

When Sensitivity Conflicts with Closure

JOSHUA KAMINASH

The traditional tripartite account of knowledge as justified true belief is vulnerable to the Gettier counterexamples, in which although an agent holds a justified true belief, he clearly does not possess knowledge, since his belief is true merely in virtue of some coincidental circumstance in the world unconnected to his reasons for holding the belief. In these cases, it seems that the agent lacks knowledge because there is some crucial connection missing between his internal belief and the relevant external state of affairs in the world. By demonstrating that internal justification fails to guarantee this connection, the Gettier cases are apt to inspire the idea that the proper connection between a belief and the world might be secured in some other fashion, unmediated by justification.¹ For someone taking up this suggestion, the challenge is to articulate in what particular manner a true belief about the world must be connected with the world in order for it to count as knowledge.

Robert Nozick locates the source of the crucial connection in an agent's propensity to maintain true beliefs even as conditions in the world

¹That is, in addition to the fact that a true belief's being justified is not a sufficient condition for knowledge, the idea is that it's also not a necessary one. In particular, the Gettier cases are prone to inspire this idea since they are prone to inspire the idea that internal justification, because it is fallible, is fundamentally ill-suited to provide the desired connection between an (internal) belief and the (external) world.

Joshua Kaminash is completing his undergraduate studies in philosophy this year at New York University. His primary interest is in metaphysics, and he is completing a thesis on the question of whether, and under what conditions, material objects exist.

change. In particular, he claims that what it is for an agent to know something is for the agent to be disposed to respond appropriately to potential changes in the world, including changes in the truth-value of the relevant proposition, across a range of counterfactual situations. Whereas someone with a mere true belief that p (where p is some proposition or other) is not expected to maintain an appropriate attitude toward p as either the truth-value of p or other circumstances in the world shift, someone who knows that p should maintain an appropriate attitude toward p even when such shifts occur. Nozick proposes the following formal account of knowledge: An agent, S , knows that p if and only if

- (i) p is true;
- (ii) S believes that p ;
- (iii) if p were not true and S were to use the same method, M , by which he arrived at his belief that p , to arrive at a belief whether (or not) p , then S would not come to believe, by M , that p ; and
- (iv) if p were true and S were to use M to arrive at a belief whether (or not) p , then S would come to believe, by M , that p .² (179)

Nozick's proposed account of knowledge purports to identify the necessary and sufficient conditions for knowledge. In order to evaluate Nozick's account, I will assess to what degree the account, when applied to individual cases, produces results that conform with our intuitions: I will regard the account as successful to the extent that it both counts as knowledge those instances of belief we intuitively regard as instances of knowledge, and does not count as knowledge those instances of belief we intuitively deny to be instances of knowledge. In this paper, I will first evaluate the account with respect to several individual cases. Second, I will attempt to explain why the account sometimes fails to deliver the proper result. Finally, I will identify an additional seemingly undesirable consequence of the account, the denial of closure, and assess whether this consequence is in fact undesirable.

Nozick's account correctly identifies as instances of knowledge many beliefs we intuitively regard as such. For example, consider my belief that I have hands, formed on the basis of visual inspection. On the assumption that my belief is true, our intuitions clearly indicate that it is an instance

²As Nozick notes, in order to be filled out, the account would need to be supplemented both with rules for specifying how, for any given case of belief formation, the relevant method is to be characterized, and with a semantics for subjunctive conditionals to elaborate conditions (iii) and (iv).

