A Defence of the Sensitivity Analysis of Knowledge

Eric Wallace

Abstract. It is widely considered that Nozick's sensitivity analysis of knowledge fails. This is largely due to arguments proposed by Sosa. In this paper, I defend Nozick's sensitivity condition on knowledge. To do this, I define and motivate sensitivity, then explain Sosa's definition of safety and its supposed advantages. For all three of Sosa's purported counterexamples, we will find that the sensitivity theorist can offer responses that are more satisfactory than those available to the proponent of Sosa's safety theory. Having motivated sensitivity and shown that it deals with these purported counterexamples better than safety, I will conclude that Sosa fails to motivate a move from sensitivity to safety.

One key question in epistemology is "what is knowledge?". Traditionally, a response to this question has been given in the form of necessary and sufficient conditions. One early analysis of knowledge was Nozick's sensitivity analysis. While this theory enjoyed a short-lived popularity, it has fallen into disrepute in large part due to the objections of Sosa. This paper aims to show that these objections miss their mark. The sensitivity analysis of knowledge comes out unscathed.

The Intuition Behind Sensitivity

The basic statement of sensitivity is as follows:

> **S**'s belief that **P** is sensitive iff, if **P** were false, **S** would not believe that **P**.[50]

The conditional on the right hand side is a subjunctive conditional. There are several proposed methods for modelling these, but Nozick opts to model them as follows: The conditional is satisfied iff, in the closest possible world where **P** is false, **S** does not believe that **P**.[51] The insight here is that beliefs that do not track truth cannot count as knowledge, and we can model truth tracking using a subjunctive conditional. A key reason for thinking that knowledge must track truth is that lucky knowledge, that is a belief that constitutes knowledge by chance, is implausible. A truth tracking condition rules out such cases of

[50]Robert Nozick, "Knowledge and Skepticism," In *Epistemology: an anthology* ed. Ernest Sosa and Kim Jaegwon, (Oxford: Blackwell, 2000), 79.

[51]Nozick, "Knowledge and Skepticism", 80.

lucky knowledge because beliefs that are true as a result of luck will not track truth.

The basic statement of sensitivity above needs to be amended to factor in methods of knowing. The amended definition is as follows:

> **S**'s belief that **P**, arrived at via method **M**, is sensitive iff, if **P** were false and **S** used **M** to come to a belief about **P**, **S** would not believe that **P**.[52]

Consider a person whose friend phones to say that they will be late. This friend is normally late, so, had their friend not phoned, they still would have assumed their friend was going to be late. While the belief formed based on the friend's testimony seems like knowledge, it fails the unamended definition of sensitivity. The amended definition classifies the belief that the friend will be late, formed based on testimony, as sensitive because if their friend had not been late, they would not have believed that they were going to be late via their testimony. I will use this amended definition of sensitivity throughout the paper. Sosa agrees with Nozick that knowledge must track truth. He also agrees that sensitivity is at least initially plausible as an attempt to capture truth tracking.[53]

Safety and Its Advantages

Sosa characterises safety as follows:

> **S**'s belief that **P** is safe iff, if **S** were to believe **P**, it would be so that **P**.[54]

Sosa's safety condition aims to capture some of the appeal of Nozick's sensitivity condition while circumventing some perceived problems. A safe belief tracks truth in the sense that it could not easily have been wrong. When working out whether a belief is safe, one considers all nearby[55] possible worlds

[52]Ibid., 81.

[53]Ernest Sosa, "How to Defeat Opposition to Moore," In *Philosophical Perspectives* no. 13 (1999) 141, 147; Sosa opts to use Nozick's unamended condition. Presumably this is because he did not think considerations of methods would be relevant to his criticisms. As we will see later, they are.

[54]Sosa, "How to Defeat Opposition to Moore," 141-153. Some use a definition that is amended in a similar way to Nozick's methods amendment. I opt to use this simpler definition for clarity, but everything I say also applies to the amended definition.

[55]One might worry that there is no principled way of working out how nearby a world must be to be relevant. Even though Sosa insists that the relevant alternatives theorist responds to an analogous point, I set this worry aside. This is because this seems to be an instance of a more general problem in the analysis of knowledge. Examples: How relevant does a relevant alternative have to be? How reliable does a cognitive mechanism have to be? How good does justification have to be? How closely tied to an epistemic virtue does a belief have to be? Given the generality of the problem, we had all better hope it can be fixed.

where **S** believes that **P**. If it is the case that **P** in these worlds, then **S**'s belief that **P** is safe.

