Why There is no Justified Belief At Demon Worlds

Thomas D. Senor
Department of Philosophy
318 Old Main
The University of Arkansas
Fayetteville, AR 72701

Phone: 479-575-3551
Fax: 479-575-2642
Email: senor@uark.edu

Abstract:
The New Demon World Objection claims that reliabilist accounts of justification are mistaken because there are justified empirical beliefs at demon worlds—worlds at which the subjects are systematically deceived by a Cartesian demon. In this paper, I defend strongly verific (but not necessarily reliabilist) accounts of justification by claiming that there are two ways to construct a theory of justification: by analyzing our ordinary concept of justification or by taking justification to be a theoretic term defined by its role in the theory of knowledge. The former route is not promising because of the splintered nature of our ordinary concept of justification—or perhaps because there is no single such concept. On the other hand, if justification is defined by the role it plays in the theory of knowledge, then there is good reason to think that justification must be strongly truth-conductive since the term was introduced by Edmund Gettier to play the primary role in converting true belief into knowledge. And if that is right, then there will be no justified empirical belief at demon worlds. The real lesson of the demon world is then turned on its head: justification does not supervene on what one shares with one’s deceived doppelganger.

Keywords: Justification, knowledge, demon world, reliabilism, Gettier, externalism.
Introduction

The New Demon World Objection (NDWO) to reliabilist theories of justification goes as follows: consider a cognitive doppelganger of yourself who is the inhabitant of a world governed by a Cartesian demon. She has just the perceptual and introspective experience you have at the actual world, and inasmuch as is possible, she forms just the beliefs you form in our world. Now consider any given perceptual belief of yours that you take to be as well justified as any belief you have that about the external world. Your twin at the demon world has exactly the same history of experiences and beliefs you do, follows the same rules and norms you follow, and is equally confident of her parallel perceptual belief. But in her case, the belief is not only false but is unreliably formed since she is the victim of systematic deception. The intuition the proponents of the NDWO invite us to share is this: given that your belief is justified and that your twin’s experience and background system of beliefs are just like yours, then her belief must be justified too. But since her beliefs are emphatically not the result of reliable processes, then reliability can’t be relevant to the justification of your beliefs (even if they turn out to be reliably formed). In a word, reliability is not necessary for justification.¹

Although consensus is never reached in epistemology, there can be little doubt that the NDWO has been taken as decisive by the vast majority of epistemologists. Indeed, no less a reliabilist than Alvin Goldman has granted its impact and changed his view in light of it.² Michael Bergman, who argues in his recent book for a robust form of externalism, also grants

¹ The canonical version of the NDWO is found in Cohen 1984.
that beliefs formed at demon worlds are justified. It is against this strong tide that this paper attempts to swim. I will argue not only that the NDWO is not a problem for reliabilism but furthermore than that the real lesson of the NDWO is strikingly pro-externalist (although not necessarily pro-reliabilist).

Section I: What is Justification?

I begin by noting the many points of epistemic agreement that the reliabilist will doubtlessly share with the NDWO proponent. To this end, consider my belief that Eleanor, my dog, is currently lying on the floor. The belief is formed on the basis of visual perception: Eleanor is about three feet away from me, the light in the room is normal, the belief coheres well with my background beliefs, etc. My twin at the demon world will have his own belief based on a phenomenologically similar experience and imbedded in a precisely parallel doxastic system. Here are some significant points of agreement between the proponent of the NDWO and the reliabilist regarding my belief and my twin’s belief: (i) both fit our experiential evidence, (ii) both cohere with our background beliefs, (iii) both were responsibly formed and (iv) both beliefs are undefeated. There are, of course, two other points of agreement, one of which is simply part of the case: my twin’s belief is unreliably formed while mine is reliably formed; and my twin’s belief does not count as knowledge but my belief does.

The question at issue, then, is whether the salient property that my belief has that my twin’s lacks is the property of epistemic justification. But how do we decide which property that is? Prima facie, it would seem that there are two different ways to figure this out. First, we might reflect carefully on the word ‘justified’ as applied to beliefs and see if that term plausibly picks out a single property. Another way of putting what I take to be essentially the same point

---

3 Bergman 2007
is that we might attempt to analyze our ordinary concept of justification with the hope of discovering its conceptual parts. Although the first way of putting the point and the second differ slightly (the first is linguistic, while the second is conceptual, analysis), they both would root the answer to the question about the relevant property in the ‘everyday’ or non-philosophical contexts.

