

Epistemic Value and the New Evil Demon

B.J.C. Madison

(Forthcoming in *Pacific Philosophical Quarterly*)

Draft Version – Do Not Cite Without Approval

Abstract: In this paper I argue that the value of epistemic justification cannot be adequately explained as being instrumental to truth. I intend to show that false belief, which is no means to truth, can nevertheless still be of epistemic value. This in turn will make a good prima facie case that justification is valuable for its own sake. If this is right, we will have also found reason to think that truth value monism is false: assuming that true belief does have value, there is more of final epistemic value than mere true belief.

1) Introduction:

Epistemic justification is valuable; it is a good that warrants our pursuing. But what explains the value of justification? Recent tradition has it that justification's value is to be understood purely *instrumentally*: while justification is a good, it is only good as a means to something else which has final value, such as true belief.

In this paper I will argue that the value of epistemic justification cannot be adequately explained as being instrumental to truth. I intend to show that false belief, which is no means to truth, can nevertheless still be of epistemic value. This in turn will make a good prima facie case that justification is valuable for its own sake. If this is right, we will have also found reason to think that truth value monism is false: assuming that true belief does have value, there is more of final epistemic value than mere true belief.

2) Epistemic Justification: Of Mere Instrumental Value?

My belief that there are now fewer than five hundred people before me is *justified*; if, on the other hand, I were to now believe, sitting here in Cape Town, that there are now forty-five planes on the runways of Heathrow, that belief would be *unjustified*, even if by some chance that belief turned out to be true. The kind of justification at issue in these examples is *epistemic*, rather than say moral or pragmatic. Whereas moral justification concerns, for example, right actions, and pragmatic justification concerns the usefulness of perhaps beliefs as well as actions, it is widely thought that what individuates epistemic justification from other sorts is some kind of connection to truthⁱ. What exactly the connection is between truth and justified belief, however, is far from clear.

In addition to exhibiting a truth-connection, epistemic justification (or ‘justification’ for short, unless I indicate otherwise) is a valuable property for a belief to have. Justified beliefs are a good -- they are better than unjustified ones, from an epistemic perspective. If I were to hold a belief about the number of people currently in the room, in at least one clear sense, it would be *better* to believe that there are fewer than five hundred people, rather than believe that there are. Reflecting on our epistemic practices reveals that we often consciously strive to acquire justified beliefs. In addition, we criticize ourselves and others for holding unjustified beliefs, or for lacking sufficient justification for believing as we do, even if there are no practical or moral implications for believing without justification. These considerations suggest that, from an epistemic point of view, if one holds a belief, it is always better if that belief is only held *if* justified.

So, on the face of it, justified beliefs are valuable, from an epistemic point of view, independent of any moral or practical concerns. But to say that justified beliefs are of value is not to commit to holding that the value of justification always, or even sometimes, outweighs other considerations, all things considered. The kind of value justified beliefs seem to have is *pro tanto*: they have value, but their value is not a decisive reason to pursue or hold them -- sometimes it might be best overall if an unjustified belief is held, if for example doing so was morally required of one. Or to take another example, it might be best overall to hold an unjustified belief when there is something of great practical value at stake. For an extreme illustration of this, consider William James' case of a climber who would not be able to summon the strength required to save his life and leap a vast chasm, unless he believed that he was able to make the jump. It seems that, all things considered, the climber ought to believe that he can make the jump, even if that belief is epistemically unjustified. So it is consistent with holding that epistemic justification is of value, even of great value in our intellectual lives, without being implausibly committed to the idea that the value of justification need always trump other things of value.

Now, it is one thing to claim that epistemic justification is of value, but it is of course quite another to ask how the value of epistemic justification is to be explained. As Laurence Bonjour asks, 'Why should we, as cognitive beings, *care* whether our beliefs are epistemically justified? Why is such justification something to be sought and valued?' (Bonjour, 1985, p.7) Bonjour thinks that the answer to these questions is obvious. He writes,

Once the question is posed this way, the following answer seems obviously correct, at least in first approximation. What makes us cognitive beings at all is our capacity for belief, and the goal of our distinctively cognitive endeavors is *truth*: We want our belief to correctly and accurately depict the world. [...] The basic role of justification is that of a *means* to truth, a more directly attainable

mediating link between our subjective starting point and our objective goal. [...] If epistemic justification were not conducive to truth in this way, if finding epistemically justified belief did not substantially increase the likelihood of finding true ones, then epistemic justification would be irrelevant to our main cognitive goal and of dubious worth. It is only if we have some reason for thinking that epistemic justification constitutes a path to truth that we as cognitive beings have any motive for preferring epistemically justified beliefs to epistemically unjustified ones. Epistemic justification is therefore in the final analysis only an instrumental value, not an intrinsic one. (Bonjour, 1985, pp. 7-8; see also p. 157 for another clear expression of this position)

As is made clear above, Bonjour thinks that the value of justification is instrumental to the end of truth. However, upon reflection, we will see that this is false: the value of justification is not exhausted by its being instrumentally valuable in securing truth.