Sosa gives several arguments that safety has advantages over sensitivity. I will focus on Sosa's counterexamples as his other arguments have been discussed elsewhere. To give reason for maintaining sensitivity in the face of safety, I will show that sensitivity gives better responses than safety.

False Belief Counterexample

I start with Sosa's false belief counterexample because he believes it to be the most compelling, and I agree with him.[56] Sosa asks us to consider a subject **S** who knows some proposition **P** and believes that they do not falsely believe that **P**.[57] I shall call **S**'s belief that they do not falsely believe that **P** their metabelief.[58] Sosa makes the following argument:

> (P1) All metabeliefs about propositions that constitute knowledge are knowledge.
>
> (P2) All metabeliefs about propositions that constitute knowledge are insensitive.
>
> ---
>
> (C3) Some knowledge is insensitive.

Should Sosa successfully show that some knowledge is insensitive, this is sufficient to show that the sensitivity theory is false. Sosa thinks that (P1) must be true because a minimally attentive and reflective agent who knows **P** surely must know that they do not falsely believe that **P**. He defends (P2) on the grounds that in the most likely scenarios where I falsely believe that **P**, I will still believe that **P**. Based on that, I will form a belief that I do not falsely believe that **P**, making my metabelief insensitive.

(P1), however, is not true. To see this, consider someone who has beliefs about the world formed in a manner such that some of them constitute knowledge. They then form metabeliefs not based on reflection upon how they formed their beliefs about the world but based on an arrogant conviction that all their beliefs are correct. Intuitively, it does not seem that a belief formed purely on this arrogant conviction can be knowledge. Such a metabelief not only does not intuitively feel like knowledge, but also is not safe, so the safety theorist cannot maintain that it is knowledge. Considering this, we must revise (P1) to factor in that some metabeliefs about knowledge are not knowledge as follows:

[56] Ibid., endnote 11.

[57] Ibid., 145.

[58] I introduce this nomenclature for brevity and legibility, though I appreciate that the term 'metabelief' is usually used more broadly.

> (P1*) Some metabeliefs about propositions that constitute knowledge are knowledge.

The new argument with (P1*) replacing (P1) is still valid. Consider someone who knows that they have over-brewed their tea via the method of tasting it, and suppose that they form a belief that they know that they do not falsely believe that they have over-brewed their tea via the method of reflecting on their tasting the tea. To evaluate whether this metabelief is insensitive, we must consider whether the agent would still hold their metabelief in the closest possible world where they do falsely believe that the tea is over-brewed *via the method of reflecting on their tasting it*. One plausible candidate for such a world is a world where the agent misremembers how long the tea has been brewing, forms the belief that the tea is over-brewed when it is not via deduction from their knowledge of how long it takes tea to over-brew, and then forms their metabelief by reflecting on their deduction. Importantly, in such a world the agent does not form their metabelief via the same method. In other words, the example meets Nozick's methods definition of sensitivity. The example above is an example of a sensitive metabelief. Hence, (P2) is false. If we amend (P2) as follows:

> (P2*) Some metabeliefs about propositions that constitute knowledge are insensitive.

Sosa is faced with the problem that the argument is now invalid.

Now I shall compare how safety responds to these cases. Suppose that knowledge is safe true belief. Consider a case where I believe that my friend will do as I asked them based on their testimony. Suppose that this belief is knowledge. In all likely scenarios where I believe that they will do as I ask, they will do as I ask. Suppose also that I believe that I do not falsely believe that they will do as I asked, not because I reflect on their testimony and their honesty, but because I believe myself to be a master manipulator, which I am not. Intuitively, it seems that my metabelief is not knowledge because it is formed based on a belief that is false. In all close worlds my metabelief arises only when I believe that they will do as I ask, and, by stipulation, there is no close possible world where I believe that they will do as I ask and they do not do so. Hence there is no close possible world where I hold my metabelief, and it is false. The metabelief is a safe true belief, so, for the safety theorist, it is knowledge. Yet it does not seem to be knowledge as it is based on false considerations. This does not show that safety is not a necessary condition for knowledge, but it does show that a fourth condition on top of safety, truth, and belief would be required to eliminate such a case. Hence, this case illustrates a problem facing the safety theorist that does not face the sensitivity theorist.

