9-2012

False Negatives

Steven Luper
Trinity University, sluper@trinity.edu

Repository Citation
Luper, S. (2012). False negatives. In K. Becker & T. Black (EDs.), *The Sensitivity Principle in Epistemology* (pp.207-226). New York: Cambridge University Press. doi: 10.1017/CBO9780511783630.016

FALSE NEGATIVES

In *Philosophical Explanations*, Robert Nozick suggested that knowing that some proposition, *p*, is true is a matter of being "sensitive" to *p*'s truth-value. It requires that one's belief state concerning *p* vary appropriately with the truth-value of *p* as the latter shifts in relevant possible worlds. Nozick fleshed out this sketchy view with a specific analysis of what sensitivity entails. Famously, he drew upon this analysis in order to explain how commonsense knowledge claims, such as my claim to know I have hands, are true, even though we do not know that skeptical hypotheses are false. His explanation hinged on rejecting the principle that knowledge is closed under (known) entailment. In this essay I will criticize Nozick's view of knowledge as "sensitivity" to truth-value. In doing so, I mean to undermine his case against the closure principle and against the claim that we do not know that skeptical hypotheses are false.

I begin with a review of Nozick's notion of sensitivity and his analysis thereof.

Sensitivity

According to Nozick (1981, p. 176 and p. 208), if a subject S is to know that some proposition, *p*, is true, it is necessary for S's belief concerning whether or not *p* is true to be "sensitive" to the "truth-value of *p*;" that is, S's belief state concerning *p* must vary consistently with the truth or falsity of *p* in salient possible situations, so that, for example, in all of the salient worlds in which *p* holds, S believes that *p* holds. Nor is it enough for S to be sensitive to only one of *p*'s truth-values; knowing *p* requires sensitivity to *p*'s truth, but it also requires sensitivity to *p*'s falsity. Let us assume that, for S to know *p*, there really must be some sort of variation between *p*'s truth-value and something related to S's belief state vis-à-vis *p*. Call this required variation, whatever it turns out to be, *epistemic sensitivity*. Let us also coin terms for the two specific sorts of sensitivity to which Nozick referred: the requisite sensitivity to *p*'s truth we can call *truth sensitivity,* and the requisite sensitivity to *p*'s falsity we can call *falsity sensitivity*.

Nozick said that, for subject S to know *p*, the requisite sensitivity vis-à-vis *p*'s truth is roughly that in the close possible worlds in which *p* is true, S has a matching belief; that is,

$p \rightarrow$ S believes *p*.

(Here the arrow symbolizes the subjunctive conditional.) Following current practice, we can call this the *adherence condition*; or, when more precision is needed, we may call it the *truth* adherence condition. To defend his view, Nozick (1981, p. 175) drew on "the case of the person in the tank who is brought to believe, by direct electrical and chemical stimulation of his brain, that he is in the tank and is being brought to believe things in this way." Nozick said that this Vatted person does not know he is Vatted, since "the operators of the tank could have produced any belief," including the false belief that he is not Vatted. So he will not meet the adherence condition.

Given that truth sensitivity is constituted by truth adherence, one might have expected that falsity sensitivity would amount to falsity adherence. In that case, Nozick's account of epistemic sensitivity would look like this:

Subject S is epistemically sensitive to *p* iff:

$p \rightarrow$ S believes *p* (truth adherence), and

$\sim p \rightarrow$ S believes $\sim p$ (falsity adherence).

Call this the *adherence account*. However, this is not the position that Nozick defended. The falsity adherence condition requires that S have a *correct belief* in the close $\sim p$ worlds. By contrast, Nozick required something weaker, namely that S *fail to have an incorrect belief* in the close ~*p* worlds. He said we should analyze falsity sensitivity in the following way:

~*p* → ~(S believes *p*) (avoidance; more precisely, false positives avoidance).

(The above condition, which I here label the avoidance condition, has come to be called the sensitivity condition, but the latter term is misleading since both of Nozick's main conditions are implicated in his account of epistemic sensitivity, and the most important requirement for epistemic sensitivity, in his view, is that belief should follow truth.) Initially, he defended the avoidance condition on the grounds that it explains why knowledge fails in Gettier cases. Nozick added to his defense in a footnote, where he claimed that his condition is superior to an alternative, namely the falsity adherence condition. The defense he offered is this: If we analyze truth sensitivity as adherence to truth (as he wished to do), and falsity sensitivity as adherence to falsity, as the adherence account does, we imply that, to know *p*, our relation to ~*p* must be as strong as our relation to *p*. Yet "knowledge that *p* involves a stronger relation to *p* than to not-*p*" (1981, p. 682 n. 15). Since the avoidance condition is *weaker* than the adherence condition, the former constitutes a superior account of falsity sensitivity. Accordingly, Nozick opted for the following (preliminary) account of epistemic sensitivity:

S is epistemically sensitive to *p* iff:

p → S believes *p* (truth adherence), and

~*p* → ~(S believes *p*) (avoidance).

Call this the *belief version of the tracking account*.