of knowledge. On that same assumption, Nozick's account indicates that I know that I have hands provided that conditions (iii) and (iv) are satisfied. In order for (iii) to be satisfied, it must be the case that if I didn't have hands and I were to use visual inspection to arrive at a belief whether or not I had hands, then I would not, by that method, arrive at the belief that I had hands. I will interpret the subjunctive conditional sentences of (iii) and (iv) using the same intuitive possible worlds semantics that Nozick employs.³ Accordingly, (iii) is satisfied if and only if in all the closest possible worlds in which I don't have hands (and in which I arrive at a belief, by visual inspection, as to whether or not I have hands), I *don't* come to believe (by that method) that I have hands.⁴ Apparently, the closest worlds in which I do not have hands are those worlds that otherwise resemble the actual world but in which my hands have been removed.⁵ Since visual inspection is still reliable in those close worlds, I will correctly perceive myself as handless. So (iii) is satisfied. In order for (iv) to be satisfied, it must be the case that in all nearby worlds in which I have hands and arrive at a belief whether or not I have hands by the same method, I come to believe by that method that I have hands. And since my method of belief formation is reliable in nearby worlds, I do arrive, by that method, at the belief that I have hands in all those worlds. So (iv) is also satisfied, and therefore Nozick's account rightly indicates that I know that I have hands.⁶

Nozick's account also successfully handles some potentially problematic cases. For example, imagine that Jane is a disembodied brain (DEB), electrically stimulated by a computer to have various experiences and beliefs.⁷ On a whim, the scientist on the overnight shift modifies the computer

³Although Nozick here does not endorse any particular semantics for subjunctive conditionals, he informally employs a (deliberately sketchy) possible worlds semantics in order to elucidate the subjunctive conditionals contained in the account.

⁴The degree of closeness of a world (to the actual world) corresponds with the degree to which that world resembles the actual world.

⁵Or maybe those in which I was born without hands. At any rate, any world that is extremely *dissimilar*, such as any world in which I am a brain in a vat, is not a close world.

⁶Of course, a skeptic might not regard this as a desirable result; however, as I noted, I am here regarding the account as successful to the extent that it corresponds with our intuitive conception of knowledge.

⁷I have chosen to use the term 'disembodied brain', rather than the traditional 'brain in a vat', since the second 'brain in a vat' case, discussed below, involves the occasional removing of brains from their vats, and I wish to avoid causing any distracting concerns about what it might mean for a 'brain in a vat' to not be in a vat.

program in a way that causes Jane to form the true belief that p : she is a DEB. Condition (iv) requires that in all nearby p -worlds (in which she employs the same method of belief-formation) she believe that p . But since the capricious scientists could have easily chosen to stimulate her in such a way that she would not have believed that she was a disembodied brain (in fact, that is what they have always done, prior to tonight), it is clear that it is not the case that in all close p -worlds she believes that p . Thus condition (iv) is not satisfied, and the account appropriately indicates (in concordance with our intuitions) that she lacks knowledge that p .⁸

However, Nozick's account fails with respect to the following DEB case. Dr. Strangelove is a very conscientious brain-envatter and has always felt a duty toward the DEBs under his supervision to not conceal from them the circumstances of their lives. Strangelove has now perfected a sturdy, tamper-resistant, and hyper-reliable mechanism for causing true beliefs in his envatted subjects: belief-causing prongs. The metal prongs are customarily affixed to the bottom of the vats, projecting upward; when a temporarily unenvatted brain is lowered onto the prong within the liquid-filled vat, the prong becomes firmly inserted into the brain stem. Whenever a brain is envatted and empronged, the prong, by a chemical reaction with the brain tissue, causes the subject to form (and to maintain, so long as she is envatted and empronged) the belief that she is an empronged and envatted DEB. Although the prongs are designed for use within vats, it is also possible to emprong an unenvatted brain. However, the prong is made of a highly reactive alloy whose chemical properties radically change when any portion of the prong (or even of the semi-porous brain into which it is inserted) is exposed to air. As a consequence, when an unenvatted brain is empronged, the prong causes the subject to form the (false) belief that she is an embodied human being.

Now take Jennifer, an envatted and empronged DEB. As a consequence of her empronging, she holds the true belief that p : she is an envatted and empronged DEB. In order for (iii) to be satisfied, it must be the case that in all the closest *not*- p -worlds in which Jennifer arrives at a belief whether or not p by being empronged, she does not believe that she is an empronged and envatted DEB. Clearly, the closest *not*- p -worlds in which she arrives at a belief whether or not p by being empronged are the closest worlds in which

⁸Unlike in a classic Gettier case, Jane's belief is not disconnected from the circumstances in the world that make it true. In fact, her being a DEB is one of the causes of (or at least, causal conditions for) her forming her present belief. Nonetheless, Jane clearly lacks knowledge here since, even if the relevant circumstance in the world is part of the reason she holds the belief, that circumstance is clearly not connected to her belief in the right kind of way.