In addition, in the central thought experiment I shall consider, since there are no other plausible ends present that justification could be instrumental to (e.g. such as knowledge), this then will give us *prima facie* reason to think justification is of final value. If this is correct, we will also find that there is good reason to think that a closely related influential view about epistemic value is false. What has been variously labeled ‘Veritism’ⁱⁱ, ‘Epistemic Value T-Monism’ⁱⁱⁱ, or just ‘Epistemic Value Monism’^{iv}, holds in common that truth is the only fundamental epistemic good.^v So while on these accounts there may be epistemic goods other than truth, these things are only *instrumentally* good, relative to the end of truth. However, if my central argument is sound, this will make a good *prima facie* case that epistemic-value monism is false: there is more of fundamental epistemic value than mere true belief.^{viii}

3) The New Evil Demon Problem(s):

Descartes' evil demon hypothesis can be used as the basis for a classic skeptical argument. The familiar argument, of course, begins as follows: it seems possible that there could be an evil demon of great power, so great, that he could ensure that our beliefs about the external world are false, based on non-veridical perceptual experiences. The challenge then is this: how do we know that we are not the victims of such an evil demon? The further claim is that if we cannot rule out the possibility of such a demon, then we lack knowledge of the external world. We can call this skeptical challenge the Old Evil Demon Problem.

But consider another form of argument, which has come to be known as the New Evil Demon Problem. We are asked to once again consider a possible world inhabited by an evil demon of great power, so great that he could ensure that the subjects of that world have beliefs about the external world that are false, based on non-veridical perceptual experiences. A question that can now be asked is not a skeptical one, but rather, the following: do the subjects of that demon world have *justified* beliefs? If so, are their beliefs justified to the same extent as their counterparts in a non-demon world?

All parties ought to agree that the victims of the evil demon do not have *knowledge*, since their beliefs are false. But still, many will judge that the subjects hold beliefs that are epistemically justified, and justified to the same extent as their counterparts in the actual world we inhabit^{viii}. From this initial evaluative judgment, different lessons have been drawn.

From the judgment that justification could still be present in a demon world, some epistemic internalists have then argued by inference to the best explanation that

justification therefore must be an internal matter: factors external to the subject's conscious awareness, for example, such as the reliability of the process that gave rise to the belief, are not necessary for its justification^{ix}.

Independently of such a positive conclusion, a more modest negative conclusion has often been drawn: as the case was originally presented, it was offered as an argument against process reliabilism (the view that a belief is justified IFF it is the product of a reliable belief-forming mechanism^x). The New Evil Demon problem has been offered as a counterexample to the reliabilist's claim that reliable belief formation is a necessary component of epistemic justification.

My central aim in this paper is draw a different moral from the New Evil Demon case, and to show the significance of this case for how we ought to think about epistemic value. Before doing so, however, it is worth briefly saying a few words about the judgment that victims of the evil demon enjoy beliefs that are epistemically justified, and justified to the same extent as their counterparts in a non-demon world. As presented, this is an intuitive judgment about a possible case. If this judgment is to have probative value, does it require further *argumentative* support?

It is difficult to say. Consider Gettier cases, or thought experiments like the trolley problem offered in support of consequentialism, or those offered in debates about personal identity that ask if a person could survive a brain transplant. In each instance, a possible case is offered: judgments are then made about the case. Usually it is not thought that we need to *argue* in support of the judgment that a subject lacks knowledge in a Gettier case, or that the morally right action in the trolley case is divert the trolley and thereby kill one person to save ten for example, or that a person

could possibly survive a brain transplant. We might go to *explain* the judgment, or see what follows from it, to see if the consequences are acceptable, for instance. But other than by offering more cases, it is hard to see how one might go about *arguing* for the truth of such fundamental judgments. My hypothesis is that the judgment of the presence and sameness of justification between subjects and their demon-world counterparts is likewise fundamental. But that the judgment has this fundamental status need not be a bad thing.

This is because it is important to see that our claim to sameness of justification is an *evaluative* thesis, not an *explanatory thesis*.^{xi} It says only that twins internally alike are justificationaly alike; not *why* or in virtue of what are they so justified. As Nico Silins has pointed out, this is a virtue of evaluative theses in general, since what supports them is pure intuitive plausibility, not motivations from theoretical judgments about the nature of justification. (Silins, 2005, p. 386) As relatively pre-theoretical judgments, they can be used as data in philosophical theory building, as I intend to do here. Such data should be used unless we have good reason to reject it.

Nevertheless, for those who might not yet feel the pull of the New Evil Demon intuition, I shall offer three sets of considerations in its favor: that the victims of the New Evil Demon are rational and reasonable in believing as they do, and this implies that they are justified; that they believe what they ought to, given how things seem to them, which in turn entails that they are justified; and finally, that the New Evil Demon victims can be epistemically praised and blamed in the same way as their non-demon world counterparts is best explained by the fact that like us, their beliefs are justified.