The second way one might decide the matter of which property (or properties) are constitutive of justification is to figure out what role justification is supposed to play in one’s epistemological theory and then determine which property in fact plays that role. This method would not be particularly beholding to the ordinary use of the term ‘justified’ or even to the analysis of our concept of justification.

What are the prospects for making progress on the question of justification at demon worlds by employing either of the above strategies? I submit that there is overwhelming reason to think that the first method will get us nowhere. A lesson to be gleaned from the analytic epistemology literature in the 1980s and early 90s with their clashing of theories of justification, and differences regarding internalism and externalism, is that there is no univocal ordinary understanding of ‘justification.’ Or if there is a univocal understanding, it is deeply incoherent or disjunctive. For it apparently implies that a belief that is responsibly held is prima facie justified, a belief that is evidentially supported by the agent’s belief system is prima facie justified, and a belief that is likely to be true is prima facie justified. Many have thought of each of the aforementioned conditions of justification that it is not only sufficient but necessary as well. Yet a belief may have any one of these epistemic virtues (being responsibly held, evidentially supported, and likely to be true) while lacking the others. We use the term ‘justified’ (or ‘rational’ or ‘reasonable’) in varied and fundamentally distinct ways. My belief is justified
when I’ve done the best I could in holding it; my belief is justified when it fits my evidence; my belief is justified when it is probably true. All of these claims get support from reflection on the ‘ordinary concept’ of justification. But such accounts of justification will give us myriad cases of beliefs that both are (because they satisfy one of the paradigms) and aren’t (because they fail to satisfy other allegedly necessary features) justified.4

If we are to abandon the project of using our ordinary language and pedestrian intuitions regarding justification as a means of developing an account of what it is to be justified, then we are left with the project of figuring out what role the concept of justification is supposed to play in our epistemological theory. And how are we to do that?

Thankfully, and perhaps surprisingly, there is a pretty straightforward answer to this question. Edmund Gettier’s 1963 paper “Is Knowledge Justified True Belief?” is justly famous for showing that a certain plausible account of knowledge is insufficient.5 What hasn’t been generally recognized is that it was Gettier who introduced not only his compelling counterexamples but also the claim that the traditional view of knowledge is that it is justified true belief. The two philosophers that Gettier cites when constructing his version of the traditional view of knowledge are A.J. Ayer6 and Roderick Chisholm.7 According to Ayer, one knows that p if and only if (i) p, (ii) one is sure that p, and (iii) one has the right to be sure that p. Chisholm, on the other hand, says that S knows that P iff (i) p, (ii) S accepts that p, and (iii) S has adequate evidence that p. So what Gettier identifies as the justification condition can be reasonably described as whatever it is that Chisholm had in mind when he said that knowledge

4 The point that there is more than one concept of epistemic justification goes back at least as far as Alston 1985. See also Alston 2005.
5 Gettier 1963
6 Ayer 1956
7 Chisholm 1959
requires ‘adequate evidence’ and what Ayer identified as ‘the right to be sure.’ But we shall be led astray if we ignore the fact that the justification condition was, according to Gettier, thought by both Ayer and Chisholm to be sufficient with true belief to generate knowledge. If Gettier was right, the assumption of Ayer and Chisholm was that once you had true belief all you needed for knowledge was the right to be sure or adequate evidence for the proposition in question. So, in short, Gettier introduced the term ‘justified’ in contemporary epistemology as a theoretical term: a belief is justified if and only if it would lead to knowledge in typical cases where it was true.  