What should we make of Nozick's account? A critic might well resist the case Nozick offered for preferring the tracking account over the adherence account. If, as Nozick said, knowing *p* requires sensitivity to both of *p*'s truth-values, it is not clear at all that we should accept his epistemic asymmetry thesis—his claim that knowing *p* involves a stronger relation to *p* than to ~*p*. Why isn't the required sensitivity the same for both truth-values, as the adherence account implies? However, this is not the sort of criticism I intend to offer. Instead, I will suggest that Nozick's tracking account and the adherence account are both fatally flawed. I will also suggest something stronger, namely, that Nozick's basic idea of knowledge is mistaken: knowing that a proposition is true does not require sensitivity to its truth and falsity.

Before I make my case, I will briefly discuss how Nozick's account might best be formulated.

Methods and Qualifications

The belief version of the tracking account *suppresses* any mention of the methods by which beliefs are formed (or sustained). However, as Nozick noted, by explicitly *revealing* methods we can allow for bad methods that weren't, but might have been, used. A bad method is a method whose use will not position us to know things. We can distinguish two possibilities: the bad method is one which S might (but won't) use if *p* holds, or the bad method is one which S might (but won't) use if ~*p* holds. Here are illustrations of these possibilities (based on Nozick's Grandmother case on p. 179):

Child 1: I believe my child is alive because I see her playing tennis, but also because she told me she is immortal and cannot die.

Child 2: I believe my child is alive because I see her playing tennis, but if she weren't alive I'd believe she was through wishful thinking.

In both of these cases I know my child is alive, despite the fact that I might have relied on a bogus method of belief formation. If I were relying on wishful thinking, I would not know my child is alive. But I am not relying on wishful thinking. I see her.

Nozick responded to this line of criticism by revising his account so as to reveal methods. His revisions created further difficulties, discussed in Luper 1984, p. 29, where I suggest the following improved version of the tracking account:

S knows *p* only if S believes *p* via a method M such that:

p & S applies M → S believes *p* via M (truth adherence), and

~*p* →~(S believes *p* via M) (avoidance).

Take the Child cases: if my child were not alive, I would not see her playing tennis (even though wishful thinking, or my belief that she is immortal, would prompt me to believe she is still alive).

Let me now suggest another improvement. The formulation just mentioned *attaches belief*. Like the belief version of the tracking account, this generates difficulties. It would be better to *detach* belief, as follows:

S knows *p* only if S believes *p* via a method M such that:

p and S applies M → M indicates *p* (truth adherence), and

~*p* →~(M indicates *p*) (avoidance).

The reason we should detach belief is that Nozick's avoidance condition does not function as intended unless we do so. For example, it is disarmed in the following case:

Life: I believe I am not dead because my cat is in a hat.

If I were dead, I couldn't believe anything, so I could not possibly fail to meet the consequent of the original avoidance condition. Yet my being alive has nothing to do with cats or hats. If I were dead, my cat might not be in a hat. I do not meet the restated avoidance condition.

Once we reveal methods of belief formation (or sustainment), we bear the burden of spelling out what methods are, what it is for them to indicate that a proposition is true or false, how they are related to the beliefs they facilitate, and whether they must pass through some sort of 'internalism' filter. I will not take up that burden (for discussion, see Luper 1984, Alfano 2009, and Becker 2009a, 2009b). In what follows I will assume that we can describe methods on the model of rules that take the general form, if *q* holds, then *p* is true. Such a method M indicates *p* if and only if *q* holds. Here *q* might be a disjunction; if so, M indicates *p* if and only if the one disjunct or the other holds. I will also assume that a means by which *p* is verified (or falsified) qualifies as a method for believing *p* (or ~*p*).

One other complication arises: as Nozick noted (albeit in discussing the preliminary formulation of his account), his truth adherence condition is too easily met by methods that would indicate that *p* is both true and false if *p* were true. We might ensure consistency by adding the following condition:

[M indicates *p* →~(M indicates ~*p*)] and [M indicates ~*p* →~(M indicates *p*)] (the consistency requirement),

or by building consistency into Nozick's condition, as follows:

p and S applies M → M indicates *p* and ~(M indicates ~*p*)

Given the worries that led us to reformulate the tracking account, it seems best to restate the adherence account as well. First, we can reveal methods and detach belief:

p and S applies M → M indicates *p* (truth adherence)

~*p* and S applies M → M indicates ~*p* (falsity adherence).

Second, we will need to eliminate the use of contradictory methods. This is especially pressing in connection with the adherence account since, as stated, it allows us to know *p* by relying on a method that always indicates that *p* is both true and false. To eliminate inconsistent methods, we can add the consistency requirement. Alternatively, we can build consistency in, as follows:

> *p* and S applies M → M indicates *p* and ~(M indicates ~*p*) (truth adherence)
>
> ~*p* and S applies M → M indicates ~*p* and ~(M indicates *p*) (falsity adherence).

(Note that, as long as we avoid contradictory methods, the falsity adherence condition is stronger than the false positives avoidance condition.)

Undetected Instances

I said that Nozick's account and the adherence account share a flaw. That flaw will become apparent if we consider the following method of belief formation:

> *Gappy Fever Detection* (GFD): if someone is alive and her temperature is over 101°F then she has a fever; if she is alive and her temperature is less than 95°F then she does not have a fever.