First, as Cohen noted when originally offering the New Evil Demon case, the victims in the demon world seem to be perfectly well rational and reasonable in believing as they do, given the grounds on which they hold their beliefs. Cohen asserts that ‘reasonable’ and ‘rational’ are ‘virtual synonyms’ for ‘justified’, so that if one concedes that the subjects are rational and reasonable, one is also thereby conceding they are justified too^{xiii}. (Cohen, 1984, p. 283) And we should say that the subjects in the demon world are rational and reasonable in holding the beliefs they do, given how, or on what basis, they believe as they do. Their beliefs are not *groundless*; rather, they are based on reasons: e.g. perceptual beliefs are held on the basis of having certain perceptual experiences; inferential beliefs are arrived at through the same rules of reasoning (deductive; inductive; abductive) that their non-demon world counterparts use. If our beliefs are reasonable and rational if formed in these ways and on the basis of these grounds, then so are the beliefs of our counterparts who form their beliefs in the same ways. If the relevant beliefs are reasonable and rational, then this is at the very least strong evidence that they too are justified.

A second consideration in favour of the New Evil Demon intuition arises from reflecting on the connections between the notions of justification and epistemic oughts. Take my counterpart who has recently been deceived by a demon. Since I believe that I am typing on a computer, so does he. Now we can ask: what attitude *should* my counterpart take towards the proposition that he is typing on a computer? Given that it seems to him like he is typing on a computer, and given that there is nothing he is aware of to suggest he is mistaken about this, then he should believe that he is typing on a computer (there is not even anything he *should* be aware of, given his epistemic position; it is not that he is ignoring counter-evidence, for example).

Since my counterpart believes what he ought to believe, it seems correct to say that he is justified in so believing, just as I am now justified in believing that I am typing at a computer, since that is what I ought to believe, given how things currently seem to me.

If someone were to resist this line of thought and deny that my counterpart ought to believe as he does, what would the alternative be? Should he *disbelieve* that he is typing at a computer? Should he *suspend judgment* about whether or not he is typing at a computer? Surely not: given how things seem to the subject, it would be quite improper to disbelieve or suspend judgment that things are as they seem. Doing so would find the subject in a quasi-Moorean absurd situation of thinking or uttering things like the following:

DENIAL: it seems to me that P, and I have no reason to doubt that P, but I believe that not-P;
or

SUSPENSION: it seems to me that P, and I have no reason to doubt that P, but I shall suspend belief as to whether or not P.

The absurdity of DENIAL and SUSPENSION is reason to hold that a subject should obey something like the following epistemic norm: if it seems to S that P, then S should believe that P (in the absence of defeaters)^{xiii}. The mere fact that P is false is not sufficient for it to be a defeater for one's justification for believing that P, since that would have the implausible consequence that no one can ever be justified in believing falsehoods^{xiv}. Like us, victims of the New Evil Demon believe what they ought to believe, given how things seem to them. That is to say, that they, like us, have justification for believing as they do.

I shall offer a third and final consideration in support of the intuition that New Evil Demon subjects have justified beliefs, and share sameness of justification with their non-demon world counterparts. It has been widely noted that victims of the New Evil Demon are *blameless* in believing as they do. In fact, that they are blameless has been offered as an error theory by some of those who deny that New Evil Demon victims have justification for their beliefs^{xv}. Their basic line is this: granted, subjects systematically deceived by an evil demon are *blameless* in believing as they do, but that is not to say that they have *justification* for believing as they do. Further, blamelessness is not justifiedness, and although they may often go together, we ought not confuse one for the other. So while there may be *something* epistemically positive going for the subjects and their beliefs in the demon world, that something is mere blamelessness, not justification.

Let us grant that justification and blameless are not identical properties, for at least this reason: being blameless is insufficient for being justified^{xvi}. Still, it does seem to be the case that justification and blamelessness are very often co-present: whenever a subject is justified in believing as they do, they are also epistemically blameless. Since not property identity, what best explains the co-presence of justification and blamelessness?

My suggestion is that what best explains blamelessness in a demon world is that which justifies those subjects in believing as they do. By reflecting on clear cases of epistemic justification, we can see that blamelessness is a pre-condition for justification, since if a subject is epistemically blameworthy in believing as he does, those conditions defeat the justification he would have otherwise enjoyed, had the subject not been epistemically blameworthy.

For example, suppose a subject S believes, upon reading this information in a departmental memo produced by the usually highly reliable departmental chair, that he is invigilating two exams that week. The subject would be justified in believing this proposition on that basis. Suppose further that the departmental secretary subsequently sends out a memo which contradicts the first: it says that S is to invigilate *three* exams that week. But what if because plans were already made, or because it was otherwise inconvenient to invigilate the extra exam, S wantonly disregards the second memo, and continues believing that he is invigilating two exams that week, and believes it as strongly as he did having only read the first memo?

Such a belief would now be unjustified, and the explanation of this fact is obvious: by disregarding the conflicting evidence for the reasons he does, S is epistemically irresponsible. Such irresponsibility is epistemically blameworthy – S's believing can rightly be criticized on this basis. The grounds of the blameworthiness defeats the justification that S would have otherwise had; remove the grounds of the blameworthiness, holding all else fixed, and justification is restored. Being epistemically blameless, therefore, is a necessary pre-condition for believing with justification.