Yet, alas, cases aren’t always typical; sometimes justified true belief fails to count as knowledge. Smith might have overwhelmingly good evidence that Jones will get the job, know that Jones has ten coins in his pocket and thus conclude that the person who will get the job has ten coins in his pocket. But if the fates are epistemically cruel yet vocationally kind, it will turn out that although Smith really does have objectively strong evidence for both of these propositions, the former belief is false—i.e., Jones won’t get the job. For, as luck would have it, Smith—who also has ten coins in his pocket—will get the job that they both seek and so his belief that the person who will get the job has ten coins in his pocket is justified and true but it is not an instance of knowledge. Or maybe Smith has great evidence for thinking that Jones owns a Ford and a knack for making as many valid inferences from his well-justified beliefs as he can. So he reasons that since Smith owns a Ford, then the proposition that “Smith owns a Ford or Brown is in Barcelona” must be true, even though he has no evidence at all regarding Brown’s

---

8 Jack Lyons has suggested that what I really have in mind is that “justification” is introduced by Gettier as a “theoretically-tweaked version of a pre-theoretical term” rather than a “purely theoretical term.” While I think this is right, the phrase Lyons suggests is too unwieldy to be useful, so in the context of this paper please understand “theoretical term” in the way Lyons suggests.
location. Again, the epistemic fates conspire against Smith and Jones does not own a Ford but Brown turns up in Barcelona; Smith has a justified true belief but fails to know.

The assumption that was clear in the Gettier examples (and indeed in the large literature that they spawned) but which I fear has been lost in the meantime is that the subjects in those cases were very close to having knowledge. In the overwhelming majority of cases, if (a) you have the very good inductive evidence of the sort that Smith has that Jones will get the promotion, (b) you also know that Smith has ten coins in his pocket, and (c) you infer from these premises that the person who will get the promotion has ten coins in his pocket, you’ll not only end up with a true belief but with knowledge too. The same goes when you begin with a belief that p that is grounded in very good evidence and use disjunction introduction to infer (p v q). Of course, there are many more familiar cases from the Gettier literature (think of Nogott, Grabbitt, etc.). I submit that all of these are cases in which the subject fails to have knowledge but in which he “missed it by that much” (as Maxwell Smart would say).

**Section II: Justification and Truth**

What does it mean to say that a Gettiered justified true belief is very close to being knowledge? Intuitively, it means that beliefs that are true and justified in the way that the particular belief is justified are usually (to put the point rather too weakly) instances of knowledge. What does this tell us about justification and the role it plays in knowledge? I think there are two lessons here. The first is admittedly contentious but I think the second point should be granted on all sides

Here, then, is the contentious claim:

[A] Necessarily, the great majority of justified beliefs are true.
To repeat: I’m understanding justification as a theoretic term introduced by Gettier; [A] is not a claim about our ordinary, pre-theoretical concept of justification (if there is a single such concept, which I doubt). The sort of justification enjoyed by the subjects in the Gettier cases was very robust; indeed, it was so strong that in the absence of the Gettierized circumstance, the beliefs would *clearly* count as knowledge. If a belief justified in a certain way usually turns out to be an instance of knowledge, then that justification is objectively strong; beliefs grounded in the sort of evidence Smith has in the Gettier cases will generally be true. And this is surely crucial to the force of the examples and hence is an essential feature of the property that Gettier was intending to isolate. Had he picked a case in which the truth connection was absent or greatly diminished, one would have been able to deny that the justification condition was indeed satisfied.

Consider, for example, this case: Ethelene is twenty years old, has been raised in a fundamentalist community, and has had little access to the outside world. She learns in her science class that God has promised never to destroy (pretty much all of) humanity with water again; God will use fire the next time should he be so provoked. Ethelene’s science teacher reasons that since humanity is more sinful than ever, God will destroy humanity again and, as promised, he’ll do it with fire. But since God’s active wrath tends to be preceded by a warning, God will cause the earth’s temperature to increase as a sign of what is to come. Based solely on this inference, Ethelene’s science teacher tells her class that the temperature of the planet is increasing and that unless humanity changes its ways, this will continue. Ethelene believes her science teacher and what he says is indeed true: climate change is happening, the earth’s temperature is increasing, and unless humanity changes its dependence of fossil fuels, this will continue. So Ethelene has a true belief that is *in a sense* justified. Would this have been a
convincing Gettier case? I submit that it would not. For in response, one could have plausibly objected that Ethelene’s sorry epistemic situation regarding science claims keeps her from having *objectively* good reason for beliefs of this sort. And if she’s not in a good position to get at the truth, then she’s not in a good position to have justified beliefs in the sense that is fundamental to the theory of knowledge. Notice that such an objection to the original Gettier cases would not have worked. The subjects in those cases had objectively good reason for their true beliefs that nevertheless failed to count as knowledge.

