where sensitivity don't work

Mark Anthony Dacela
De La Salle University, Manila

Abstract

Robert Nozick (1981, 172) offers the following analysis of knowledge (where S stands for subject and *p* for proposition):

> **D1** S knows that p $=_{df}$ (1) S believes *p*, (2) *p* is true, (3) if p weren't true, S wouldn't believe that p (variation condition), and (4) If p were true, S would believe it (adherence condition).

Jointly, Nozick refers to conditions 3 and 4 as the *sensitivity condition*: for they require that the belief be *sensitive* to the truth-value of the proposition—such that if the proposition were false, the subject would not have believed it, and if the proposition remains true in a slightly different situation, the subject would have still believed it. In other words, they ask us to consider the status of the belief in close possible situations (those that obtain in close possible worlds); specifically, in situations that would obtain if the proposition is false, and in those in which it remains true. Condition 3 specifies how belief should vary with the truth of what is believed, while condition 4 specifies how belief shouldn't vary when the truth of the belief does not vary. I will discuss some notable problem cases for Nozick's analysis and then look at why the sensitivity condition he proposes fails in these cases.

Keywords
sensitivity, modal epistemic condition, nozick, epistemology, justification

Introduction

Robert Nozick (1981, 172) offers the following analysis of knowledge (where S stands for subject and *p* for proposition):

> **D1**   S knows that p =_{df} (1) S believes *p*, (2) *p* is true, (3) if p weren't true, S wouldn't believe that p (variation condition), and (4) If p were true, S would believe it (adherence condition).

Jointly, Nozick refers to conditions 3 and 4 as the *sensitivity condition*: for they require that the belief be *sensitive* to the truth-value of the proposition—such that if the proposition were false, the subject would not have believed it, and if the proposition remains true in a slightly different situation, the subject would have still believed it. In other words, they ask us to consider the status of the belief in close possible situations (those that obtain in close possible worlds); specifically, in situations that would obtain if the proposition is false, and in those in which it remains true. Condition 3 specifies how belief should vary with the truth of what is believed, while condition 4 specifies how belief shouldn't vary when the truth of the belief does not vary. I will discuss some notable problem cases for Nozick's analysis and then look at why the sensitivity condition he proposes fails in these cases.

How Does Sensitivity Work

Nozick (1981, 173) appeals to the possible-worlds accounts of subjunctive conditionals to explain how sensitivity works (*p* stands for the antecedent and *q* for the consequent):

> The subjunctive is true when (roughly) in all those worlds in which *p* holds true that are closest to the actual world, *q* also is true (Examine those worlds in which *p* holds true closest to the actual world, and see if *q* holds in all these). Whether or not *q* is true in *p* worlds that are still farther away from the actual world is irrelevant to the truth of the subjunctive.

From the above passage, we deduce the following steps in checking if the given case satisfies the sensitivity requirement:

> Step 1: Check for *close possible worlds* (worlds similar to the actual) in which the antecedent holds (call these *p* worlds).
> Step 2: Check if in those worlds the *consequent* also holds (call these *q* worlds).

Nozick also tells us to shy away from *far-away* worlds:

> **D2** The subjunctive 'if it were *p* then it would be *q*' is true = Df. (1) All *p*-worlds close to the actual world are *q* worlds.

Conversely, if there's at least one *p-world* that is not a *q*-world the subjunctive is false:

> **D3** The subjunctive 'if it were *p* then it would be *q*' is false = Df. (1) At least one *p*-world close to the actual world is not a *q*-world.

Now consider again Nozick's conditions (recall: *p* stands for proposition and *S* stands for subject):

> (3) If *p* weren't true, S wouldn't believe that *p*.
> (4) If *p* were true, S would believe *p*.

Given D2, we see that (3) asks us to consider those worlds close to the actual in which the given proposition is false (worlds in which the antecedent holds). We are to check if in these worlds the subject does not believe the proposition. If so, then, the subjunctive '*if p weren't true, S wouldn't believe that p*' is true. D3 tells us that the belief is not sensitive if in at least one of those *relevant worlds* the subject falsely believes the proposition.

