could be that the inferred belief that someone in Smith's office owns a Ford enjoys some unmentioned epistemic credentials sufficient to make it count as knowledge. As we saw, the I-heuristic favors an interpretation where the inference is just an elementary logical inference from two justified beliefs. One of this belief is explicitly mentioned, the other is easily accommodated, following the principles of the I-heuristic. Now for the resulting belief to count as knowledge, it would be necessary to add further premises to the inference. However, such premises are not as easy to accommodate. The I-heuristic does not compel us to introduce them. This possibility is then ruled out by the Q-heuristic which favors an interpretation such that what is not mentioned is simply not present unless required by the I-heuristic: if the writer had desired to describe a case where Smith's inferred belief had such additional epistemic features, she would have mentioned them. Since she did not, the reader understands (and is expected to understand) the scenario as describing a situation where such additional epistemic features are just not there.

On the other hand, it could be that Smith turns out to know that someone owns a Ford by having antecedently formed such a belief on totally different grounds, for example on the basis of the reliable testimony of Brown himself. Although there is absolutely nothing unusual or abnormal about knowing this proposition in such a way, the addition of this information to the interpretation of the scenario is not required by the I-heuristic: unlike the addition of the missing belief in the enthymematic reasoning, adding this state of knowledge would not strengthen the connection between the events described in the scenario. Moreover, if Smith had such a nontrivial piece of knowledge in the situation the writer intends to describe, she would explicitly mention it, otherwise the addressee would have no chance to include it in the interpretation she builds on the basis of what is explicitly said or written. Remember that the "m-intention" of the author is not only to lead the addressee to imagine a particular situation, but also to lead the addressee to imagine that situation *as a result of the expression of that very intention*. More precisely, this pragmatic enrichment is a consequence of the Q-Principle. An alternative description of a case where Smith is explic-

itly said to have in addition some independent source of knowledge for the proposition that someone in his office owns a Ford would be informationally stronger than (G). If the author of the description intended to convey such a case, uttering (G) would violate the speaker's maxim of the Q-heuristic. For the author of the description presumably knows whether in the situation she intends, Smith does or does not have such an independent source of knowledge. The recipient corollary of Q-heuristic then prompts the addressee to assume that the author made the strongest statement consistent with what she knows about the described situation. If the description does not mention any independent source of knowledge, then the preferred interpretation will be one in which Smith does not have any independent source of knowledge for that proposition. Thus all k-deviant realizers are excluded from the default interpretation of the scenario.

The conclusion we reach, then, is that the content of the modified scenario (GH_3) designed by Grundmann and Horvath to exclude all possible deviant realizations, can be derived as the content of the default interpretation of the original scenario (G), on the basis of independent pragmatic principles.

Of course, one might resist this conclusion by questioning the cogency of the theoretical framework within which this default interpretation was derived. Even though Levinson's (2000) theory of presumptive meanings is arguably an influential framework for explaining pragmatic enrichments of the kind we are interested in, it is by no means the only one, and it has not remained uncriticized. For example, some critics are skeptical about the very idea of generalized conversational implicatures (Carston 2002, p. 111). Others question the psychological plausibility of Levinson's account of Q-inferences (Noveck and Sperber 2007) or the homogeneity of his I-inferences (Jaszczolt 2005, pp. 40-43). However, the derivation offered above of a default interpretation equivalent to (HG_3) on the basis of the scenario (G) does not depend on the contentious features of Levinson's theoretical framework. It relies on features of default interpretations, associated in Levinson's framework to his I- and Q-heuristics which are in themselves uncontroversial enough and should be accommodated by rival frameworks. For example, a Relevance Theorist will prefer to explain the same phenomena with the Communicative Principle of Relevance rather than Levinson's heuristics. But it is plausible that coherence, normality and ontological parsimony maximize relevance in most contexts, and in the present case in particular. Similarly, Relevance Theory can accommodate the scalar and exhaustivity inferences explained by the Q-heuristic (Carston 1998).