The Garbage Chute Counterexample

I will now consider Sosa's garbage chute case. In this counterexample, a person releases a bag of garbage down a garbage chute, and the garbage makes it to the bottom of the chute. The person believes that the garbage makes it to the bottom of the garbage chute. Sosa claims that this is a case of knowledge that is not sensitive as, if the garbage had somehow snagged, the person would have believed that it made it to the bottom of the chute.[59] There are three issues with this counterexample. First, it is not safe, so the safety theorist cannot claim that it is a case of knowledge. Second, we can construct structurally similar cases that are not knowledge. Third, cases like the garbage case are cases where belief does not track truth.

To be safe, the belief in the garbage chute case would need to be such that in all close possible worlds where the agent believes that the garbage is at the bottom of the chute, the garbage is at the bottom of the chute. The cases where the garbage does not get to the bottom of the chute are very unlikely so, Sosa claims, the worlds where these cases hold are not close worlds. Hence, even though the agent would believe that the garbage is at the bottom of the chute in cases where it snags, the belief is safe. This is all well and good so long as the worlds where the garbage snags are distant worlds. These scenarios being unlikely does not seem like good reason for thinking that these are distant worlds. To see this, suppose for a reductio that worlds at which unlikely events occur are distant. Imagine that I run a raffle with 500,000 of my friends. It is a fair raffle and exactly one of us will win. The chance that any of us win is 0.0002%, which is unlikely. In all worlds where our raffle takes place, one of us will win. That event had a 0.0002% chance of occurring, so all worlds are worlds where an unlikely event occurs. By our supposition, this means that all worlds are distant worlds. Given that all worlds are distant, no worlds are close. Hence, for any belief I hold, there is no close world where I am mistaken. Hence, all beliefs would be safe, simply because my friends and I ran a raffle. This is absurd, so our assumption that worlds at which unlikely events happen are distant was false. The safety theorist must now concede that the unlikeliness of the garbage snagging does not show that worlds where it snags are distant.

Standardly, worlds that we choose to model as close to the actual world are worlds that are objectively similar to the actual world.[60] If we adopt this view, then worlds where the garbage snags and very little else is changed are close as they are only slightly different to the actual world. All that varies is the location of the garbage and the exact shape of the chute. Given that worlds where the garbage snags are close and they are worlds where the subject in Sosa's example still believes that the garbage is at the bottom of the chute, the subject's belief that the garbage is at the bottom of the chute is not safe.

[59] Ibid., 145.

[60] This view is not without its issues, but it fairs much better than Sosa's unlikelihood view.

Hence, the safety theorist must concede that the garbage chute case is not a case of knowledge. Hence, the safety theorist cannot maintain that it is a counterexample to sensitivity.

Aside from the belief in question not being safe, we had better hope there is independent reason to think that the garbage chute case is not a case of knowledge. Without independent reason, it is hard to see why the garbage chute case should not be considered a counterexample to both sensitivity and safety.

First, it should be acknowledged that not everyone shares the intuition that this is a case of knowledge. I, for example, was very confused by Sosa's assertion that this belief is knowledge.

The case I propose is inspired by Lewis's talk of stakes when ascribing knowledge. We are less willing to tolerate margin for error in ascriptions of knowledge when the stakes are higher. Some choose to explain such cases in terms of context. As Sosa himself opposes contextualist theories of knowledge, this option is not available to Sosa. The following is structurally similar to a garbage chute case but with much higher stakes. Imagine someone working in a nuclear power plant. When they press a button, safety rods slide down a chute and stop the nuclear reactor overheating and killing thousands of people. Only very rarely do the rods snag and not get to the bottom of the chute. Suppose that the worker presses the button and forms the predictive belief that the rods will be at the bottom of the chute, and the disaster has been averted. Also, suppose that the rods do indeed make it to the bottom of the chute. Intuitively, it seems wrong to ascribe knowledge to the worker. Someone who claims knowledge because there is only a 0.001% chance that thousands will die appears to be following the wrong epistemic norms.

Following Sosa in supposing that contextualism is implausible, we have three options. (I) We can accept that garbage chute cases and high stakes cases are knowledge. The implausibility of this claim in the latter case and its incompatibility with otherwise plausible constraints on knowledge (for example sensitivity, safety, relevant alternatives, reliabilism) give good reason to reject this option. (II) We can accept that neither are knowledge. (III) We can deny that the cases are structurally similar. While it may be a failure of imagination on my part, I cannot see how this could be done. For those who do not have the intuition that the garbage chute case is knowledge, there is no issue here; (II) is the obvious choice. For those who are unconvinced, there are several considerations in favour of this option. First, (I) — admitting that the reactor case is knowledge — looks dire. Second, when one is pressed on whether one *really* knows that the garbage is at the bottom of the chute, many are inclined to admit that they do not; they were speaking loosely. The same consideration sits nicely with the reactor case. Third, if we find the initial claim that knowledge tracks truth plausible, then we are committed to rejecting the claim that cases like the garbage chute are knowledge. Consider both the garbage case and the nuclear reactor case. Once the garbage is put in the chute or the button is