Similarly, (4) asks us to consider close possible worlds in which the given proposition is true (again, worlds in which the antecedent holds). We are to check if in these worlds the subject still believes the proposition. If this is so, then, the subjunctive '*if p were true, S would believe p*' is true. Given D3, we can deduce

that the belief is not sensitive if in at least one of these *relevant worlds* the subject does not believe the proposition, even if it is true in these worlds.

Problems for Sensitivity

Sensitivity is proposed as necessary condition for knowledge. In other words, sensitivity theorists hold that knowledge requires sensitive beliefs. Proponents of these two conditions drop the justification condition but preserve the belief and truth conditions of the traditional definition of knowledge, or the justified true belief account:

> D4   S knows that p = $_{Df.}$ (1) S believes *p*, (2) *p* is true, and (3) S is epistemically justified in believing *p*.

Here are some problem cases for Nozick's analysis.

Garbage Chute (Vogel and Sosa)

Vogel (1987), and later Sosa (1999), present this case as a counter-example to sensitivity:

> **Garbage Chute.** Sam throws a trash bag down the garbage chute of her condo. Some moments later she believes that *the trash bag is in the basement* (p).

In *Garbage Chute*, Sam's belief that *the trash bag is in the basement* (p) amounts to knowledge (intuitively) yet it fails to be sensitive. Imagine a world where *p* is false. Suppose that the thrash bag got stuck somewhere in the chute. In this imagined situation, Sam would still believe *p*. The variation condition is not satisfied: *If it were false that the thrash bag is in the basement, Sam would still believe that it is.* The belief is not sensitive, and we get the wrong result: Sam does not know *p*.

The problem is this: while the proposition would have easily been false, the subject would not have easily varied her belief. The intuition that this is an instance of knowledge seems motivated by Sam's strong justification for believing *p*. This was not made explicit since *sensitivity* is not concerned with

justification. But I do think we need to clarify this point. We cannot just assume that what we have here is an instance of knowledge at the expense of Nozick's theory. To take this objection seriously, we need to understand why we have the intuition that in this case *Sam knows*.

What could Sam's justification be for believing that *p*? It could be that throwing a trash bag in the chute is something that she has done several times before, and that every time she does this, she would later see the bag in the basement (only this time she didn't). This slightly modified version of *Garbage Chute* still works as a counterexample. Even with a strong justification, in a close possible world where the bag got stuck somewhere in the chute, Sam would still *justifiably* believe *p*.

This problem extends to a family of cases where the subject has good reasons for believing a proposition that could have easily been false. Think of a case in which someone believes that her friend had read the email she just sent; or someone believing she returned a book she borrowed. Consider as well an instance in which someone thinks she already texted a friend, and another in which someone believes she did not have money left in her wallet. These are situations in which we could easily be justified in believing something of which we are not quite sure. Except perhaps in times when we go out of our way to double-check if our belief is true (e.g. you can ask your friend if she got your email or text and you can go to the library and check if you did return the book), or when circumstances force us to do just that (you are really hungry so you have to check your wallet if you have money to buy food). But usually, we just assume that our beliefs in these situations are true, and often they are. However, they could have easily been false. Your email could have easily gone straight to the spam folder; you could have easily returned the wrong book; and you could have easily forgotten about the money (inside your wallet) that you saved for emergencies.

In the examples cited above, it is possible that you have a strong but non-conclusive evidence for your belief. Sensitivity seems to fail in cases of this sort. If we suppose that in the situations cited, you have good but non-conclusive reason to hold belief *p* (where *belief p* refers to the belief in question), then given sensitivity, you don't know *p* in all instances. Had your email gone straight to the spam folder, you would have still believed that your friend read it; had you returned the wrong book, you would have still believed that you returned the right one; If you had money left in your pocket, you would have still believed

that you don't. In all instances, your belief is not sensitive. Thus, sensitivity seems to fail in cases where the subject has strong but non-conclusive evidence for *p*.

Life on Mars (Gellman)

Now consider Jerome Gellman's (2004) Gettiered case:

> **Mars**. Suppose that it follows from a law of nature that S believes that there is life on Mars if and only if S is in brain state B10 (a_1). So there is this lawful correlation between this mental state of S and this state of S's brain. Sam knows nothing of this correlation. Sam believes that *there is life on Mars* (m). Sam has solid evidence (e_1) for his belief that *there is life on Mars* (m). It is false that there is life on Mars (not-*m*). Sam believes that *he is not in brain-state B10* (not-*b*). S has solid evidence (e_2) that he is not in brain state B10. S has acquired a habit of inferring disjunctions with I am in brain state B10 as the second disjunct, when S feels warranted in believing the first disjunct. Therefore, S then infers from that *there is life on Mars* (m) that: There is life on Mars or I am in brain state B10 (p).