Thus, Grundmann and Horvath's suggestion that (GH_3) corresponds to the way (G) is spontaneously interpreted finds some independent justification

from pragmatic principles of default interpretation. But Grundmann and Horvath's suggestion is more specifically that *epistemologists* will interpret (G) in this way. Is that additional suggestion plausible?

7 Default Gettier cases without expertise

The principles used to derive the deviance-proof interpretation of the scenario are pragmatic principles which explain how competent speakers interpret discourse in their natural language. These principles, i.e the Q- and the I-heuristic, will be operative for most competent speakers. No step in the derivation of the interpretation we gave in the previous section seems to require the possession of some expertise in epistemology. So Grundmann and Horvath's additional suggestion regarding the role of expertise is prima facie unmotivated.

It might be objected that the I-heuristic, and default interpretations in general, draw on background beliefs about the world to construct a preferred interpretation. A key component of these background beliefs is the possession of stereotypical representations for various objects, properties and relations. Now, although the principles governing the activation of prototypes in the interpretation of discourse are relatively stable across subjects, it is well-known that the content of stereotypical representations may vary from (groups of) subjects to (groups of) subjects and as a function of the degree of exposure to, and familiarity with, their referents (e.g. Giora 2003). And it is quite plausible that expertise has an effect on stereotypical representations. For example, a musicologist and a musically illiterate teenager will probably have different stereotypes of XXth century classical music. This will presumably make a difference to the way they interpret pieces of discourse mentioning this kind of music. Presumably, the representations of the musicologist will be more accurate, both in the sense that they will be more specific and that they will be closer to the way XXth century classical music actually sounds. Should we not expect the same kind of effect in the case of Gettier scenarios? After all, should the expert epistemologist not have more accurate stereotypes of *knowledge* and *justified belief* than the layperson? Although this is far from being implausible, it is hard, in the present case, to see how these better stereotypes would give experts in epistemology a decisive advantage over laypersons. For the types of epistemic states and processes described in the scenario (visual perception, memory, elementary logical inference) are familiar states and processes, to which any normally socialized human being is exposed. So the default interpretation of these de-

scriptions should be fully available to the average layperson and it is hard to see how expert epistemologists should have a significantly higher probability of accessing spontaneously this interpretation than the layperson.

One might also object that expert epistemologists have an advantage less due to their background knowledge, than to some procedural knowledge of how thought experiments work. For the the original scenario is embedded in the distinctive practice of thought experimentation which is quite foreign to the layperson. A skill in thought experimentation may make a difference in the interpretation of the scenario, in that it may help to better grasp the point intended by the author of the thought experiment. Levinson himself acknowledges that the I-heuristic is sensitive to "what [the reader] judges to be the speaker's m-intended point" (Levinson 2000, p. 114). The ultimate m-intended point of the author A of the (G) scenario, when she communicates it to her reader, is presumably that her reader (i) imagines a possible case of true justified belief without knowledge, (ii) thinks that A intends her to imagine a possible case of true justified belief without knowledge and (iii) thinks that A intends the fulfillment of (i) to be based on the fulfillment of (ii). An expert epistemologist, who is used to making conceptual distinctions between various epistemic states may then have a better way to fulfill (ii) and (iii), and thus have a lower probability of deviating from the intended Gettier case. For example, a reader who clearly grasps the dialectical situation and sees clearly that the purpose of the thought experiment is to refute JTB will presumably be careful to exclude jtb-deviant realizers and k-deviant realizers from his interpretation of the case. Even if this is true, such a familiarity with the dialectics of the thought experiment is far from being a *minimal* requirement to form an interpretation of the case which excludes both types of deviant realizers. As we saw in the previous section, the Q- and I-heuristics guarantee that the author fulfills this intention by presenting to her reader the scenario (G). So, even if expert epistemologists may have more resources to rule out deviant realizations, compared to laypeople, it does not follow that the layperson does not have sufficient resources to do so as well.

