

Sensitivity and Necessity

Simon Gaus

(SEN) S sensitively believes that p via method M just in case S would not believe that p via M if p were false.

- (SEN) as a necessary condition for justification: S's belief that p is justified only if it is sensitive.
- (SEN) as underminer/good grounds for rational belief
- (SEN) as a necessary condition for knowledge: S knows that p only if S sensitively believes that p

1. Sensitivity in Gettier cases

[Jill] works in a production plant. She hears a continuous, ringing noise. She has heard similar sounds before and has learned that, in general, these sounds are associated with a malfunction in one of the machines. She therefore comes to believe that one of the machines is malfunctioning at the moment. Unbeknownst to Jill, however, her earlier perception of a sharp, ringing noise was produced by a form of temporary tinnitus. Coincidentally, a machine was also malfunctioning at the time.

Intuition: Jill's belief is justified and true, but does not constitute knowledge.

Explanation: Jill's belief does not satisfy *sensitivity*: She would have believed that a machine is malfunctioning even if there had not been a malfunctioning machine.

Upshot: Jill's case is a paradigmatic example for the sort of case which motivates the claim that *sensitivity* is necessary for knowledge.

2. The Necessity Objection

[Actual Jill] ... (like Jill, but believes "actually, a machine is currently malfunctioning")

[Moral Carla] wishes to figure out whether she may break a promise to a friend to meet her for a coffee in order to be with her sick grandmother. She engages in ordinary moral reasoning, looks at her own intuition, and speaks with people that are widely respected as moral authorities. All of this points to the conclusion that she may indeed break the promise in order to be with her sick grandmother. Unbeknownst to Carla, however, she has been the subject of a Truman-show like experiment aimed at testing whether an arbitrarily chosen belief – that one may break a promise in order to help one's grandmother – can be instilled by systematically manipulating all parts of that person's upbringing, education and social environment. The moral authorities, her parents and school teachers have all been part of this experiment.

[Mathematical Bob] wishes to figure out the sum of 250, 3210 and -3418. He takes his calculator – which, so far, has proved itself to be very reliable –, types in the numbers and gets "42". Based on this, Bob forms the belief that the sum of 250, 3210 and -3418 is 42. Unbeknownst to Bob, however, his calculator has been

manipulated in such a way that whenever it is asked to calculate the sum of 250, 3210 and -3418, it gives as a result Bob's current age in years; and Bob is 42 at the moment.

Problem: These cases throw doubt on whether truth-sensitivity is of any epistemological import.

- (1) Mathematical Bob's, Moral Carla's and Actual Jill's beliefs are necessarily true.
- (2) Therefore, the counterfactual conditionals "If [belief-content] had been false, they would not have believed [belief-content]" is trivially true.
- (3) Therefore, Mathematical Bob's, Moral Carla's and Actual Jill's beliefs are truth-sensitive by default.
- (4) Therefore, truth-sensitivity cannot explain why those cases are not cases of knowledge. [i.e. cannot account for the epistemological difference between these cases and their 'good' mirror cases].
- (5) But whatever explains why the original Jill-case is a case of knowledge must also be what explains why Mathematical Bob's, Moral Carla's and Actual Jill's belief are not knowledge.
- (6) So truth-sensitivity cannot be what explains why original Jill does not have knowledge. [and the same holds for other apparent examples for the epistemological import of *sensitivity*].

3. Rebutting the Necessity Objection

Observation: We often treat counterfactual conditionals which we believe to have necessarily false antecedents as substantial, informative claims and take our reasons for beliefs to depend on their being true.

- (A1) "If Gödel's incompleteness theorem were invalid, someone would have found out by now. But no one has. So Gödel's theorem is not invalid."
- (A2) "Gödel was a brilliant deceiver. He could have made even invalid theorems appear valid to even the best of today's logicians. Thus, even if Gödel's proof were invalid, and even given how much attention has been paid to it, no one would have found out by now"
- (B1) "If water were not H₂O, my teachers would not have told me this and my fellow philosophers would not use the claim as the standard example for a necessarily true proposition. But they do so use it, and my teachers did tell me. So water is H₂O.
- (B2) "Your teachers told you all kinds of wrong things; and none of them were expert chemists. As regards your fellow philosophers, they are not expert chemists either, their general track record in coming up with correct descriptions of examples is not very good, and they don't really care about the details of their examples, anyway. If water had some other chemical structure, your teachers and fellow philosophers would very likely not have acknowledged this."

Conclusion: There are counterfactual conditionals with a necessary false antecedent that are neither trivially true nor trivially false and that can, if justified, serve as a good basis for otherwise unwarranted beliefs.

Objection 1: The antecedent of the counterfactual conditionals “*If Gödel’s proof were invalid, someone would have found out by now*” as used in (A), is not necessarily false. There is a world in which the proof Gödel developed is invalid – a world in which the proof contains a mistake that it actually does not have.

Reply 1: Substitute “Gödel’s proof” with “this proof” and imagine a situation in which two students look at a piece of a proof, don’t understand a particular step but remain convinced that the proof must be valid.

Objection 2: Making sense of the above examples as counterfactual conditionals would require a completely new account of how counterfactual conditionals work, or a completely new account of possible worlds that includes logically inconsistent worlds. This is a serious cost. Any view that can avoid it – by analyzing them as not being a counterfactual conditionals ‘deep-down’ or by maintaining that they are trivially true – is in one important respect superior.

Reply 2.1: (A) seems to be a counterfactual conditional just as much as any other example, and seems to be just as substantive as many other, more familiar ones. In light of this, we should regard making sense of the non-trivial truth of (A) as simply a further, underappreciated constraint on a satisfying theory of counterfactual conditionals.

Reply 2.2: Any satisfying account of the example claims above must yield the result that there is a sense in which the correctness of the utterances of (A1) and (B1) depends, among other things, on the empirical claims made in (A2) and (B2). This suffices for the sensitivity-proponent. Whatever the interpretation that allows us to make sense of utterances of (A1) and (B1), use that interpretation for the “would”-sentence that is ordinarily used to express the *sensitivity*-criterion.

Objection 3: There is principled reason why counterfactuals with logically or conceptually inconsistent antecedents come out as trivial on standard accounts: In order to assess a counterfactual conditional, we must be able to *conceive of* or *mentally simulate* the antecedent. But we cannot conceive of logical or conceptual inconsistencies.

Reply 3.1: We clearly *can* assess (A). So (A) is simply a reductio of the conjunction of the claims that we cannot conceive of logical impossibilities and that conceiving of the antecedent is necessary for assessing the counterfactual.

Reply 3.2: Tentative proposal: The sense of “conceiving” in which we cannot conceive of logical inconsistencies – and there is such a sense – is not the sense of conceiving in which we must be able to conceive of the antecedent of a counterfactual conditional. To see this, consider “If I could perceive as a bat perceives, Germany would crown me as king”. I cannot conceive of the truth of the antecedent in just the same way in which I cannot conceive of the truth of a logical inconsistency. Nonetheless, that counterfactual conditional has a non-trivial truth-value, and I know it (“false”).