Suppose we are trying to establish that Frieda has a fever using this method. If Frieda is alive and her temperature is over 101, the GFD method will indicate that she has a fever. However, people are feverish when their temperature is not over 101. For example, the temperature of 101, or 100, and everything in between, results in fever (assuming a normal temperature of 98.6). GFD will not indicate that such persons have fevers. It has gaps, in that there are cases of fever which it misses. Hence there are close worlds in which Frieda has a fever, yet, using GFD, we are not in a position to determine that she has a fever. So indication need not adhere to truth. Nor need indication adhere to falsity: If Frieda's temperature is 95, she does not have a fever, but GFD will not indicate that she does not have a fever.

None of this jeopardizes our knowing that Frieda is feverish using GFD when her temperature is over 101. Clearly we can know she is feverish using GFD when her temperature is over 101 even though we cannot when her temperature is 101 and even though we cannot know she is not feverish using GFD when her temperature is 95.

In order to describe what is distinctive about the GFD example, let us coin some terminology. Relative to a proposition *p* and a method M that is applied to *p*, a *positive* is an instance of M indicating that *p* is true, while a *negative* is an instance of M indicating that *p* is false. A *false* positive is a case in which M indicates *p* yet *p* is false, while a false negative is a case in which M indicates ~*p* yet *p* is true. In a *true* positive (negative), M indicates *p* (~*p*) and p is true (false). By contrast, *an undetected instance of p* is a case in which *p* is true but M fails to indicate *p*; similarly, *an undetected counterinstance of p* is a case in which *p* is false but M fails to indicate ~*p*. The method discussed above, GFD, fails to detect some instances of *p* as well as some counterinstances. People have fever when their temperature is 100; this is an instance which GFD misses. Similarly, people do not have fever when their temperature is 95; this is a counterinstance which is missed by GFD. But while GFD misses some instances and counterinstances, it does not give us false positives or negatives. Hence what the GFD example reveals is this: we can know things using methods that miss instances or counterinstances or both.

Nothing said here hinges on some special feature of the GFD method. I can know that an item on the clothes rack is blue by seeing that it exactly matches a particular blue color swatch, but if the item I am examining were not the same shade of blue as the swatch, this method would not tell me that it is blue. I can determine that an item in my salad bowl is a vegetable by establishing that it is a carrot, but this method would not work if applied to broccoli.

I conclude that knowledge does not entail adherence to truth-value. The truth adherence condition, which is part of the tracking account and the adherence account, is false. The falsity adherence condition, which is part of the latter, is also false. (We should also reject the view at which we arrive if we contrapose the conditions that constitute the adherence account, namely:

S is epistemically sensitive to *p* iff:

~(M indicates *p*) → ~*p*, and

~(M indicates ~*p*) → *p*.

This view also rules out undetected instances and counterinstances.)

Although examples like the ones I have offered strongly suggest that adherence conditions are unacceptable, the matter may well prove to be controversial. Certainly there are theorists who say things that are inconsistent with the position I have taken. Examples include Duncan Pritchard (2002), who suggests coupling (a belief version of) truth adherence with the safety condition, discussed below. Another example is Nozick himself, and everyone else who reasons as he did about his Vatted brain case (e.g., Mark Alfano 2009 assumes, in discussing this case and others, that we cannot know things when we violate a truth adherence condition). According to Nozick (p. 177), the reason the Vatted brain fails to know it is Vatted is because the truth adherence condition is violated. It is not clear what method of belief formation is employed in the case, but, according to Nozick, there is a failure of knowledge because in close worlds in which the brain is Vatted it might not be brought to believe it is Vatted. Nothing like this can be good grounds for saying that the brain does not know it is Vatted. The same goes for what Nozick says in the next paragraph about Gilbert Harman's Dead Dictator case (1973, pp. 142-154). Here Nozick's diagnosis is once again that the truth adherence condition is violated. (But does the newspaper reader use the same method in the situation in which he believes the dictator is dead as he does in the possible situation in which he believes the dictator is not dead? It is not obvious. To help Nozick out, let us stipulate that the reader's method is: if a recent newspaper says the dictator is dead then the dictator is dead; if a recent newspaper says the dictator is not dead then the dictator is not dead.) Yet if the Vat and Dead Dictator cases are examples of ignorance, it is not because of a failure of adherence.

Still, even if we reject the adherence conditions, there may be a way to salvage Nozick's claim that epistemic sensitivity entails sensitivity to truth and falsity. In the next section attempts at salvage will be made, and rejected.

False Negatives

Assuming that we must reject Nozick's view of epistemic sensitivity, what should we put in its place? For falsity sensitivity, we might consider Nozick's own suggestion (modified appropriately), namely:

~*p* →~(M indicates *p*) (false positives avoidance).

After all, this component of the tracking account does not require adherence to truth value. But what should we adopt as our account of truth sensitivity, which, Nozick said, is the more important component of epistemic sensitivity? We might consider the corresponding avoidance condition:

p →~(M indicates ~*p*) (false negatives avoidance).

The equation of epistemic sensitivity with the conjunction of these two conditions we might call the *avoidance account*. (An interesting feature of this approach is that it handles the problem of inconsistent methods all by itself. Any method that indicates that *p* is both true and false will fall afoul of one avoidance condition or the other.)