The following equations then obtain:

1. Sam believes that *there is life on Mars* (Bs*m*).
2. If Sam is in brain-state B10 then Sam believes that there is life on mars (a_1).
3. S has acquired a habit of inferring disjunctions with I am in brain state B10 as the second disjunct, when S feels warranted in believing the first disjunct (a_2)
4. Sam does not know L_1 (-Ksa_1).
5. Sam's evidence supports his belief that *there is life on Mars* ($e_1 \rightarrow m$).
6. Sam's evidence supports his belief that *he is not in brain state B10* ($e_2 \rightarrow$ not-*b*).
7. Given a_2, Sam believes that his belief that *there is life on Mars* implies the belief that *there is life on Mars on I am in brain state B10* {Bs [m $\rightarrow$ (m ∨ b)]}.
8. Sam believes that *there is life on Mars or I am in brain state B10* (Bs*p*).
9. There is no life on Mars (not ~m).

In *Mars*, Sam believes (among other things) that *either there is life on*

mars or he is in brain state B10 (p). Notice that this belief is in the form of a disjunctive. I should note two things about this type of proposition: (1) A disjunction is true if in case both or one of the disjuncts is true and (2) It is false, if both disjuncts are false. Since Sam believes that *there is life on mars* (m), and given that 'S believes that there is life on Mars iff S is in brain state B10' (a_1), it follows that the second disjunct of *p* is true (Sam is in brain state B10). So, Sam's belief that *p* is true. Gellman (2004, 281) thus claims that in *Mars*, the belief and truth condition are satisfied.

Gellman (2004, 281) further claims that Sam's belief satisfies both the variation and adherence condition of sensitivity. Here are the reasons. First, the worlds closest to the actual world in which *p* is false, both disjuncts would be false (since a disjunctive proposition is false iff both disjuncts are false). Meaning, there is no life on Mars and Sam is not in brain state B10. However, the laws of nature in the actual world also hold in the closest world and so a_1 (S believes that there is life on Mars iff S is in brain state B10) will hold. We should also consider as fixed Sam's evidence for his belief that *he is not in brain state B10* (not-*b*). So in this imagined worlds the following obtains:

1. S believes that there is life on Mars iff S in brain state B10 (a_1).
2. S has acquired a habit of inferring disjunctions with I am in brain state B10 as the second disjunct, when S feels warranted in believing the first disjunct (a_2).
3. There is no life on mars (~*m*).
4. S is not in brain state B10 (not-*b*).
5. Not-*p* (This follows from 2 and 3).
6. $e_2 \rightarrow$ not-*b*

If *not-b* (4), then, given a_1, Sam does not believe that *there is life on mars* (m). And given his evidence (e_2) for the belief that *he is not in brain state B10* (not-*b*), Sam will continue to disbelieve the second disjunct of *p* (that he is in brain state B10). Also, we have no reason to suppose that in this imagined situation, where Sam does not believe both *m* (there is life on Mars) and *b* (that he is in brain state B10), Sam would believe *p* (the disjunction of *m* and *b*). This also follows from (2). Thus, in the closest world where *p* is false, Sam does not believe it. The variation condition is satisfied. Sam's belief is sensitive to the *falsity* of *p*.

The second reason is this, in the world closest to the actual world in which

p is true, either m or b, or both are true. So the following obtains:

1. S believes that there is life on Mars iff S in brain state B10 (a_1).
2. Either there is life on mars (m) or S is in brain state B10 (b).
3. Sam has solid evidence for his belief that there is life on mars ($e_1 \rightarrow m$)
4. S has acquired a habit of inferring disjunctions with I am in brain state B10 as the second disjunct, when S feels warranted in believing the first disjunct (a_2)

If *m* (there is life on Mars) is true, given Sam's solid evidence (e_1) for *m* and his habit of inferring disjunctions with I am in brain state B10 as the second disjunct when he feels warranted in believing the first (a_2), Sam would still believe *p*. If *b* (Sam is brain state B10) is true, then, given the correlation between the belief that there is life on Mars and mental state B10 (a_1), S would believe that there is life on Mars (*m*). And again, given a_2, Sam will form and accept *p*. In the close possible worlds then in which *p* is true (close m-worlds or b-worlds), Sam would still believe *p*. The same holds in close worlds in which *m* and *p* are both true. The adherence condition is satisfied. Sam's belief is sensitive to the truth of *p*.