This conclusion is consistent with the empirical evidence gathered by experimental philosophers on Gettier cases. If some expertise were required to access the default interpretation, one would expect the layperson to be more prone to be misled by deviant realizations than experts. However, differences between laypeople and experts are rarely observed and remain small whenever they are observed.[15] In particular, Machery et al. (2016) failed to

[15]See Machery (2017, chapter 2) for a useful review of the literature. See Nagel et al. (2013) for a study of lay responses to Gettier cases.

replicate Weinberg et al.'s (2001) initial finding that intuitions about Gettier cases vary with cultural background, and found that lay subjects from four different cultures (American, Brazilian, Indian and Japanese) agreed with the philosophical consensus. If some epistemological expertise played such a decisive role, we should expect those lay subjects to deviate more from the philosophical consensus. Second, in the few reported Gettier cases where laypeople dominantly ascribed knowledge (Starmans and Friedman 2012; Turri 2013), it is difficult to explain this by a deviant understanding of the scenario. The effect appears to come from the presence of "authentic" (as opposed to "apparent") evidence in the scenario (Starmans and Friedman 2012), or from a failure to distinguish clearly the respective sources of good and bad epistemic luck in the case (Turri 2013). None of those explanations involve a deviant interpretation of the scenario. So, contrary to the Expert Interpretation Hypothesis, there are independent reasons to think that expertise in epistemology does not provide a decisive advantage to avoid deviant realizations in the interpretation of standard Gettier cases.

8 Solving the Content Problem

If the default interpretations of Gettier scenarios like (G) rule out their deviant realizations, then an elegant solution to the Content Problem is available: the logical form of the intuitive judgment elicited by scenario (G) is not (NEC) but

(D-NEC) $\Box\forall x\forall p(DG(x,p) \rightarrow (JTB(x,p) \wedge \neg K(x,p)))$

where the binary predicate DG holds between x and p just in case x is related to p as Smith is related to the proposition that someone in his office owns a Ford in a *default interpretation* of (G).

The logical form of the judgment is still a necessitation, like (NEC). But since default interpretations rule out all deviant realizers, the Problem of Deviant Realization does not arise. It is then easy to construct an argument **DG** which is just like Argument **G** above, except that the predicate DG is substituted everywhere for G. The argument **DG** is clearly valid and sound, as the reasons to doubt the truth of the second premise no longer hold.

This solution improves on the limitations of Grundmann and Horvath's earlier proposal. First, it does not rely on an *ad hoc* hypothesis connecting the scenario to a modified version of it which is immune to deviant realizations. Second, this solution is not restricted to a particular thought experiment, but straightforwardly generalizes to other thought experiments,

since the pragmatic principles relied upon are general enough to be applied to any other scenario.

Whether this solution generalizes *correctly* to all other thought experiments is another matter. All that has been shown so far is that the default interpretation of a particular scenario, i.e. (G), rules out the deviant realizations of that particular scenario. It is of course in itself a rather weak basis for an inductive generalization to all scenarios of all philosophical thought experiments. One way to argue for the general adequacy of this solution is to extend our inductive base by considering other examples of scenarios, the literal content of which admits deviant realizations, while their default interpretations do not. This would require reviewing a large number of cases, which cannot be done here. But I will present one such case, with the hope to confer at least some plausibility to the general adequacy of the view. Consider the following standard description of a trolley case:

(T) Albert is the driver of a runaway tram which he can only steer from one narrow track on to another. Five men are working on track A. One man is working on track B. Anyone on the track he enters is bound to be killed. (cf. Foot 1967, p. 3)

It is usually assumed in discussions of this case, that this scenario elicits at least the following intuitive judgment:

(10) Albert is morally allowed to steer the tram to track B.