Nozick himself, it appears, accepted the false negatives avoidance condition. He committed to it when he revised his adherence condition so as to deal with contradictory methods. That is, in accepting (his version of) the following condition,

p and S applies M → M indicates p and ~(M indicates ~p)

he committed to:

p →~(M indicates ~p) (false negatives avoidance).

We have already seen that the adherence condition is unacceptable, but Nozick could have given it up and still retained the false negatives avoidance condition.

However, he probably would not have done so. If he had, he would have been committed to the avoidance account, and this view is inconsistent with his claim that knowing p involves a stronger relation to p than to ~p. If we play by his rules, we can retain the false negatives avoidance condition as our account of truth sensitivity, but then we will have to jettison the last piece of Nozick's original analysis. We will have to replace the false positives avoidance condition with something weaker, such as the requirement that, were p false, M *might* not indicate p. No one will want to go down this road. (We might also look into pairing Nozick's false positives avoidance condition with something stronger than the false negatives avoidance condition, but the only candidate in sight is Nozick's truth adherence condition, which we have already rejected.)

Alternatively, if the two avoidance conditions strike us as being the right account of epistemic sensitivity, we could reject Nozick's asymmetry thesis, and announce the discovery that knowing p involves an equally strong relation to p and to ~p. This would deprive Nozick of his reason for rejecting the two adherence conditions, but that is of no concern, since we have found an excellent reason to reject them anyway.

However, even if we were willing to reject the asymmetry thesis, I doubt that the avoidance account would have many takers. It faces a serious objection, which stems from the way it handles false negatives.

The difficulty I have in mind can be illustrated using either of the following methods of belief formation:

> *Positive Fever Detection* (PFD): if someone is alive and her temperature is over 101 then she has a fever, and if she is alive and her temperature is between 95 and 101 then she does not have a fever.
>
> *Chickenpox Virus Detection* (CVD): If someone has shingles, he is infected with the chickenpox virus; if he does not have shingles, he is not infected with that virus.

If we use PFD to establish that Frieda has a fever, or CVD to establish that she is infected with the chickenpox virus, we will not go wrong; that is, it will not give us false positives. Both methods may, however, give us false negatives—PFD when, for example, a person's temperature is 99, and CVD when she has childhood chickenpox. (I am assuming that children who get chickenpox will get shingles, if at all, only much later in life).

I suggest that we can know that Frieda has a fever using PFD, or the virus using CVD. What we cannot do is know that she does *not* have a fever using PFD, or that she does not have the virus using CFD. (Applied to the proposition that Frieda does not have a fever, PFD gives false positives: it is positive relative to this proposition when it says she does not have a fever, and falsely so when her temperature is 100.) We can know things even though the methods on which we rely might give false negatives. It suffices that our methods would not give us false positives.

If this is correct, then we must reject the avoidance account, since it relies on the false negatives avoidance condition. Given this condition, we could not know that Frieda has a fever using PFD. If someone has a temperature of 100, she is feverish, but PFD would indicate that she is not.

Let me counter some reasoning that might suggest a different conclusion about false negatives. Up to now I have used the term *positives* in connection with methods for determining whether some proposition *p* is true: positives occur when the method indicates *p*, and negatives when it indicates ~*p*. Contrast this terminology with a more common way of speaking, according to which a positive occurs when a method indicates that something *x* has some property F, and a negative occurs when M indicates that *x* does not have F. Let us call the latter *property* positives and negatives, and the former *propositional* positives and negatives. I have claimed that knowing things does not require our using methods that avoid false propositional negatives. However, my claim might seem false if we fail to distinguish between propositional and property negatives.

To see why, suppose once again that we are using PFD as our method of deciding whether or not Frieda has a fever. Earlier it was noted that, so employed, PFD will give false propositional negatives but not false positives. We can add that PFD will give false property negatives (when her temperature is 100) but not false positives. However, matters are interestingly different if we apply PFD to the proposition that Frieda does *not* have a fever. Applying PFD to *this* proposition changes what counts as a propositional positive or negative. PFD will give false propositional positives (e.g., at 100) but not false negatives. Yet the change in application does not alter what counts as a property positive or negative. It remains true that PFD gives false property negatives (at 100) but not positives. We can sum up this way:

(1) S knows *p* via M only if M does not give false propositional positives. It does not matter if M gives false propositional negatives.

(2) S knows that *x* is F via M only if M does not give false property positives when applied to *x* and F. It does not matter if M gives false property negatives.

(3) S knows that *x* is *not* F via M only if M does not give false property negatives when applied to *x* and F. It does not matter if M gives false property positives.

My claim is that knowing that a proposition is true never requires the use of a method that avoids false propositional negatives, not that it never requires the use of a method that avoids false property negatives. (In the appendix I discuss (2) at further length.)

My view concerning false negatives is more controversial than my stance concerning undetected instances. Plenty of theorists have said things that are inconsistent with it. Here are some examples.

The first is Nozick himself. Earlier it was noted that his official reason for denying that the Vatted brain knows it is Vatted is that there is a violation of adherence. He explains knowledge failure in the Dead Dictator case in the same way. However, in some passages he explains the failure of knowledge by appealing to false negatives. The Vatted brain might easily have ended up falsely believing that it was not Vatted (p. 176); the newspaper reader might have ended up falsely believing that the dictator was not dead (p. 177). (Had he read the other papers, he would not know that the dictator was alive, but that is because he would not *believe* the dictator was alive. It does not follow that he fails to know the dictator is alive when he does believe it via reading a newspaper.) Nozick's seeming equivocation is understandable; as noted earlier, in reformulating his adherence account, he ended up requiring *both* truth adherence *as* *well as* false negatives avoidance. However, given what I have said about undetected instances and counterinstances, neither of Nozick's explanations is correct.