Jennifer is empronced but unenvatted. Since in those worlds the emproncing causes her to believe that she is an embodied human, and not an empronced and envatted DEB, (iii) is satisfied. In order for (iv) to be satisfied, it must be the case that in all nearby p -worlds in which Jennifer is empronced, she forms, by her emproncing, the belief that p . In all p -worlds, Jennifer is envatted and empronced. The method of emproncing is hyper-reliable—whenever it is performed on an envatted brain, it always succeeds in producing the belief that p . Therefore, in *all* p -worlds (not just the close ones), Jennifer forms, by her emproncing, the belief that p . So condition (iv) is satisfied. Therefore, according to Nozick's account, Jennifer knows that she is an empronced and envatted DEB. This is clearly not the desired result, however. Apparently, Nozick's account bizarrely implies that it can make a difference, with regard to whether a belief constitutes knowledge, whether the belief is caused by computer-directed electronic stimulation or by a chemically reactive metal prong.

With respect to Jennifer's case, Nozick's account is too liberal, since it declares that she has knowledge where actually she has none. The following case demonstrates that Nozick's account is also sometimes too restrictive. Imagine that Professor Goodbar is supervising an exam being taken by his students. Bad Adam is a skilled cheater; despite his serial cheating, he has never been caught. His technique involves sitting in the middle of a row of seats and stowing a cheat-sheet underneath the seat in front of him that is visible only from his immediate vantage point. Good Judy, who would never dream of cheating, happens to be seated beside him. At some point during the exam, Judy politely raises her hand in order to summon Goodbar to ask him a clarificatory question. As Goodbar approaches her seat, however, he stumbles over Judy's umbrella and onto Adam's lap, from which position he observes Adam's fraud. Intuitively, it seems clear that Goodbar, upon observing Adam's cheating, comes to know that p : Adam is cheating. However, in order for (iv) to be satisfied, it must be the case that in all nearby p -worlds in which Goodbar employs the same method of belief formation, he thereby forms the belief that p . Goodbar apparently formed his belief that p by visually inspecting his surroundings. However, since Adam had concealed his cheat-sheet so well, it was only by great fortuitousness that Goodbar came to believe that p . If any one of the string of lucky circumstances (Judy sitting next to Adam, Judy asking a question, Judy leaving her umbrella in the aisle, Goodbar tripping and falling in exactly such a way onto Adam, etc.) were even slightly altered, then Goodbar would not likely have come to believe that p . Accordingly, a survey of nearby p -worlds reveals that Goodbar does not believe that p in all such worlds. So condition (iv) is not satisfied, and the account indicates, contrary to our strong intuitions on the matter, that Goodbar does *not* know that p .

A defender of Nozick's account might claim that our analysis here is flawed since a more precise description of Goodbar's method of belief formation could have yielded a different result. The defender might propose that we say that Goodbar arrived at his belief by visually inspecting his surroundings *while occupying a suitable vantage point*, for example. With the aid of such a description, (iv) would apparently be satisfied, since in all nearby p -worlds in which Goodbar employs that method, he apparently comes to believe that p . But for the defender to redescribe the method in such a manner seems egregiously contrary to the spirit of Nozick's theory of knowledge. For condition (iv) expresses the idea that an agent's belief counts as knowledge only if his belief is secure enough that it could not be shaken by any slight changes in external circumstances. While it may be true that Nozick's account specifies no rule for how an agent's method is to be described, the defender's proposed description, by contrast with my proposed description, is highly unintuitive. And it is unintuitive because the phrase "while occupying a suitable vantage point" appears to denote a relational property of Goodbar that does not pertain to any (internal) method of belief formation. Imagine one close possible world, in which all the minute circumstances that conjoin to cause Goodbar to trip and fall remain constant, but in which, at the moment he falls, the lights in the room flicker and consequently he fails to see the cheat-sheet. In view of such a possibility, should Goodbar's method be properly characterized as "visual inspection of his surroundings, *while occupying a suitable vantage point, and while under suitable lighting conditions*"? Many external factors have the potential to impair an agent's ability to exercise his cognitive faculties; if the agent's method must be conceived of as encompassing the whole range of external conditions such as are optimal for the agent to exercise his cognitive faculties, condition (iv) would be rendered virtually vacuous.