Furthermore, if most justified true beliefs turn out to be knowledge, then the justification must have a verific component to it; it must be that beliefs formed that way are in most cases true. It is important to note that [A] does *not* commit one to reliabilism. For not only is [A] is not an analysis of our ordinary concept of justification it is also not an account of the *nature* of the theoretic concept. [A] implies only that the kind of justification that is enjoyed in the Gettier cases is necessarily connected to truth. Recognizing this point is consistent with any one of many different theories of the nature of justification. One could be an evidentialist, say, and think that justification should be analyzed in terms of strength of evidence; the only constraint that [A] would impose is that the evidence necessary for justification (when “justification” is construed as a theoretical term introduced by Gettier) must carry with it the implication that beliefs believed on the basis of evidence *like that* are generally true. But I don’t see any reason for the evidentialist to resist this. For if justification is the positive epistemic component of knowledge, then it must be hooked up with truth in some pretty straightforward way. And evidentialism has the theoretical means to acknowledge this. What’s true of evidentialism is true for any plausible account of justification. Indeed, it is a condition of adequacy on such accounts that they be capable of explaining why it is [A] is true.
As I said when I introduced it, I grant that [A] is contentious. Still, I’m not sure how it can be reasonably resisted. The Gettier examples simply wouldn’t have the force they clearly have if justification in such cases didn’t generally lead to true beliefs; and since the term is defined by its starring role in the concept of knowledge, this verific component is essential to the concept.

Another way of seeing why [A] has to be right is its relation to a second claim that will be defended in the next section of the paper. Again, in the original Gettier examples, Smith’s justified true belief fails to be an instance of knowledge because of highly unusual, knowledge-defeating circumstances. But that means in non-Gettier scenarios, true beliefs that are justified in the sense introduced by Gettier are instances of knowledge. In other words (and here comes the second, less contentious claim):

[B] Necessarily, the vast majority of justified true beliefs are instances of knowledge.

With the caveat that justification is understood as a theoretical term introduced by Getter to refer to the property that takes us very close to knowledge, I don’t see how [B] can be denied (more on this presently). Furthermore, the truth of [A] provides a nice explanation of the truth of [B]. Cases of justified true belief are typically cases of knowledge because necessarily, justified beliefs are true. If there were not a strong connection between justification and truth, then we shouldn’t expect justified true beliefs to typically be instances of knowledge. And if the truth connection were infallible, then there couldn’t be Gettier examples. So what [B] seems to require is a strong but fallible truth connection and that is precisely what [A] asserts.

It is a straightforward consequence of [A] that there are no justified beliefs at demon worlds. For there is no property that a subset of my doppelganger’s beliefs have which is such that, necessarily beliefs that exemplify that property are true. So in any world in which there is
the radical split between belief and truth that one finds at demon worlds, there will be no justified belief.

Still one might not be convinced that even if necessarily, justified true beliefs are knowledge, then justified beliefs are generally true. So let’s drop that and move on to what I think should not be contentious—viz., the claim that justified true beliefs are instances of knowledge.

Section III: Justified True Belief and Knowledge

[B] is false only if the justification condition as construed by Gettier is consistent with justified true beliefs failing to be instances of knowledge. Notice that if the beliefs in the Gettier examples were justified in only a weak sense and were true, the cases would fail to do what Gettier wanted them to do. As we saw above, this is because if the justification condition isn’t sufficiently robust, then the defender of the justified-true-belief view of knowledge could claim that they aren’t counterexamples because the justification condition hasn’t been sufficiently satisfied. Gettier’s genius is in coming up with examples where none but the skeptic would claim that beliefs justified to the degree and in the way that these beliefs are justified would fail to be knowledge if true in typical circumstances. So beliefs in Gettier cases are justified in such a way that true beliefs formed in that way count as knowledge in typical circumstances—i.e., in the great majority of actual and counterfactual circumstances true beliefs that are formed in the same way or under the same justifying conditions as those beliefs count as knowledge.