So in *Mars*, Sam's belief is sensitive. Gellman (2004, 282) however claims that Sam does not know *p*, given that the case is similar in form with Gettier's second case (Gettier, 1963):

> **Ford.** Smith remembers that Jones has at all times in the past owned a car, and always a Ford. Also, Jones just offered Smith a ride while driving a Ford. Smith thus believes that Jones owns a Ford (p). But unknown to Smith Jones does not really own a Ford. Jones at present is driving a rented car. Smith has another friend, Brown, whose whereabouts he does not know. Smith then realizes that his belief that Jones owns a Ford entails the following: either Jones owns a Ford, or Brown is in Boston (s), either Jones owns a Ford, or Brown is in Barcelona (q), either Jones owns a Ford, or Brown is in Brest-Litvosk (r), and so believes all three. Unknown to Smith, Brown is in Barcelona.

In *Ford*, Smith's deduces his belief that *either Jones owns a Ford or Brown is in Barcelona* (q) from the belief that *Jones owns a Ford* (p). The second disjunct of *q* (Brown is in Barcelona) makes it true. But Smith does not really know that

Brown is in Barcelona. Also, *p* is false. Jones does not own a Ford. Smith luckily deduces a true belief (belief q) from a false one (belief p). It is only by sheer coincidence that Smith's belief *q* is true. So, Smith does not know *q*.

Similarly, Gellman (2004, 282) claims, in *Mars*, Sam deduces his belief that *either there is life on Mars or I am in brain state B10* (p) from his belief that *there is life on Mars* (m). The second disjunct of *p* (I am in brain state B10) makes it true. But Sam does not know that he is in brain state B10 (b). And, *m* is false. There is no life on Mars. Sam deduces a true belief (belief p) from a false one (belief m). It is only by sheer coincidence then that Sam's belief *p* is true. So while Sam's belief that *p* is sensitive, it is not an instance of knowledge. Sam does not know *p*.

Mars demonstrates that sensitivity is not sufficient for knowledge. But it also reveals a more serious problem. In *Mars* we are required by sensitivity to check a world where the subject has a different brain state. Note that in the actual world, Sam believes that there is life on Mars. Given that there is a correlation between this mental state and brain state B10, we have to suppose that Sam is in brain state B10. In the closest world where Sam's belief that *either there is life on Mars or I am in brain state B10* (p) is false, Sam is not in brain state B10. Now you can say that a difference of this sort is not really significant, so you might think that it is not far-fetched to say that the world where Sam has a different brain state is epistemically relevant, at least in the sense Nozick thinks close possible worlds are. But consider another case that has a similar pattern:

> **Dog Man.** Smith has strong evidence for the following beliefs: I am not a dog (r) and I am drinking coffee (q). From *r* and *q* Smith deduces that *either I am not a dog or I am not drinking coffee* (p).

To check if the subjunctive *"if p were false then S would believe that p"* holds in this case, we have to imagine a world where both disjuncts of *p* are false. So that means checking the closest world where *Smith is a dog drinking coffee*. Note that in this particular case there is a good reason to think that Smith knows *p*: Smith validly deduces *p* from true beliefs he holds on the basis of strong evidences. It seems in this case (and in most cases that involve disjunctive beliefs which negation would hold in strange worlds similar to a world where Smith is a dog drinking coffee) the sensitivity requirement is not necessary.

Mad, Bad, and Dangerous (Briggs and Nolan)

We now turn to three cases highlighted by Rachel Briggs and Daniel Nolan (2012). They say that in these examples, the subjects have an epistemically relevant disposition to believe the relevant proposition if it is true, and not to believe it if it is false. However, in these examples, the subjects could easily lose such disposition, and this change would have influenced their attitude toward the proposition. I think this diagnosis hits the mark, but I have more to say about these cases.