If this is correct, we also have reason to think that what explains epistemic blamelessness can in part also explain justification: the grounds and conditions that make a belief justified are also the ones that make the belief epistemically blameless. In our example above, without defeaters present, S is justified in believing that he is invigilating two exams on the basis of reading the first memo. Holding this belief on the basis of those grounds is also epistemically blameless.

If epistemic blamelessness enables epistemic justification, then of course the mere presence of blamelessness does not entail the presence of justification. Nevertheless, if we ask what explains the blameless believing in the demon world, it seems that its best explanation is that the subjects are justified in believing as they do.

In short, the thought is this: since New Evil Demon victims share epistemic blamelessness with their non-demon world counterparts in the ways that they do, then this is some reason to think that they share epistemic justification: the very factors that account for epistemic blamelessness also make it the case that epistemic justification is present too. Combine this with the facts that the New Evil Demon victims' beliefs are reasonable and rational, and that they are believing as they ought, given how things seem to them: these three considerations offer support in favor of the evaluative judgment that New Evil Demon victims share sameness of justification with their non-demon world counterparts.

While more might be said in defence of the New Evil Demon intuition, hopefully a cumulative case has been made sufficient for the purposes at hand: our counterparts in the demon world share sameness of justification with us. But I want to now draw a different and new moral from the case of the New Evil Demon.

Consider once again the possible world being inhabited by an evil demon of great power, so great that, for example, he could ensure that the subjects of that world have beliefs about the external world that are false, based on non-veridical perceptual experiences. Again, setting aside the skeptical worry, we can ask the following, in addition to the question of the presence of epistemic justification: do the subjects of that demon world have beliefs that are of *epistemic value*? If so, is this value of the same sort that our beliefs enjoy in the actual, non-demon world?

I make the evaluative judgment that they do: despite being victims of an evil demon, they nevertheless have beliefs that are epistemically valuable, and there is a sameness of value that is shared with their counterparts in the actual world. The justified perceptual belief of the demon victim who believes that there is a candle before him because that is what he seems to see is *good*, from an epistemic point of view; likewise, the subject who justifiably infers further beliefs through valid rules of inference likewise holds beliefs which are epistemically *good*.

That these beliefs are of epistemic value is an intuitive judgment, like the judgment that justification is present. As I have been emphasizing above, there is an important difference between evaluative and explanatory hypotheses, and explaining the source of this value is a separate question from the judgment that it is present. But it seems that some of the promising possible ways of explaining the source of this value include noting that, like us, the victims of the evil demon base their perceptual beliefs upon their perceptual experiences. Also, like the good reasoners among us in the actual world, they do not make hasty generalizations and they obey the canons of good inductive inference. Our counterparts in the demon world are no more confident of their beliefs than their evidence warrants^{xvii}. If our beliefs cohere with each other, so do the beliefs of our demon-deceived counterparts. These are valuable properties for beliefs to have: both for us, as well as for our demon-deceived counterparts. So it seems that if we can have justified beliefs of epistemic value, so can our demon-deceived counterparts.

To make this central point clear, it is also worth considering a variation of Cohen's 1984 New Evil Demon case. In that case he compared two subjects, both of whom were in the demon world, as opposed to comparing a subject in the actual world with his demon-victim counterpart. Cohen's pair of subjects seemed

epistemically quite different, despite both having unreliably produced false beliefs: the first, A, was a good reasoner, in that he proportioned his belief to the evidence, he did not make hasty irrational generalizations, he was sensitive to defeaters, and so on; the other subject, B, did none of these things. B failed to carefully survey evidence, jumped to conclusions, and was generally epistemically irresponsible. Cohen concluded, ‘There is a fundamental epistemic difference between the beliefs of A and the beliefs of B. But the Reliabilist does not have the theoretical means to display this difference. I would claim that the distinction between the beliefs of A and B is marked precisely by the concept of justified belief.’ (Cohen, 1984, p. 283) But now, in addition to asking whether A’s beliefs differ in justification from B’s, we can ask the following: do any of the subjects of that demon world have beliefs that are of *epistemic value*? Do all the subjects of the demon world have beliefs that are of the *same* epistemic value? As to the first question: it seems so - I contend that A’s beliefs are epistemically good, whereas B’s are not. This is so, even though both subjects hold false beliefs (the demon sees to that). As to the second question: it seems not – A’s beliefs are epistemically good, whereas B’s are not.

Reflecting on these cases we can see that demon-deceived subjects can have beliefs of epistemic value, despite the fact the demon systematically ensures that the beliefs held in the demon world are false. Given that this is so, the value of justification is not exhausted by being instrumental to truth. But besides truth, is there any other plausible candidate that could explain the allegedly purely instrumental value of justification that can nevertheless be present in a demon world?

Perhaps one might initially wonder if justification is of purely instrumentally value as a means to *knowledge*, until one sees that most knowledge is also impossible in the demon world, as are all other factive states – the demon sees to that by ensuring

the falsity of what is believed. If justification can obtain in the demon world and is of value, and there are no reasonable candidate goods that could explain justification's instrumental value, then this makes a good prima facie case that justification is valuable for its own sake.