However, it is not too difficult to construct a deviant realization of this case. Consider a possible world w_1 where (T) literally holds, but where there are on track B five other men who are not working but merely watching the worker referred to in (T). In such a possible world, steering to track B would now cause the death of six men, instead of five if Albert steered the tram to track A. But then (10) seems false in w_1. So w_1 counts as a deviant realizer of (T). However, this possible world, and in fact any possible world where there are more than one man on track B, is ruled out from the default interpretation of (T) by the Q-heuristic. If the author had in mind a case where there are n people, where $n > 1$, (T) would be insufficiently informative. Moreover, the author could easily have provided the missing information. Since she did not, she presumably had in mind a case where there is at most one person on track B.

Another deviant realizer one might think of is a world w_2 where the number of workers on tracks A and B are exactly 5 and 1, respectively, but

where the tracks are connected in such a way that track B loops back to track A, while track A does not loop back to track B, as shown in Figure 1.

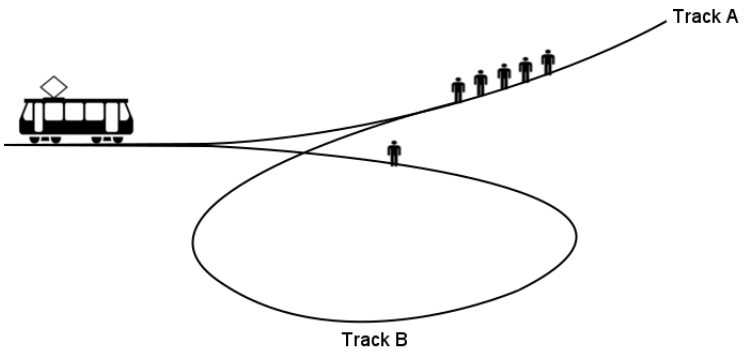

Figure 1: A deviant realizer of (T).

The literal content of (T) is true in w_2. Steering to track A will cause the death of five workers, but steering to track B will cause the death of one plus five workers. In that possible world, (10) is also false, so that this world will count as a deviant realizer. Again, interpreting (T) in such a way is prevented by the I-heuristic, since such a configuration of tracks is somewhat abnormal and would require an explicit description, had the writer had such a configuration in mind. Without constituting of course a conclusive proof of the proposed solution, this example may help make this general adequacy plausible at least.

There are additional general reasons to be confident about this generalization. A decisive counterexample would be a thought experiment scenario (S) which, just like the Gettier case and the Trolley case, elicits a clear intuition, and such that the default interpretation of (S) does not rule out all deviant realizers. But it is hard to see how such a thought experiment could elicit a *clear* intuition. Thus, unless a decisive counterexample is found, we can be confident that the proposal generalizes to most thought experiments that generate a stable intuition, if not all.

Conversely, the current proposal may also be useful to explain why some thought experiments fail to offer a stable intuition. For it entails that scenarios that are liable to prompt diverging default interpretations among different groups of readers or hearers. This may happen in several ways. If, for what-

ever reason, one scenario involves a concept to which the two populations attach different stereotypes, the resulting judgments might go in opposite directions. Or it could be that the scenario is not specific enough about a certain aspect of the case which various subjects interpret by default in opposed ways. I do not claim that all variability in the judgments elicited by thought experiment scenarios have to be explained in that way, as there are presumably other sources for this instability, such as framing effects or order effects.[16] The present proposal provides however a new way of explaining the variability of judgments elicited by some problematic thought experiments.[17]

At this point, one might still find the current proposal *ad hoc* in another way: why should we care in the first place about the *default interpretations* of scenarios, rather than their literal content? There is however a principled reason for that. The author of a thought experiment seeks to introduce a hypothetical singular case by description. The description may be purely verbal or verbal-cum-pictorial. In any case, the descriptive resources are usually insufficient to provide a complete description of the case. The only way the designer of a thought experiment can communicate successfully her intended case to an audience is to give the smallest amount of information about the case which guarantees that the reader will grasp the case as it is intended (*modulo* some irrelevant differences like the difference between Smith being dark-haired, short-haired or bald in the Gettier case).