The second example is Keith DeRose (1999, p. 204), who, in defending a version of David Lewis's (1996) contextualist approach to knowledge, claimed that "an important component of being in a strong epistemic position with respect to P is to have one's belief as to whether P is true match the fact of the matter as to whether P is true, not only in the actual world, but also at the worlds sufficiently close to the actual world." On this view, avoiding false negatives is as important as avoiding false positives. That cannot be right. Recall, for example, that we can know that Frieda has a fever using PFD; in close worlds in which it indicates she has a fever, she does. But in worlds that are at least as close PFD misleads us into thinking that she does not have a fever. Here we know something even though there are close worlds in which our belief as to whether *p* is true does not match the fact of the matter. Needless to say, there need not be a match in the other direction of fit either—from fact to matching belief. There are circumstances in which we can know that Frieda has a fever using GFD, as well as circumstances in which we can know that she does not. However, there are worlds that are at least as close in which Frieda has a fever yet GFD will not so indicate, as well as worlds in which she lacks a fever yet GFD will not so indicate; in such worlds, we will lack a belief that matches the fact of the matter.

Here is a third example: according to Alvin Goldman (1979, p. 178), justification, as he analyzes it, is necessary for knowledge, and he analyzes it roughly as follows (1979, p. 183, and compare 1986; what I say about Goldman's rough account is also true, mutatis mutandis, of his more refined analysis): "The justificational status of a belief is a function of the reliability of the process or processes that cause it, where (as a first approximation) reliability consists in the tendency of a process to produce beliefs that are true rather than false." A process that causes the belief *p* is reliable in Goldman's sense only if it lacks the tendency to produce the belief ~*p* when *p* holds; that is, a process will be unreliable, hence incapable of positioning us to know *p*, if it tends to produce false negatives—or rather a high proportion of them as compared to true beliefs. (Kelly Becker's version of Goldman's account, requiring that beliefs be formed by a process that "produces a high ratio of true beliefs in the actual world and throughout close possible worlds," faces the same difficulty [Becker 2009b, Chapter 2].)

Sensitivity to False Negatives and Positives

According to Nozick's approach, a method is sensitive to the truth of *p* just when its indication status holds firm in some way in close worlds in which *p* is true; it is sensitive to the falsity of *p* just when its indication status holds firm in some way in close worlds in which *p* is false. (Here 'holding firm' might be its consistently indicating *p*, or its consistently failing to indicate ~*p*, for example.) I have suggested that this conception is mistaken. However, it may be possible to replace Nozick's account of sensitivity with one that is defensible. On Nozick's view, what methods indicate must somehow match truth-value in salient worlds. Perhaps we should reverse the direction of fit, so that truth-value must somehow match what methods indicate in salient worlds. Let us see where this idea takes us.

On the view under consideration, truth sensitivity demands that *p* be true in salient worlds in which our method indicates that *p* is true, while falsity sensitivity requires that *p* be false in salient worlds in which our method indicates that *p* is false. Call this the *reversed sensitivity account*. Note that on this view truth sensitivity will always be a restriction against false positives, while falsity sensitivity will always be a restriction against false negatives. Note, too, that adopting the reversed sensitivity account will force us to reclassify Nozick's own avoidance condition. He offered it as an account of falsity sensitivity. On the revised account, it is an account of truth sensitivity.

If we take this view of epistemic sensitivity on board, another analysis of epistemic sensitivity is worth considering. We arrive at this analysis by contraposing the conditions that constitute the avoidance analysis, giving us the following:

S is epistemically sensitive to *p* iff S believes *p* via a method M such that:

M indicates $p \rightarrow p$, and

M indicates $\sim p \rightarrow \sim p$.

The first of these conditions has come to be known as the 'safety' condition (I defended a version of it in Luper 1984). I will call it the *positive safety* condition. The second I will call the *negative safety* condition. The account consisting of both I will call the *combined safety account*. The positive safety condition says that in the closest worlds in which M indicates *p*, M is correct, so that if we are relying on M, the belief we form concerning *p* will be true. The second condition, negative safety, adds that M must be correct in the close worlds in which M indicates ~*p*. (Hence, like the avoidance account, the combined safety account handles the possibility of inconsistent methods all by itself.)

Unfortunately, we cannot salvage Nozick's conception of epistemic sensitivity using the combined safety account. Perhaps we could live with the fact that it violates Nozick's asymmetry thesis. (In fact, we could restore asymmetry by wedding positive safety or Nozick's avoidance condition, now seen as ensuring sensitivity to truth, with a weaker condition ensuring sensitivity to falsity). However, on the assumption that we can know things using methods that give false negatives, we must reject the combined safety account, since the negative safety condition rules out false negatives. So even reversing the direction of fit required for epistemic sensitivity does not enable us to rescue Nozick's view.