With all of this behind us, the argument that beliefs at demon worlds aren’t justified can be presented rather quickly. Subjects at standard demon worlds have only false external world beliefs. But more than that, the beliefs of the victims of the demon world are also false at all the
worlds surrounding the world they inhabit. For the demon is cunning and powerful at those worlds too; so my doppelganger’s beliefs are not only actually false (where actuality is indexed to his world) but generally counterfactually false as well. But recall that given the way Gettier introduced the term ‘justification,’ a justified true belief is Gettiered iff it is knowledge at most nearby worlds. So to test to see whether my doppelganger’s belief would count as knowledge at the closest worlds at which it is true, we’ll have to ignore the fact that at the great, great majority of nearby worlds the belief is false and so hence not knowledge. We’ll need to find worlds in which the demon screws up and accidentally lets the very rare true belief through. At such worlds, the subject would have a true belief but it would not count as knowledge. And as long as the relevant worlds are all demon worlds, then in none of them would the subject know. Why would a true belief not count as knowledge at demon worlds? Because it is purely a lucky accident that the belief is true. Recall Carl Ginet’s well-known barn façade case. The subject doesn’t know because even though this particular belief that he’s seeing a barn turns out to be true, it is entirely a matter of luck that this is the case; in the great majority of relevant cases, if he forms a similar belief in similar circumstances, his belief will be false. Things are very much the same for the demon on those rare occasions when a true belief sneaks through—in fact, given the power and cunning of the demon, such lucky cases will be at most very rare and likely only at worlds a fair distance from the world my doppelganger inhabits. If the subject of the barn façade cases lacks knowledge, then all the more so does the lucky doppelganger.

So whereas Gettier examples are cases in which the subject is very close to knowledge (i.e., in the great majority of situations where the subject has that kind of justified true belief, the subject also has knowledge and in the nearest worlds, the great majority of beliefs that are true and justified in that way are instances of knowledge), there are no cases in which the subject of a
demon world has a true belief that is justified in such a way that in nearby worlds, the majority of beliefs justified in that way are cases in which the person has knowledge. This is so because there are no nearby worlds at which the subject as knowledge; and even at more remote worlds where the subject has true belief, that belief fails to count as knowledge.

This means that demon world examples are not Gettier cases. So the failure of the subjects to know in instances of accidental true belief at demon worlds can’t be a matter of those beliefs failing the fourth, Gettier-defeating condition of knowledge. But obviously, the belief and truth conditions of knowledge are satisfied. What’s missing is the satisfaction of the justification condition. Beliefs formed at demon worlds aren’t justified because beliefs formed via demonic deception cannot be instances of knowledge should they be true.

Here’s the main line of the argument I’ve been giving:

1. ‘Justification’ in epistemology is a theoretical term introduced by Gettier to mean that which with true belief was traditionally thought to be sufficient for knowledge (and which with the addition of a fourth condition is sufficient).

2. Any true belief that is justified in the above sense will be an instance of knowledge in the great majority of actual and counterfactual cases.

3. No true belief had at a demon world will be an instance of knowledge in the great majority of actual and counterfactual cases.

C1. Therefore no true belief is justified at a demon world.

4. The accidental truth of a belief is irrelevant to its justificatory status.

C2. Therefore, no belief is justified at a demon world.

So there are no Gettier cases at demon worlds—and that this is so is not because there are no true beliefs at demon worlds. There can be true beliefs even in the midst of general, mass
deception; but even when such beliefs are true (and rational), true beliefs justified that way are not knowledge in the closest worlds where those beliefs are true (since all such worlds have demons).

At the risk of redundancy, let me put the same point slightly differently. Since the term ‘justified’ was introduced by Gettier as a theoretical term to mean “that which makes true belief knowledge in non-Gettier conditions,” and since this kind of justification, when combined with true belief gets you knowledge in most actual and counterfactual situations, it follows that beliefs are justified at demon worlds only if those beliefs count as knowledge in the nearest worlds at which they are true. Yet in those nearest worlds, the subject’s true beliefs don’t count as knowledge since they are in effect extreme barn façade-style cases. Since those beliefs don’t count as knowledge in any of the nearest worlds at which they are true, these aren’t justified in Gettier’s sense.