In the case of Goodbar and that of Jennifer, the agent arrives at a true belief as a result of good fortune. What appears to importantly distinguish the cases epistemically, however, is that whereas Goodbar arrives at his true belief by his own cognitive accomplishment, Jennifer does not. In particular, it is Goodbar's cognitive accomplishment that intuitively appears to be knowledge-conferring, and it is precisely the *absence* of cognitive accomplishment that seems to cause us to reject the possibility that Jennifer has knowledge. However, condition (iv) is apparently ill-suited to differentiate instances of agents whose good luck merely *gave them occasion* to exercise their own epistemically useful cognitive faculties from instances of agents whose good luck caused them to form a true belief in the absence of any relevant cognitive accomplishment. As Jennifer's case demonstrates, Nozick's account doesn't require that knowledge be a result of cognitive accomplishment; what is crucial for the account is that the method of belief-formation

be reliable. On the other hand, Goodbar's case demonstrates that the account's emphasis on reliability of method causes the account to deny that some genuine instances of knowledge, attained by the agent's own cognitive accomplishment, are in fact instances of knowledge.

In addition to these defects, however, Nozick's account has at least one much more radical and potentially damning consequence: it requires the rejection of closure. Closure is the principle that knowledge can be transmitted across known entailments:

Where p and q are propositions, if S knows that p , and knows that p entails q , then S knows that q .

Closure implies, for example, that if I know that James is in Barcelona, and I know that 'if James is in Barcelona, then he is not in France,' then I know that James is not in France. Closure is so highly intuitive a principle that it is hard to imagine that it could be wrong. Nonetheless, Nozick's account of knowledge is inconsistent with closure since the account indicates that (for certain instances of propositions p and q) an agent may know that p , and may know that p entails q , yet still cannot know that q .⁹ For example, the account indicates that I can know that I have hands (as was demonstrated earlier). Also, the account indicates that I can know that the following entailment holds: if I have hands, then I am not a DEB.¹⁰ Therefore, if closure is true, then I can know that I am not a DEB. However, Nozick's account indicates that I *cannot* know that I am not a DEB. For assume that I truly believe that I am not a DEB, on the basis of all my actual evidence. Condition (iii) requires that in all the closest worlds in which I am a DEB and I arrive at a belief whether or not I am a DEB on the basis of all my actual evidence, I come to believe that I am a DEB. But the worlds closest to this world in which I am a DEB (and I arrive at the belief whether or not I am a DEB by actual evidence) are the worlds in which as little as possible is different from this world; in those worlds, therefore, I will go on believing

⁹One potential variant of closure, which is even harder to imagine to be false than the one we are considering, is that if the first two conditions hold, then the agent should know that q . Nozick's account, of course, is inconsistent even with this weaker principle.

¹⁰Let the proposition z be: If I have hands, then I am not a DEB. In all possible worlds, z is true. Therefore, (iii) is trivially satisfied. Now let's say that an agent S believes z just by reasoning. In all close z -worlds (which is to say, all close worlds) in which S forms a belief whether or not z by reasoning, S will come to believe, by reasoning, that z (since S 's reasoning faculties will not significantly differ in close worlds). Therefore, (iv) is also satisfied.

that I am *not* a DEB.¹¹ Since condition (iii) is not satisfied, I do not know that I am not a DEB.¹²