Let me sum up what I have said so far. I considered four accounts of epistemic sensitivity (five if we count the contraposition of the adherence account):

Tracking account:

S is epistemically sensitive to *p* iff S believes *p* via a method M such that:

p and S applies M $\rightarrow$ M indicates *p* (truth adherence), and

$\sim p \rightarrow \sim$(M indicates *p*) (false positives avoidance).

Adherence account:

S is epistemically sensitive to *p* iff S believes *p* via a method M such that:

p and S applies M $\rightarrow$ M indicates *p* (truth adherence), and

~*p* and S applies M $\rightarrow$ M indicates ~*p* (falsity adherence).

Avoidance account:

S is epistemically sensitive to *p* iff S believes *p* via a method M such that:

$p \rightarrow \sim$(M indicates ~*p*) (false negatives avoidance), and

$\sim p \rightarrow \sim$(M indicates *p*) (false positives avoidance).

Combined safety account:

S is epistemically sensitive to *p* iff S believes *p* via a method M such that:

M indicates $p \rightarrow p$ (positive safety), and

M indicates $\sim p \rightarrow \sim p$ (negative safety).

I noted that knowledge is consistent with undetected instances, so we can rule out the first two of these accounts. I also suggested that knowledge is consistent with false negatives, so we can also rule out the avoidance and combined safety accounts. And I said that every account of truth sensitivity in sight (e.g., the first conditions of the first three of the above accounts) forces us to accept overly strong conditions for knowledge, which supports the conclusion that truth sensitivity, as Nozick understands it, is simply not a requirement for knowledge. Now suppose we recast his account of epistemic sensitivity, reversing the direction of fit, so that truth-values must match what methods indicate rather than the other way around, and so that truth sensitivity consists in eliminating false positives, while falsity sensitivity consists in eliminating false negatives. So recast, Nozick's conception is still false. We can know things using methods that give false negatives, so sensitivity to falsity is not requisite for knowledge.

Nothing said so far forces us to deny that Nozick's avoidance condition is a necessary condition for knowledge. And the same can be said for the positive safety condition, and the combination of the two. Neither condition requires that we eschew methods that miss instances or that give false negatives. However, now shorn of the support it might derive from Nozick's account of epistemic sensitivity, his avoidance condition has lost much of its appeal.

Still, there is more to be said about avoidance and safety.

Safety and Avoidance

Is there any reason to say that the safety and avoidance conditions are *both* necessary for knowledge? That is, should we accept the following combination:

S knows *p* only if S believes *p* via a method M such that:

M indicates $p \rightarrow p$ (safety), and

$\sim p \rightarrow \sim$(M indicates p) (avoidance)?

Not that I can see. Someone wedded to the latter might want to add the former on the grounds that avoidance is useless when it comes to necessary truths (we vacuously meet the avoidance condition no matter how we come to believe necessary truths). But if the safety condition really did handle necessary truths properly, we would need to look elsewhere for grounds to retain the avoidance condition. Perhaps we might be struck by the thought that combining the two gives us conditions that cover both directions of fit: safety requiring an appropriate truth-value to match state of indication (here the indication that *p* is true), and avoidance requiring an appropriate state of indication (here the absence of erroneous indication) to match truth-value. However, it is hard to see why we should want to cover both directions of fit (especially if in one direction we only insist on the absence of erroneous indication). Much more striking is the fact that both conditions act to eliminate false positives. One cannot defend the avoidance condition against the safety condition, or vice versa, on the grounds that knowledge requires the elimination of false positives since *both* conditions concern false positives. Also striking is the fact that the two conditions make identical demands on worlds that are close to the actual world, when *p* actually holds. Of the worlds close to the actual world, safety demands that, throughout,

M indicates *p* only if *p*.

Of the same worlds, avoidance requires that, throughout,

$\sim p$ only if $\sim$(M indicates *p*).

Obviously these demands are identical.

Some (e.g., DeRose 1999; for discussion, see Black and Murphy, 2007) have also offered examples like the following against the avoidance condition:

Hands: by virtue of being appeared to handly, I believe the following: I am not being led to falsely believe that I have hands via being appeared to handly.

If I were being led to falsely believe that I have hands via being appeared to handly, I would still be appeared to handly, and led thereby to believe I am not being led to falsely believe that I have hands via being appeared to handly. So I cannot meet the consequent of the avoidance condition. But I have no trouble meeting the safety condition.

To these points, I have little to add in defense of the safety over the avoidance condition. What I said in an earlier essay (Luper 1984) still seems correct. The salient way in which the two conditions differ concerns *which* false positives they exclude. No one (except perhaps theorists who equate knowledge with certainty, as, e.g., Unger 1975 does) thinks that it must be *impossible* for a method of belief formation to give false positives if it is to serve as a means to knowledge. The safety condition says that methods are viable only if they do not give false positives in the closest worlds in which they give positives at all, which includes the actual world. The avoidance condition requires that methods not give false positives even in the closest worlds in which *p* is false. But the closest worlds in which *p* is false might be remote indeed. They might be worlds in which the laws of physics are very different or even nonexistent. Does it really matter that we get false positives in such worlds? If physical laws were extremely different, there might be things that from the outside exactly resemble butterflies or doorstops but whose inner lives resemble those of human beings. In some such worlds there might even be a butterfly or doorstop whose inner life exactly matches your own. Don't you know you are not such a thing? If so, we must reject the avoidance condition, but not the safety condition.