Consider this first:

> *Mad*. William has read stories in the style of *H.P Lovecraft* about nightmarish creature called gz'hurs. He believes that *there are no gz'hurs waiting to burst through the fabric space and time to seize control of his mind* (p). As a matter of fact, he is right. However, the metaphysics of gz'hurs dictates that if there were such creatures, they would already be animating his empty husk, while leaving all of his beliefs intact in order to conceal the transformation from his loved ones... until it was too late (Briggs and Nolan 2012, 314 – 315).

In *Mad*, if it were that William's belief that *there are not gz'hurs waiting to burst through the fabric space and time to seize control of his mind* (p) is false, he would still believe that this is the case; since in a close world where there are gz'hurs that are controlling William's mind, all his beliefs, including *p*, are left intact. So William's belief is not sensitive, yet it appears that he knows *p*.

Riggs and Nolan did not explain why William *knows* in this case. But just like in *Garbage Chute*, we can motivate this intuition by appealing to William's strong but nonconclusive evidence for *p* (and perhaps the absence of any evidence for the presence of gz'hurs). Inserting this modification will not change the result: if it were that *p* is false, William would still believe it. *Mad* also has similar elements with most skeptical hypotheses. For instance, in the *brain in a vat* scenario, if you were a brain in a vat you would not believe that you are. Similarly, in *Mad*, if gz'hurs are controlling your mind, you still won't believe that they are. Given sensitivity then, we cannot know that *gz'hurs are not controlling our mind*, just as we cannot know that *we are not BIVs*. Yet it seems intuitive to

say that in the case of the make believe *gz'hurs*, this is the wrong result.

Now you might say that *scenarios* like these do not really count, since they involve odd creatures existing in strange worlds. But consider this case:

> **Mom.** Hours before knowing the result of his exam, Tom already decided that whatever the outcome is, he will tell his mom that he passed. Tom makes the cut by just one point. So when his mom asked how he did in his exam, he said, truthfully, that he passed. His mom shortly after formed the belief that *Tom passed the exam* (p).

In *Mom*, Tom's mom knows that *he passed the exam* (p). But if Tom did not make the cut—say he missed one item—his mom would still believe that he did. So belief *p* is not sensitive. Now, there are no weird creatures here. Just your average student making sure his mom is happy, even if it means lying to her. Belief *p*, arguably, cannot be sensitive, given that Tom has already decided to tell his mom that he passed, regardless of the actual result of his exam. So similar to *Mad*, we have a knowledge case here that involves a belief that cannot be sensitive.

Next, consider this case:

> **BAD**. Annabella Milbank believes that *she is decent and upstanding* (p). And indeed she is; her belief flows from her decent and upstanding character. She could, however, be easily corrupted. If corrupted, she would become so desensitized to decency that she would continue to believe she was decent and upstanding (albeit unfairly maligned by puritanical society). So Annabella's belief is not sensitive to the proposition that she is decent and upstanding. Nonetheless, Annabella appears to know that she is decent and upstanding (Briggs and Nolan 2012, 315).

In *Bad*, Annabella truly believes that *she is decent and upstanding* (p). However, there is a close possible world where she is not decent and upstanding, yet still believes that she is. It is a world where she is desensitized to decency, so much so that she continues to believe that she is decent and upstanding even if she is not. So belief *p* is not sensitive: if it were that *p* is false, S would still believe *p*. Annabella does not know *p*, yet it appears that she does.

Why must we think that Annabella knows *p*? We can assume that she has strong reasons to believe that *she is decent and upstanding* (p) given that she actually is. In fact, Briggs and Nolan (2012, 315) reason that, "her belief flows from her decent and upstanding character." Her character would greatly determine her actions, and this would influence how others act toward her. Perhaps then people compliment her for her character. Or maybe she's just very conscious of her words and actions (and let's say that all of these exemplify her character). Either way, strong evidence supports Annabella's true belief. That motivates the intuition that *she knows p*. This does not change the fact that Annabella could be easily corrupted. In modal terms, this means that there's a close world where she *falsely* believes *p*. But why would she not, just as easily, let go of this belief? It says in the example that if Annabella gets corrupted, she would be *desensitized to decency*. And this would make her believe that she is *decent*, even if she's not.