Indeed, I suggest that what this variation of the New Evil Demon problem points to is some fundamental epistemic value apart from mere truth. Whether this value stems from responsiveness to evidence, from epistemic responsibility, from epistemic blamelessness, from epistemic rationality, or from some other source, remains an open question that requires further investigation^{xviii}. But to be sure, the value of epistemic justification is not exhausted by its instrumental value, which in turn suggests that epistemic-value monism is false: we have reason to think it false that the only thing of fundamental epistemic value is true belief. False belief can still be of epistemic value, even though it falls completely short of reaching the goal of true belief^{xix}.

The argument in favor of the hypothesis that justification is valuable for its own sake relies on a version of the New Evil Demon thought experiment. In response, therefore, it is worth looking at a very popular family of responses to the original New Evil Demon when that problem is offered as a challenge for reliabilism, and seeing if they can be used to maintain the account of the value of justification as instrumental to true belief. What these approaches have in common is the basic move of holding that justification needs to be relativized to worlds: while the victims of the demon world hold beliefs that are unreliably produced *in their* world, we correctly attribute justification to them since they are using belief-forming methods that are reliable in *our world*. According to these approaches, the victims of the evil demon are justified in believing as they do since they are using belief-forming methods that

would result in mostly true belief, had they not been the victims of the demon.

Following Juan Comesaña (2002), we can call this kind of approach Indexical Reliabilism.

Whether or not Indexical Reliabilism works as an account of how we correctly attribute justification in the demon world, in a way consistent with reliabilism, which I doubt, we will see that this same kind of position cannot account for the *value* of justified beliefs in the demon world. This in turn will defend the New Evil Demon thought experiment in showing something important about the value of justification: its fundamental value cannot be explained purely instrumentally.

4) Sosa's Reply to the New Evil Demon: Indexical Reliabilism

In response to the New Evil Demon problem about justification, a popular rejoinder has been to claim that justification has an indexical component, so that we correctly attribute justification in the demon world, since those subjects form beliefs via methods that are reliable in *our* world^{xx}.

Ernest Sosa, a prominent proponent of a non-standard form of reliabilism, couches the position as part of a virtue epistemology, where intellectual virtues are held to be stable and robust dispositions that produce a high ratio of true beliefs. Justified beliefs are then understood as those that arise through the exercise of such virtues. But according to Sosa, two quite different things are meant by 'epistemic justification', each of which he labels 'apt' and 'adroit' justification, which he defines as follows (where 'w' ranges over possible worlds):

J-Apt (For all w) [B is apt-justified in w *only if* B is acquired in w through the exercise of one or more intellectual virtues that are virtuous in w]

J-Adroit (For all w) [B is adroit-justified in w *only if* B is acquired in w through the exercise of one or more intellectual virtues that are virtuous in our *actual* world] (Sosa 2001, pp. 384-385)

Sosa understands ‘actual’ to be an indexical, where the context is set as the world of the subject who attributes justification, either through thought or speech (Sosa 2001, p. 399, ft. 12).

Given this distinction, Sosa thinks that when we judge that a subject enjoys justified beliefs in the demon world, this is true if we mean that she is adroit-justified in believing as she does (assuming that her belief-forming procedures would yield true beliefs in our actual world). If, on the other hand, we are talking about apt-justification, then the victims of the evil demon enjoy no such justification. Sosa thinks that this account makes just the kind of concessions that it ought to by trying to do justice to two kinds of intuitions we might have: while the victims of the evil demon have justified beliefs in *some* sense, it seems that they lack justification in another. Sosa explains this tension by appealing to different senses of justification.

Indexical Reliabilism has not been without its critics. Goldman has objected that there is no evidence that the folk relativize epistemic justification in this way (see section 2 of Goldman 1993). More recently, it has been argued such proposals suffer problems from the perspective of the philosophy of language (Ball and Blome-Tillman 2013).

Further still, I suggest that there is something going well with the victims’ beliefs, even relative to *their* world, not ours. For example, like us, the victims of the evil demon base their perceptual beliefs upon their perceptual experiences, they do

not make hasty generalizations and they obey the canons of good inductive inference, they are no more confident of their beliefs than their evidence warrants, their beliefs cohere with each other, etc. – all of which are valuable properties for beliefs to have. Since the victims of the evil demon world can be set up as our subjectively indistinguishable counterparts, if our beliefs enjoy justification, so do theirs, and on the same bases, and these positive epistemic qualities obtain relative to their *own* world, not ours. Given this, we ought to judge not only sameness of justification, but also sameness of value.

Nevertheless, suppose one grants for the sake of argument that Sosa's proposal is an adequate response to the New Evil Demon problem about justification. Does it help with the version of the problem I raised above for those who think that the value of justification is entirely instrumental on truth? It seems that it does not. Applying the basic strategy, one would have to hold that the victims of the evil demon have beliefs that are epistemically valuable since, while they are not instrumental to truth in *their* world, they *would* be an instrumental means to truth in *our* world. However, whether or not something is instrumentally valuable depends essentially on its environment, and how effectively it is a means to the given end in that environment. Outside of such an environment, instrumental value disappears. I will now argue that instrumental value does not relativize in the way a Sosa-style approach would require, and so this account cannot explain the value present in a demon world case.