pressed, the belief formation methods used neither interact with the progress of the garbage or rods, nor the things that cause garbage or rods to progress down the chute as they do (e.g. the roughness of the chute). There is no sense in which the belief formation of the agent tracks the facts of the matter except in a crude statistical way. Hence, we have reason to reject garbage cases as knowledge: beliefs in garbage chute cases (and structurally similar ones) are not knowledge because they do not track truth.

To reiterate, Sosa's counterexample is not safe, so it cannot give reason to accept safety over sensitivity. Also, it is structurally like cases that do not seem like knowledge, and the claim that it constitutes knowledge is incompatible with the claim that knowledge tracks truth, which is a key motivation for safety.

The Necessary Propositions Counterexample

Sosa believes that necessary propositions cause problems for sensitivity.[61] Consider some proposition that is necessarily true, e.g., Fermat's Last Theorem (FLT), and a subject who believes that proposition, such as Fermat. Fermat believes FLT; FLT is true; but is Fermat's belief sensitive? Sosa believes that there is no way of working out whether Fermat's belief is sensitive because we cannot evaluate the conditional 'if FLT were false, Fermat would not believe FLT via the same method'. This is because there is no closest world where FLT is false, as FLT is true at all possible worlds. This is not problematic for two reasons. First, there are many proposed ways that we can model counter-possible conditionals like the one above. Second, Sosa conflates the technical apparatus with which we are modelling language with the language. Subjunctive conditionals are claims in language that we model using a possible world semantics. The failure of the modelling tool to evaluate the conditional in some cases is not reason to believe that the conditional cannot be evaluated or is not well defined. As an analogy, consider a mathematician who wishes to model a physical system. She sets up an equation to do this and the equation accurately models the system. She then finds that for certain values the equation requires her to divide by zero. In this case she ought to consider this reason to think that her modelling method is flawed, not reason to think that the world exhibits strange behaviour at those values. In practice, this is what mathematicians do. Similarly, the failure of our methods to evaluate the conditional 'if FLT were false, Fermat would not believe FLT via the same method' tells us that our modelling methods are flawed, not that the conditional cannot be evaluated. Hence, necessarily true propositions do not cause problems for the sensitivity theorist. They do, however, cause problems for the safety theorist.

Consider Fermat, who believes FLT, which is necessarily true, but he does so because he believes all conjectures that are named after him. Presumably, a

[61] Ibid., 146.

belief formed based on vanity alone is not knowledge as it does not track truth. Despite this, the safety theory declares this belief to be knowledge. Fermat believes that FLT is true; it is true; and there is no close possible world at which Fermat believes FLT, but it is false. Note: this is a slightly different problem to the one Sosa claimed the sensitivity theorist faced. Sosa worried that the sensitivity conditional could not be evaluated in cases of necessary truths. The safety conditional, on the other hand, can be evaluated; it just gives the wrong results. The safety theorist must say that Fermat's belief is knowledge when it is not. To avoid such problems, the safety theorist must introduce some fourth condition to deal with beliefs about necessary truths. In the absence of such a fourth condition, this gives reason to support sensitivity over safety.

Conclusion

Given that I have motivated sensitivity as a necessary condition for knowledge and shown that it deals with Sosa's counterexamples better than safety, I conclude that Sosa does not give reason to adopt safety over sensitivity. In other words, the sensitivity theorist can give adequate responses to all of Sosa's counterexamples. While there are other attacks on sensitivity theory that need to be dealt with, I have shown that many counterexamples thought to rule out sensitivity do not rule it out. There are other purported counterexamples to sensitivity, which I aim to deal with in future work, but I take the considerations above to conclusively show that Sosa's purported counterexamples pose no serious problems for the sensitivity theorist.

Bibliography

Nozick, Robert (1981). "Knowledge and Skepticism". In Sosa, Ernest; Kim Jaegwon (ed.) *Epistemology: an anthology*. 2000. Oxford: Blackwell, 79-101.

Sosa, Ernest. "How to Defeat Opposition to Moore". *Philosophical Perspectives*, no. 13 (1999): 141-153.