Finally, consider this example:

> **DANGEROUS**. Adolf believes, having studied his symptoms and consulting with medical experts, that *he has a rare brain condition that is fatal before the age of 5 years in 99.99% of cases* (p). In fact, he is right; he's one of the one in 10,000 who have the disease but survive. However, in most nearby possible worlds here he has the disease, he dies before the age of 5 years – long before he entertains the proposition that he has the disease. Therefore, it is not true that if Adolf had had the disease, he would have believed he did. (He might well have died at 5 years.) So Adolf's belief is not safe. Nonetheless, Adolf knows he has the disease (Briggs and Nolan 2012, 315).

In *Dangerous*, Adolf truly believes that *he has a rare brain condition that is fatal before the age of 5 years in 99.99% of cases* (p). However, there's a close possible world where he dies long before he forms belief *p*. In this world, *p* remains true, but Adolf no longer believes it. This case violates the adherence condition of sensitivity: *if p were true, S would believe it*. The belief is not sensitive to the truth of the proposition. So, Adolf does not know *p*. But this is the wrong result.

How do we motivate the intuition that *Adolf knows*? Again, imagine Adolf having strong reasons to support his belief. For instance, suppose that his doctor reliably informed him that he has a disease that is fatal before the age of 5 in

99.99% of cases. Adolf's true belief would then be grounded on solid evidence. Yet it remains to be true that if *had had the disease* (in the situation describe above), *he would not believe he did*. He would die clueless before he hits the age of 5. Reimagining this case does not change the result. If sensitivity is necessary for knowledge, then Adolf does not know p.

Let's make explicit the difference between the actual world (@) and the close relevant world (w_1) at play in this case. First, in @, Adolf has the disease but luckily he's not dead before he turned 5. In w_1, he has the same disease but he dies before the age of 5. There is no way he would have formed belief p in w_1, since, presumably, he would not have entertained the proposition that *he has a brain disease* in such a juvenile age. To understand why this possible world is close, you have to consider that 99.99% of the time, people with whatever Adolf has, die before the age of 5. Adolf had 0.01% survival chance. In modal terms, this means that in close worlds, Adolf died before he turned 5. He got lucky. In the actual world at least, he's not as lucky in the nearby worlds.

This variation of luck would have significantly changed Adolf's epistemic state. If he were dead before he turned 5, he would not have accessed the sort of evidence that would warrant the belief that *he has a rare brain condition that is fatal before the age of 5 years in 99.99% of cases* (p).

Concluding Notes

The cases I discussed share the same profile: S has strong justification for believing p and p is true. Cases of these (call them A-type cases) sorts have a subject who has a strong justification for believing a true proposition, a *strongly justified belief*. Justification is strong if the subject's evidence is almost conclusive. In A-Type cases, the subject's belief is *internally justified*, insofar as the subject has access to the evidence that supports her belief. And as far as there are no (actual) defeaters in the description of the case, A-Type beliefs are *undefeated*. Thus, in the actual or referent world, the belief is both internally justified and undefeated, such that no contrary evidence is given or accessible to the subject. Proponents of these counterexamples count as close worlds those in which the subject holds the same evidence, so the actual belief-characteristics are extended to these worlds. Worlds in which the subject does not believe the proposition are excluded from the set of relevant worlds, while worlds in which

"*the proposition is true and the subject believes it*" are counted as close. A-Type beliefs will turn out to be insensitive, whenever the subject remains internally justified in believing the proposition in question.

Works Cited

Briggs R. and D. Nolan. 2012. Mad, Bad, and Dangerous Knowledge. *Analysis* 72: 314-316.
Gellman, Jerome. 2004. A New Gettier Type of Refutation of Nozick's Analysis of Knowledge. *Principia* 8: 279 -283.
Gettier, Edmund. 1963. Is Justified True Belief Knowledge?. *Analysis* 23: 121-123.
Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.
Sosa, Ernest. 1999. How to defeat Opposition to Moore. *Philosophical Perspectives* 13: 141 – 153.
Vogel, J. 1987. Tracking, Closure, and Inductive Knowledge. In S. Luper-Foy (ed.), *The Possibility of Knowledge: Nozick and his Critics*. Totawa, NJ: Rowman & Littlefield.