The implications of the safety and avoidance conditions are very similar except when it comes to skeptical scenarios. As for whether we know that skeptical hypotheses are false, I presume that people disagree in their intuitions. Several theorists, like Unger and Becker, think it is counterintuitive to claim to know the falsity of skeptical hypotheses. Other theorists (among them Sosa; see 2003, p. 180, note 3) think it counterintuitive to claim that we do not know the falsity of skeptical hypotheses. But virtually everyone agrees that the closure principle, suitably formulated, is highly plausible. Since the safety condition is friendly to closure, while the avoidance condition is hostile to it (as is the combination of both of these conditions), it is best to opt for the safety condition. Of course, this argument is not conclusive; nowhere in this essay have I defended the claim that the safety condition is a *sufficient* condition for knowledge (I deny it is sufficient in Luper 1984). Conceivably, some additional necessary condition, alone or in conjunction with safety, will be incompatible with closure.

Closure, Skepticism, and False Positives

In closing, let me allay a suspicion which critics might entertain. Suppose I was mistaken when I said that knowledge is consistent with false negatives. Will it follow that knowledge is not closed under entailment? Safety theorists, including myself (Luper 1984), defend closure and claim that we know that skeptical hypotheses are false; some of their critics call them 'dogmatists.' Does 'dogmatism' hinge on the position we take concerning false negatives?

If knowledge is indeed inconsistent with false negatives, we will need to reconsider the avoidance and the combined safety accounts. This will once again raise the question of closure. However, suppose (as Nozick did) that we can know things using one-sided belief formation methods: methods that cannot indicate that *p* is false, but that can indicate that *p* is true. Then 'dogmatism' is easily defended, and damage to the closure principle is easily limited. Let me explain.

First a concession: if the avoidance account is correct, it seems best to reject dogmatism and closure. Accepting the use of one-sided methods will not help. The first condition of the avoidance account, the false negatives avoidance condition, will be met automatically by one-sided methods. So, for one-sided methods, the avoidance account reduces to the simple avoidance account. That means that even when we employ one-sided methods, often we are in no position to know things that follow from things we know. Also, even when our methods are one-sided, we do not know that skeptical hypotheses are false.

Now suppose that the combined safety account is correct. Its second condition, the negative safety condition, will be met automatically by those who use one-sided methods. So for one-sided methods the combined safety account reduces to the simple safety account, and knowledge via one-sided methods is closed under entailment. Moreover, as long as we employ one-sided methods, we may come to know that skeptical hypotheses are false. Thus even if we deny that knowledge is consistent with false negatives, we can remain unrepentantly dogmatic. We need only adopt the combined safety account rather than the avoidance account.

While we are musing about the consequences of overturning the assumption that knowledge is consistent with false negatives, let me add something else: we probably would want to stop speaking of knowing that *p* is true and instead speak of knowing *whether or not p* is true. The reason is this: my assumption about false negatives was my grounds for rejecting the avoidance account and the combined safety account. Now, on both of these accounts, knowing *p* involves a relation to ~*p* that is just as strong as the relation to *p*. So if knowledge is *not* consistent with false negatives I see no reason to accept Nozick's asymmetry thesis. And if Nozick's thesis *is* false, it seems reasonable to say that knowledge is really a matter of knowing *whether or not* some proposition *p* is true, as opposed to (simply) knowing *p* is true. The conditions for knowledge could be stated accordingly. If we accept the avoidance account, we might state our position as follows:

> S knows whether or not *p* is true only if S's belief concerning *p* is arrived at via a method M such that:
>
> > $p \rightarrow$ ~(M indicates ~p) (false negatives avoidance), and
> >
> > ~$p \rightarrow$ ~(M indicates p) (false positives avoidance).

Similarly, if we accept the safety account, we might formulate our stance as follows:

> S knows whether or not *p* is true only if S's belief concerning *p* is arrived at via a method M such that:
>
> > M indicates $p \rightarrow p$ (positive safety), and
> >
> > M indicates ~$p \rightarrow$ ~p (negative safety).

(Either way, the truth condition would be redundant: whatever the truth value of *p*, M indicates that *p* has that value; since S believes what M indicates, S's belief is true. Note, too, that, if I am not only wrong about false negatives, but also about undetected instances, we could consider yet another account of knowing whether *p* is true: we could define it in terms of the adherence account.) Here again we can cling doggedly to dogmatism, assuming that using one-sided methods is kosher. We can also say that knowledge via one-sided methods is closed under entailment. We have only to accept the latter account of knowing whether *p* holds, rather than the former.

Appendix

I said that S's knowing that *x* is F via M does not hinge on whether M gives false property negatives. Here I offer further clarification of this claim.