The incompleteness of the description would be a problem if one thought of the relation between the scenario of the thought experiment and the resulting intuitive judgment as analogous to the relation between the assumptions of a mathematical proof and its logical conclusion. In thought experiments this link is intuitive rather than deductive. That is what makes thought experiments so useful and so fragile—useful because they allow to deal with matters hardly accessible by purely deductive reasoning from explicit assumptions; fragile because intuitive links are comparatively much more elusive than deductive links. Because this link between the scenario and the resulting intuitive judgment is not deductive, one need not restrict our attention the literal semantic meaning of the scenario. And because the hypothetical cases, unlike the explicit assumptions of a mathematical proof,

[16] See (Machery, 2017, chapter 2) for a review of the empirical findings.

[17] For example, fake barn cases are known to elicit divided judgments among experts (Horvath and Wiegmann, 2016) and as well among naive subjects (Nagel et al., 2013; Turri, 2017). One strategy to explain this instability would be to appeal to a diversity of default interpretation among subjects. Of course, the implementation and defense of such a strategy would require much more work that can be done within the scope of this paper. I am indebted to an anonymous reviewer for suggesting this example.

cannot be described exhaustively in full detail, we should expect writers and readers to rely on default interpretations to maximize the mutual understanding of the hypothetical case they jointly consider. So given the specific type of communication involved in the description of a hypothetical case, we should expect default interpretations to play a crucial role in the determination of the case being introduced, and therefore in fixing the content of the intuitive judgment regarding that case.

However, this principled motivation for introducing default interpretations in a solution to the Content Problem prompts a further objection. For this solution seems to presuppose that thought experimentation essentially involves linguistic communication. But a thinker can for sure think herself of a Gettier case, never talk to anyone about it and come to the conclusion that Smith has a true justified belief which fails to count as knowledge. In cases like this, we still have the problem of fleshing out the logical form of the judgment that the thinker may then use as a premise in an argument against JTB. However, since there is no linguistic communication of a scenario, pragmatic inferences and a fortiori default interpretations cannot play any role. Thus, the objection concludes, the solution offered by (D-NEC) does not provide a fully general solution to the content problem.[18]

Several things can be said in response. First of all, it seems fair to say that this presupposition actually accompanies the standard formulation of the Content Problem, e.g. in (Malmgren, 2011). For it is usually taken as a given that a Gettier case is presented via a scenario, i.e. a linguistic description. And then, given a Gettier case under that linguistic mode of presentation, one wants an account of the logical form of the judgment elicited by this scenario. Furthermore, the additional difficulties raised by the existence of deviant realizations also presuppose that the scenarios are given under a linguistic mode of presentation. If this diagnosis is correct, then it seems fair to say that (D-NEC) provides a solution to the Problem of Content, as it is usually discussed in the literature.

Of course, this observation does nothing to cancel the point that there is also an important variant of the Content Problem in the case of a solitary thinker. In this case, however, the issue is importantly different. In fact, it is a little bit more difficult to deal with for we have no way to describe the content of the case mentally represented by our thinker, except by *stipulating* that it is equivalent to the content of a given linguistic description. However, if the pragmatic story told so far is correct, we should refrain from equating this content with the content of a scenario like (G). According to

[18] I am grateful to an anonymous reviewer for raising this objection.

the speaker's maxim of the I-principle, the formulation of (G) obeys a principle of minimization: "say as little as is necessary". In other words (G) gives the minimum of information required for an addressee to recover a richer intended interpretation. So the content of the mental representation of the intended case should be expected to be much stronger that the content of (G). Assuming further that our solitary thinker has in mind roughly the same Gettier case as the author of the scenario (G), we can conclude that the content of the case she intends would be as strong as the literal content of (GH_3). For according to the argument offered in section 6 above, (G) is what you need to say in order to convey to an audience the content of a case literally described by (GH_3). And so, we have a reason to think that the case as intended by our thinker will not be affected by the Problem of Deviant Realizations.