Even leaving aside the violation of closure, Nozick's account produces results that are seemingly incompatible¹³: I know that I have hands, yet I do not know that I am not a DEB. This apparently absurd outcome, moreover, is no mere accident of the account. Rather, it arises as a consequence of how Nozick's theory is fundamentally conceived. Condition (iii) ensures that an agent who believes (truly) that p would respond appropriately to a change in p 's truth-value, such that the agent would not stubbornly maintain a belief that p if p were false. For convenience, let us call an agent's belief (with respect to a proposition p , and a method M) "sensitive" when it satisfies condition (iii). But the considerations that determine whether a belief is sensitive, and therefore, whether the belief could be an instance of knowledge, vary widely depending on the nature of the proposition in question. In order to assess whether I know that I have hands, my attitude toward the relevant proposition must be evaluated at fairly nearby worlds. By contrast, in order to assess whether I know that I am not a DEB, my attitude toward the relevant proposition must be evaluated at *radically different worlds*. But since in general, an agent will be more prone to form inappropriate beliefs at very far-off worlds than at close worlds,¹⁴ it should not be surprising that an agent's belief that he has hands is sensitive, but his belief that he is not a DEB is not. A defender of Nozick's account might construe the foregoing as a robust explanation, in terms of sensitivity, of why an agent can simultaneously know that he has hands and *not know* that he is not a DEB. Alternatively, the fact that Nozick's requirement that an agent be sensitive entails these seemingly incompatible results might be considered a good reason to reject his account.

On the one hand, closure seems to be too intuitively true to reject, but on the other hand, there is a way that Nozick's account seems to

¹¹ The fact that I could form a belief that I am not a DEB, by all my actual evidence, even if I were a DEB, is the basis of the skeptical hypothesis.

¹² Duncan Pritchard objects to characterizing the agent's method in this fashion, arguing that an agent's actual method of belief-formation in the actual world might involve perceptual faculties unavailable to a DEB. If so, then the account would generate the opposite result and therefore would not here generate a conflict with closure. would not here generate a conflict with closure.

¹³ Even if not logically contradictory.

¹⁴ Since in radically different worlds, there is more potential for things to "go wrong", epistemically. In the relevant instance, this is especially true: a DEB is more prone to form inappropriate beliefs than a handless person.

produce results that mirror our actual intuitions on the relevant cases. If an untutored person were asked to consider whether he knows that he is not a DEB, it would be quite natural to imagine him (provided that he were willing to entertain the question at all) reasoning counterfactually and in a way that parallels the relevant application of Nozick's account. He might say to himself, "Well, of course I *believe* that I am not a DEB, and there are certainly no indications that I *am* a DEB, but then again, if I *were* a DEB, I wouldn't necessarily believe that I was a DEB either. Maybe I *don't* know that I am not a DEB, after all. But I will still go on believing it, because it is much more rational to maintain this belief than to discard it." On the other hand, although closure considered abstractly is quite intuitive, the relevant *instance* of closure is *not* so intuitive. It is much harder to imagine our friend, upon being asked if knows that he is not a DEB, thinking to himself, "I know that I have hands. And I know that if I have hands, then I am not a DEB. Therefore, I must know that I am not a DEB." And the reason that our friend is unlikely to argue by closure in this case (even if the closure argument explicitly occurred to him), I think, is a good one: he senses that, relative to the question of whether he is a DEB, his attitude toward the proposition that he has hands is not germane. He intuitively recognizes that whatever evidence might justify his belief that he is not a DEB would have to be of an entirely different kind than the evidence that justifies his belief that he has hands. Although closure is highly intuitive both in its abstract form and in *most* of its instances, the instance of closure in quotation marks above is far from intuitively true. At the very least, if we persist in regarding an account of knowledge as successful to the extent that it mirrors our intuitions, then the fact that Nozick's account comes into conflict with closure in this circumstance should be regarded as a merit of the account.

Works Cited

This paper was produced for a course given by Professor Crispin Wright. I am grateful for his incisive analysis, and that of Max Barkhausen, on the topics discussed here.

Nozick, Robert. *Philosophical Explanations*. Cambridge: Harvard UP, 1981.
Pritchard, Duncan. "Sensitivity, Safety, and Antiluck Epistemology." *Oxford Handbooks Online*. Oxford UP. 2 Sept. 2009. Web. 1 May 2013.