When statisticians speak of positives and negatives, they use these terms in the property sense. Sometimes they evaluate a test in terms of its sensitivity and its specificity, which concern its track record of results. A test's *sensitivity* is the number of true positives (TPs) it issued divided by the sum of its true positives and false negatives (FNs), and a test's *specificity* is the number of its true negatives (TNs) divided by the sum of its true negatives and false positives (FPs); i.e.,

$$Sensitivity = \frac{number\ of\ TPs}{number\ of\ TPs + number\ of\ FNs}$$

$$Specificity = \frac{number\ of\ TNs}{number\ of\ TNs + number\ of\ FPs}$$

Here is an illustration. Suppose we have a test T for whether or not a person is sick. For simplicity, let's assume that T never omits to issue a verdict, and that the verdict will be either positive or negative. When a sick person is tested it will indicate that she is it sick (TP) or not sick—i.e., healthy (FN). Applied to a healthy person, its verdict will either be *sick* (FP) or *not sick*—i.e., healthy (TN). If T gains a track record high in TPs and low in FNs, it will be highly sensitive. If it is positive for *all* sick persons tested, hence *never* negative for them, it is perfectly sensitive. If these results are not a fluke, everyone who is sick will pass the test. However, a perfectly sensitive test might also be passed by people who are not sick. It might be (falsely) positive for some healthy people. (We can trust its negatives, but not its positives.) Here specificity comes in. If T gains a track record high in TNs and low in FPs, it will be highly specific. It is perfectly specific if it is (truly) negative for all healthy people tested and never (falsely) positive for them—if all healthy people fail it, which is compatible with its being (falsely) negative for the sick. (We can trust its positives, but not its negatives.) So all and only sick people will pass a test that is perfectly sensitive *and* perfectly specific, assuming, once again, that the test never omits to issue a verdict.

In the above reasoning, we assumed that T never omits to issue a verdict. Now let us drop that assumption, and suppose that, of each person tested, T will say she is sick (P), or that she is not sick (N), or that it will say nothing at all (O). Of a sick person it will say *sick* (TP) or *not sick* (FN) or neither. This last possibility we can label *OI* to signify that it is an omitted instance of sickness. Of a healthy person, T will say *sick* (FP) or *not sick* (i.e., healthy [TN]) or neither (OC—omitted counterinstance of sickness). Notice that such a test can overlook a sick person either by falsely saying she is not sick (FN) or by saying nothing (OI). So if we want a test that will overlook no sick people, it will need to avoid both FNs and OIs. Since a test can be perfectly sensitive even though it is compatible with OIs, sensitivity does not guarantee that no sick people will be overlooked. It does guarantee that they will not falsely test negative. But if we want T to say a person is sick only if sick, it suffices that we avoid FPs. It is all right if T omits some instances of sickness.

One might say that, ideally, a test should have a high degree of sensitivity as well as a high degree of specificity. (One might even say that, ideally, it should never omit to issue a verdict.) A test that is sensitive or specific but not both is much less informative than a test that is both. However, there is no reason to say that, to know that *x* is F, we must employ a method that is both sensitive and specific. If we are to know that *x* is F, we might need to believe *x* is F via a method that has perfect specificity, or rather that would have perfect specificity if it gained a track record. However, perfectly specific tests can issue false negatives and fail to detect cases in which *x* is F. They cannot be trusted when they say that *x* is not F. Anyone who insists that, to know *x* is F, we need a method that will not issue false negatives, is probably looking for a method that will position us to know whether it is true or false.

Steven Luper

Trinity University

I thank Curtis Brown and Kelly Becker for constructive comments on a previous draft.

References

Alfano, Mark, 2009, "Sensitivity Theory and the Individuation of Belief Formation Methods," *Erkenntnis* 70: 271-281.

Becker, Kelly, 2009a, "Margins for Error and Sensitivity: What Nozick Might Have Said," *Acta Analytica* 24: 17-31.

-----, 2009b, *Epistemology Modalized*, New York: Routledge.

Black, Tim, and Murphy, Peter, 2007, "In Defense of Sensitivity," *Synthese* 154.1: 53-71.

DeRose, Keith, 1999, "Solving the Skeptical Problem," in DeRose, K. and Warfield, T., eds., *Skepticism: A Contemporary Reader*, 1999, Oxford: Oxford University Press.

Goldman, Alvin, 1979, "What is Justified Belief?", reprinted in S. Luper, *Essential Knowledge: Readings in Epistemology*, New York: Longman, 2004. Page references are to the reprinted version.

-----, 1986, *Epistemology and Cognition*, Cambridge: Harvard University Press.

Harman, Gilbert, 1973, *Thought*, Princeton: Princeton University Press.

Lewis, David, 1996, "Elusive Knowledge," *Australasian Journal of Philosophy* 74: 549-567.

Luper(-Foy), Steven, 1984, "The Epistemic Predicament: Knowledge, Nozickian Tracking, and Skepticism," *Australasian Journal of Philosophy* 62: 26-50.

-----, 2003, *The Skeptics: Contemporary Essays*, Aldershot, UK: Ashgate Publishing, Ashgate Epistemology & Mind Series.

Nozick, Robert, *Philosophical Explanations*, 1981, Cambridge: Belknap Press of Harvard University Press.

Pritchard, Duncan, 2002, "Resurrecting the Moorean Response to the Sceptic," *International Journal of Philosopical Studies* 10: 283-307.

Sosa, Ernest, 2003, "Neither Contextualism Nor Skepticism," in Luper 2003, 165-182.

Unger, Peter, 1975, *Ignorance: A Case for Skepticism*, Oxford: Oxford University Press.