I have, then, presented two arguments for the conclusion that there is no justified belief at demon worlds. First, in the Gettier examples, beliefs formed in the way that the subject’s Gettiered beliefs are formed (and in the typical circumstances) will generally be true. So a belief that is justified in Gettier’s sense of the term will be true in the great majority of cases. But no

---

9 Putting the point this way is actually a gift to my opponent: for as mentioned earlier, the way beliefs in Gettier’s examples are formed, they will be true and count as knowledge in the great majority of actual situations and the nearest worlds; neither of these conditions can hold for beliefs at demon worlds. The best she can hope for is that the nearest worlds where the belief is true—and these are not very near worlds at all—will be mostly be worlds where the subject knows. For the sake of this section, I’m supposing that satisfying that condition would be sufficient for there to be Gettier cases at demon worlds.

10 Calling cases of true belief at demon worlds “extreme barn façade cases” is accurate but potentially misleading. It’s accurate because there is a deep similarity between them: in each case, it is just a happy and unlikely accident that the belief turns out to be true. It’s potentially misleading because barn façade cases are Gettier examples while demon world cases aren’t. In the barn cases, the subject is justified (in Gettier’s sense) in believing that she sees a barn; her visual experience provides her with the kind of reason that will generally produce knowledge. But, as we’ve seen, in the demon world cases, the subject doesn’t have that kind of reason for any empirical belief and so her belief isn’t justified.
belief at a demon world has the sort of justification that requires truth in the great majority of cases, so no belief at a demon world is justified. (As I noted earlier, accepting this point doesn’t commit one to reliabilism but only some kind of robust truth connection; one can even be an evidentialist and accept this.) My second, and stronger, argument hinges on the premise that, necessarily, true beliefs justified in the way Gettier example beliefs are justified are instances of knowledge in the great majority of cases (and in nearby worlds). But at the closest demon worlds where there is true belief, such belief is never knowledge so it isn’t justified in the sense that beliefs are justified in Gettier cases.

Section IV: The Real Lesson of the Demon World

Even if demon worlds are void of justified empirical belief, they may still contain other positive epistemic qualities. Deceived inhabitants can still be intellectually responsible, guided by epistemic norms, have subjectively good evidence, coherent belief systems, rational foundationalist noetic structures, properly functioning cognitive processes, and can even exhibit conditional reliability. But despite this plethora of epistemic goods, justification is lacking in worlds of mass deception. The reason this point has not been widely recognized, it seems to me, is that when giving accounts of justification, epistemologists have tended to reflect on the ordinary meaning of the term ‘justified’ or concept of justification. But given the flexible and perhaps inchoate nature of that term and concept, the results of such analyses are predictably diverse and incompatible. In my view, the difficulties with this methodology are two: there is no way to adjudicate between competing accounts since our common concept/ordinary understanding is multifarious and (more importantly) the relation of the ordinary concept of justification to knowledge is vague at best. If one wants justification to play a principled part in
knowledge, one will have to take Gettier’s tack and define it as the theoretical concept that plays a specific role in knowledge.

Once one understands justification in this light, the importance of demon world cases is turned upside down. For the lack of justification at demon worlds carries with it some important implications for the theory of justification. First, reliabilism can’t be ruled out on the ground that it denies justification at demon worlds; this is a virtue and not a vice of the theory. Second, there is a stronger truth-connection constraint on theories of justification than most have been willing to grant. As I’ve been at pains to point out, one needn’t think that reliabilism is true in order to honor the truth connection, but one will have to construct one’s theory in such a way that makes the connection between justification, truth, and knowledge quite robust. Third, the real lesson of demon worlds for the theory of justification is that justification doesn’t supervene on the cognitive states that one shares with one’s twin at a demon world. No matter how you beef up the internal states of subjects, as long as all the relevant properties could be had by a deceived doppelganger, you don’t have a condition that is sufficient for even *prima facie* justification. Justification has a strongly external component.

**Conclusion**

In this paper, I’ve argued for two related although logically independent claims. If either is true, then there are no justified beliefs at demon worlds and epistemic justification does not supervene on what we share with our deceived doppelgangers. Furthermore, as long as reliabilism is not construed as a theory of our ordinary concept of justification but rather as an
account of that which, together with truth, belief and an anti-Gettier fourth condition is sufficient for knowledge, the demon world objection is misguided.\textsuperscript{11}

\footnote{Thanks to Todd Long, Jack Lyons, Matt Mullins, and Andrew Arnold for comments.}
Bibliography