To see this, consider a paradigm case of something of instrumental value: money. A particular note has instrumental value since it can be exchanged for goods and services, which themselves might be of instrumental or final value. But whether or not a piece of currency has instrumental value depends essentially on the

environment one finds oneself in. In Australia, a \$20 Australian note has value since it can be used to purchase things. Take the note out of this environment, and it loses its value – a \$20 Australian note is worthless in Bolivia (assuming that it cannot be exchanged for Bolivian Bolivianos, which can then be used as a means of making purchases).

Consider also a world in which there is no system of legal tender of any sort, but a \$20 Australian note is suddenly introduced. If we judge that the \$20 note is not valuable in Bolivia, then to be sure, we ought not judge that the note is of value in the world where there is no system of currency. It will not do to say that in such environments the note is still instrumentally valuable, *since relative to Australia*, it can be used as a means to acquire things of value. Instrumental value is essentially tied to the environment in which the bearer is a means to a valuable end.

So whereas a \$20 Australian note has no value once it is removed from the right kind of monetary environment, the victims of the New Evil Demon enjoy beliefs of epistemic value, even though in that environment they fail miserably in meeting the goal of truth. The value of justification, therefore, is not purely instrumental as a means to truth. And as it seems that there are no other reasonable candidates of value that justification could be instrumental in acquiring in the demon world (such as knowledge), this thought experiment makes the prima facie case that justification is of final value.

In case the above argument involving the example of money is viewed with suspicion since money is a social institution, it is important to note that we can get the same result -- that instrumental value is essentially linked with its environment -- by considering non-socially constructed things that are of instrumental value. Take

penicillin, for example. Penicillin happens to be of great instrumental value – for example, it can be used to cure diseases such as syphilis, returning a patient to health. That penicillin is a means to the end of curing disease, however, depends essentially on the environment we find ourselves in. If bacteria mutated and became resistant to penicillin, the medicine would lose its instrumental value. If all the diseases for which penicillin was a treatment were eradicated, the medicine would lose its value. No one should hold that in such circumstances, penicillin is still instrumentally valuable, since relative to worlds where it *is* a cure, it can be used as a means to acquire things of value, such as health.

Doubts may linger about the penicillin being of no instrumental value – after all, would it not be irrational to destroy the medicine, on the off chance that a disease *might* crop up which it could effectively cure? And even in the case of the Australian bank note held in Bolivia, would it not be irrational to discard the note, on the off chance that the bearer *might* make it back to Australia, and so the note could be used to purchase something of value? If so, does this not show that both the note and the medicine are of some instrumental value, even in the environment where it is no means to a valuable end?

I agree that it would be irrational to destroy the penicillin and the banknote, all things being equal, even if they are of no instrumental use in the environments in which they find themselves. This is not because they are of instrumental value, however. It is for the purely pragmatic reason that they *might* be of instrumental value in the future. If there is no cost or disadvantage in preserving the penicillin or the note, they should be kept, just in case a situation arises where they would be useful in securing some valuable end. But to be sure, these are cases of the *potentially* valuable, not the *actually* valuable.

It is important to notice a key respect in which the bank note and penicillin differ from the subject's beliefs in the demon world: there is a non-zero objective chance, a metaphysical possibility, that the note and the penicillin might be a means to a valuable end, depending on how the world unfolds^{xxi}; however, there is *no possibility* that the justified beliefs held in the demon world could even possibly be instrumental in securing truth – the malicious demon sees to that. By stipulation, the demon ensures that all of his victims' beliefs are false, and given the demon's power, they can never, by some chance, be true. Nevertheless, despite being of no instrumental value, the victims of the evil demon can enjoy justified beliefs that are of epistemic value, as careful reflection on our thought experiment reveals. Again, as there are no reasonable candidates of what this value could be instrumental to (e.g. not truth, not knowledge, etc.), we have good reason to think that this value, therefore, must be final value.

But what of the intuition that the victims of the evil demon can enjoy beliefs that are of epistemic value? What if someone were to grant the intuition, but claimed that the intuition errs, because it is merely responding to features that *would be* valuable in worlds where the justification would be instrumental to truth? Of course it is *possible* that the judgment is mistaken; intuitive judgments are not infallible, after all. But to merely raise the general possibility of our judgments of value being mistaken is not enough to undermine their probative force. For this objection to succeed, we are owed an independent reason to think that justification is not of final value, given that we have the intuition that it is.

To sum up, as my examples show, instrumental value is essentially tied to the environment in which its bearer is a means to a valuable end. So whereas money or penicillin have no value once they are removed from the right kind of environment in

which it can buy things or cures illness for example, the victims of the New Evil Demon enjoy beliefs that are of epistemic value, even though in that environment they fail miserably in meeting the goal of truth. So at the very least, the value of justification is not exhausted by any instrumental value it could have in getting truth in some worlds. In fact, the New Evil Demon case suggests that the value of justification cannot be explained in purely instrumental terms. As I have been suggesting, the best explanation of the kind of value that justification has is that it is not purely instrumental – rather, epistemic justification is of final value – it is valuable for its own sake.