In the solitary thinker case, the Content Problem then receives a solution which is structurally similar to the (D-NEC) proposal for the case where the Gettier case is linguistically communicated:

(I-NEC) $\Box\forall x\forall p(IG(x,p) \rightarrow (JTB(x,p) \wedge \neg K(x,p)))$

where the binary predicate IG holds between x and p just in case x is related to p as Smith is related to the proposition that someone in his office owns a Ford in the hypothetical situation intended by the thinker. In fact, the two solutions to the two problems are less disconnected than it might seem. According to the pragmatic principles reviewed above, the default interpretation of a linguistic description such as (G) is supposed to reflect the case as it is intended by the author of the scenario. Of course a Gettier case can be thought of, and so intended, without being communicated. Thus (I-NEC) can be seen as a limiting case of (D-NEC), where no communication occurs. But whether or not communication occurs, necessitation is the basic relation that connects the relevant representation of the case (a mental representation case or a default interpretation) and the judgment about the case that may then be used as a premise in a philosophical argument.

9 Conclusion

Following a suggestion of Grundmann and Horvath's, I have argued that the Problem of Deviant Realization can be blocked by considering the way the scenarios of thought experiments are interpreted *by default*, rather than focusing on their literal content. This yields a pragmatic solution to the

Content Problem which improves on Grundmann and Horvath's initial suggestion, by being independently motivated and readily generalizable to other thought experiments. Another result is that expertise in epistemology plays a lesser role than Grundmann and Horvath have suggested. These results naturally invite to form a broader picture of thought experimentation where pragmatic factors play a key role. If the case made in this paper is sound, then pragmatics can offer valuable resources to those who want to understand how thought experiments, intuitive judgments and arguments mesh together to buttress or undermine philosophical views.

Although I have mostly focused in this paper on the epistemic benefits of default interpretations in Gettier cases, I should add that I do not mean to imply that they are intrinsically or always epistemically beneficial. We should be open to the eventuality that for some types of hypothetical cases, default interpretations driven by stereotypes can affect negatively the epistemic qualities of our judgments. The delineation of this class of thought experiments and the identification of the features that cause default interpretations to have this epistemically detrimental effect will however have to be reserved for another paper.[19]

References

Alexander, J. and Weinberg, J. M. (2007). Analytic epistemology and experimental philosophy. *Philosophy Compass*, 2(1):56–80.

Bach, K. (1994). Conversational impliciture. *Mind and Language*, 9(2):124–162.

Ball, B. and Blome-Tillmann, M. (2014). Counter closure and knowledge despite falsehood. *Philosophical Quarterly*, 64(257):552–568.

Bealer, G. (1998). Intuition and the autonomy of philosophy. In DePaul, M. and Ramsey, W., editors, *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, pages 201–240. Rowman & Littlefield.

[19]I would like to thank Synthese's two anonymous reviewers for valuable criticism that helped improve the paper, as well as the participants of the *Intuitions and the Expertise Defence* workshop in Aarhus (September 2017) for valuable feedback. Special thanks to Anna Daria Anna Drożdżowicz, Karen Kiil Brøcker, Anne Reboul, Samuel Schindler and Tom Schoonen for their comments on previous versions of this paper.

Carston, R. (1998). Informativeness, relevance and scalar implicature. In *Relevance Theory: Applications and Implications*, pages 179–236. John Benjamins, Amsterdam.

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*.

Fischer, E. and Engelhardt, P. E. (2017). Stereotypical inferences: Philosophical relevance and psycholinguistic toolkit. *Ratio*, 30(4):411–442.

Fitelson, B. (2010). Strengthening the case for knowledge from falsehood. *Analysis*, 70(4):666–669.

Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, 5:5–15.

Gardiner, G. (2015). Normalcy and the contents of philosophical judgements. *Inquiry : An Interdisciplinary Journal of Philosophy*, 58(7-8):700–740.

Geddes, A. (2018). Judgements about thought experiments. *Mind*, 127(505):35–67.