5) Conclusion:

By reflecting on the case of the New Evil Demon in connection with the value of epistemic justification, we have seen that this value cannot be adequately explained as being instrumental to truth. False belief, which is no means to truth, can nevertheless still be of epistemic value. As there is no plausible candidate of what could account for this allegedly purely instrumental value in the demon world, we therefore have made a good prima facie case that justification is valuable for its own sake. If this is right, we have also found reason to think that truth value monism is false: assuming that true belief does have value, there is more of final epistemic value than mere true belief^{xxii}.

Department of Philosophy
College of Humanities and Social Sciences
United Arab Emirates University
P.O. Box 15551, Al Ain
United Arab Emirates

Sources Cited:

- Alston, William. (1989) "Concepts of Epistemic Justification", reprinted in *Epistemic Justification*. (Ithaca: Cornell University Press), 81-114.
- Ball & Blome-Tillmann. (2013) "Indexical Reliabilism and the New Evil Demon." *Erkenntnis* 78, 1317-1336.
- Bonjour, Laurence. (1985) *The Structure of Empirical Knowledge*. (Cambridge, Mass.: Harvard University Press).
- Bradley, Ben. (1998) "Extrinsic Value." *Philosophical Studies* 91, 109-126.
- Cohen, S. (1984) "Justification and Truth." *Philosophical Studies* 46, 279-295.
- Comesana, J. (2002) "The Diagonal and the Demon." *Philosophical Studies* 110, 249-66.
- David, M. (2001) "Truth as the Epistemic Goal", in *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Virtue, and Responsibility*, ed. M. Steup, (Oxford: Oxford University Press), 151-169.
- David, M. (2005) "Truth as the Primary Epistemic Goal: A Working Hypothesis", in *Contemporary Debates in Epistemology*, eds. E. Sosa and M. Steup, (Oxford: Blackwell), 296-312.
- DePaul, M. (2001) "Value Monism in Epistemology", in *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Virtue, and Responsibility*, ed. M. Steup, (Oxford: Oxford University Press), 170-183.
- Dorsey, Dale. (2012) "Can Instrumental Value Be Intrinsic?" *Pacific Philosophical Quarterly* 93, 137-157.
- Goldman, A. (1992) "What is Justified Belief?" in *Liaisons*. (Cambridge, Mass.: MIT Press), 105-126.
- Goldman, A. (1986) *Epistemology and Cognition*. (Cambridge, Mass.: Harvard University Press).
- Goldman, A. (1993) "Epistemic Folkways and Scientific Epistemology." *Philosophical Issues* 3, 271-285.
- Goldman, A. (1999) *Knowledge in a Social World*. (Oxford: Oxford University Press).
- Greenough, Patrick and Pritchard, Duncan, eds. (2009) *Williamson on Knowledge*. (Oxford: Oxford University Press).
- Huemer, Michael. (2006) "Phenomenal Conservatism and the Internalist Intuition." *American Philosophical Quarterly* 43, 147-58.

- Hyman, John. (1999) "How Knowledge Works." *The Philosophical Quarterly* 49, 433-451.
- Hyman, John. (2006) "Knowledge and Evidence." *Mind* 115, 891-916.
- Kagan, Shelly. (1998) "Rethinking Intrinsic Value." *The Journal of Ethics* 2, 277-297.
- Kelly, Thomas. (2008) "Evidence: Fundamental Concepts and the Phenomenal Conception." *Philosophy Compass* 3, 933-955.
- Kvanvig, J. (2003) *The Value of Knowledge and the Pursuit of Understanding*. (Cambridge: Cambridge University Press).
- Kvanvig, J. (2005) "Truth is not the Primary Epistemic Goal", in *Contemporary Debates in Epistemology*, eds. E. Sosa and M. Steup, (Oxford: Blackwell), 285-96.
- Lehrer, K. and S. Cohen. (1983) "Justification, Truth, and Coherence." *Synthese* 55, 191-207.
- Littlejohn, Clayton. (2009) "The Externalist's Demon." *Canadian Journal of Philosophy* 39, 399-434.
- Littlejohn, Clayton. (2012) *Justification and the Truth-Connection*. (Cambridge: Cambridge University Press).
- Madison, BJC. (2010a) "Epistemic Internalism." *Philosophy Compass* 5/10, 840-853.
- Madison BJC. (2010b) "Is Justification Knowledge?" *Journal of Philosophical Research* 35, 173-191.
- Madison, BJC. (2014) "Epistemological Disjunctivism and the New Evil Demon." *Acta Analytica* 29, 61-70.
- Pritchard, Duncan; Millar, Allan; Haddock, Adrian. (2010) *The Nature and Value of Knowledge*. (Oxford: Oxford University Press).
- Pritchard, Duncan. (2011) "What is the Swamping Problem?" in *Reasons for Belief*, eds. A. Reisner and A. Steglich-Petersen. (Cambridge: Cambridge University Press), 244-59.
- Pritchard, Duncan. (2012) *Epistemological Disjunctivism*. (Oxford: Oxford University Press).
- Ronnow-Rasmussen, Toni. (2002) "Instrumental Values – Strong and Weak". *Ethical Theory and Moral Practice* 5, 23-43.
- Silins, Nicholas. (2005) "Deception and Evidence." *Philosophical Perspectives* 19, 375-404.
- Sosa, E. (1991) "Reliabilism and Intellectual Virtue" in *Knowledge in Perspective: Selected Essays in Epistemology*. (New York: Cambridge University Press), 131-145.