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

Giora, R. (2003). *On Our Mind: Salience, Context and Figurative Language*. Oxford University Press, Oxford.

Grice, H. P. (1989a). Logic and conversation. In *Studies in the Way of Words*, pages 22–40. Harvard University Press, Cambridge, London.

Grice, H. P. (1989b). *Studies in the Way of Words*. Harvard University Press, Cambridge, London.

Grundmann, T. and Horvath, J. (2014a). Thought experiments and the problem of deviant realizations. *Philosophical Studies*, 170(3):525–533.

Grundmann, T. and Horvath, J. (2014b). Erratum to: Thought experiments and the problem of deviant realizations. *Philosophical Studies*, 170(3):535–536.

Häggqvist, S. (2009). A model for thought experiments. *Canadian Journal of Philosophy*, 39(1):55–76.

Häggqvist, S. (2017). Thought experiments, formalization, and disagreement. *Topoi*, pages 1–10.

Horn, L. (2004). Implicature. In Horn, L. and Ward, G., editors, *The Handbook of Pragmatics*, pages 3–28. Blackwell, Malden MA, London.

Horvath, J. and Wiegmann, A. (2016). Intuitive expertise and intuitions about knowledge. *Philosophical Studies*, 173(10):2701–2726.

Ichikawa, J. and Jarvis, B. (2009). Thought-experiment intuitions and truth in fiction. *Philosophical Studies*, 142(2):221–246.

Jaszczolt, K. M. (2005). *Default Semantics: Foundations of a Compositional Theory of Acts of Communication*. Oxford University Press UK.

Jaszczolt, K. M. (2018). Defaults in semantics and pragmatics. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2018 edition.

Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. MIT Press.

Machery, E. (2017). *Philosophy within its proper bounds*. Oxford University Press, Oxford.

Machery, E., Stich, S., Rose, D., Chatterjee, A., Karasawa, K., Struchiner, N., Sirker, S., Usui, N., and Hashimoto, T. (2016). Gettier across cultures. *Noûs*, 50(4).

Malmgren, A. (2011). Rationalism and the content of intuitive judgements. *Mind*, 120(478):263–327.

Montminy, M. (2014). Knowledge despite falsehood. *Canadian Journal of Philosophy*, 44(3-4):463–475.

Nagel, J., Juan, V. S., and Mar, R. A. (2013). Lay denial of knowledge for justified true beliefs. *Cognition*, 129(3):652–661.

Noveck, I. and Sperber, D. (2007). The why and how of experimental pragmatics: the case of scalar inferences. In Burton-Roberts, N., editor, *Pragmatics*, pages 184–212. Palgrave Macmillan, Basingstoke.

Pagin, P. (2014). Pragmatic enrichment as coherence raising. *Philosophical Studies*, 168(1):59–100.

Recanati, F. (2010). *Truth-Conditional Pragmatics*. Oxford University Press.

Schnee, I. (2015). There is no knowledge from falsehood. *Episteme*, 12(1):53–74.

Sperber, D. and Wilson, D. (1995). *Relevance*. Blackwell.

Starmans, C. and Friedman, O. (2012). The folk conception of knowledge. *Cognition*, 124(3):272–283.

Turri, J. (2013). A conspicuous art: Putting gettier to the test. *Philosophers' Imprint*, 13.

Turri, J. (2017). Knowledge attributions in iterated fake barn cases. *Analysis*, 77(1):104–115.

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

Weinberg, J. M., Nichols, S., and Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics,*, 29(1-2):429–460.

Williamson, T. (2007). *The Philosophy of Philosophy*. Blackwell.

Williamson, T. (2009). Replies to ichikawa, martin and weinberg. *Philosophical Studies*, 145(3):465–476.

Williamson, T. (2011). Philosophical expertise and the burden of proof. *Metaphilosophy*, 42(3):215–229.

Wilson, D. and Sperber, D. (2012). *Meaning and Relevance*. Cambridge University Press.