Sosa, Ernest. (1993) “Proper Functionalism and Virtue Epistemology.” *Nous* 27:1, 51-65.

Sosa, Ernest. (2001) “Goldman’s Reliabilism and Virtue Epistemology.” *Philosophical Topics* 29: 1&2, 383-400.

Sosa, Ernest. (2003) “Beyond Internal Foundations to External Virtues” in *Epistemic Justification*. With Laurence Bonjour. (Oxford: Blackwell).

Steglich-Petersen, Asbjørn. (2014) Review of Clayton Littlejohn’s *Justification and the Truth-Connection* in *Notre Dame Philosophical Reviews*
<http://ndpr.nd.edu/news/45532-justification-and-the-truth-connection/>

Sutton, Jonathan. (2007) *Without Justification*. (Cambridge, Mass.: MIT Press).

Tucker, Chris, ed. (2013) *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism*. (Oxford: Oxford University Press).

Wedgwood, Ralph. (2002) “Internalism Explained.” *Philosophy and Phenomenological Research* 65 (2), 349-369.

Unger, Peter. (1975) *Ignorance*. (Oxford: Oxford University Press).

Williamson, Timothy. (2000) *Knowledge and Its Limits*. (Oxford: Oxford University Press).

Zagzebski, L. (2004) “Epistemic Value Monism”, in *Sosa and His Critics*, ed. J. Greco, Oxford: Blackwell, 190-198.

ⁱ The idea that it is truth that defines the epistemic is ubiquitous in contemporary epistemology; for a classic expression of this sentiment, see Alston 1989 p. 83.

ⁱⁱ Veritism is a theme that runs through Alvin Goldman’s work; see in particular Goldman 1999. Earlier, expressing a similar idea, Goldman wrote, ‘[...] I conclude that the verific approach to justificational value – in the objective consequentialist style – is the most promising. True belief is the value that J-rules [justification-rules] should promote – really promote – if they are to qualify as right. But so far, this is just a theory of *value* and a choice of objective over subjective consequentialism. We still do not have a determinate criterion of rightness.’ (original emphasis) (Goldman, 1986, p. 103) Goldman of course spells out this criterion of epistemic rightness along reliabilist lines.

ⁱⁱⁱ ‘*Epistemic Value T-Monism* True belief is the sole fundamental epistemic good. [...] the epistemic value of all other epistemic goods is instrumental value relative to this epistemic good. So, for example, on this picture the epistemic value of an epistemic standing like justification is to be understood purely instrumentally relative to the fundamental epistemic good of true belief. That is, justification is epistemically valuable, but only because it is a means to true belief; it is not, unlike true belief, epistemically valuable in its own right.’ (Pritchard 2010, p. 14) My central argument challenges both parts of this view: that true belief is the sole fundamental epistemic good, as well as that the value of justification is exhausted by its instrumental value. I intend to argue that epistemic justification is valuable for its own sake.

^{iv} Linda Zagzebski construes epistemic value monism as the view that: ‘Any epistemic value other than the truth of a belief derives from the good of truth.’ (Zagzebski 2004, p. 191)

^v Arguments in support of this position differ – for some examples, see David 2001; David 2005.

^{vi} For other criticisms of epistemic value monism, see for example DePaul 2001; Kvanvig 2003; Kvanvig 2005; Zagzebski 2004.

^{vii} I will follow recent convention and speak of the fundamental value of *true belief*. However, one might endorse a version of monism where it is *truth*, rather than true belief, that is the only fundamental epistemic good. For our purposes, however, nothing of significance will hang on this difference. For a discussion of this issue, see ch. 2 of Kvanvig 2003.

^{viii} The classic presentation of what has become known as the New Evil Demon argument is Cohen 1984.

^{ix} See for example Wedgwood 2002; Huemer 2006; Madison 2010a.

^x For classic expressions of process reliabilism, see Goldman 1986; Goldman 1992.

^{xi} See Silins (2005) p. 385 for discussion of evaluative v. explanatory theses in epistemology.

^{xii} Regardless of whether rationality and reasonableness are the same thing, it has recently been contested whether these notions are equivalent to justification. Clayton Littlejohn (2009) argues that in a New Evil Demon case, the internalist may be confusing putative justification in the demon world not with mere blamelessness, but rather, with excusability. On Littlejohn's account, excusability requires that the subject did what was reasonably expected to meet their obligations. Littlejohn 2009 section III.4 provides an in-depth examination of ways in which a subject's beliefs and actions may be blameless, namely: by being justified, by being exempt from responsibility, or by being excusable. According to this proposal, the subjects in the demon world are to be excused for their epistemic wrongdoing given their subjective perspective on the situation, even though the beliefs they hold are not justified.

Whether or not justification is distinct from reasonableness and rationality, my central point above still stands: if our beliefs and the beliefs of our counterparts are reasonable and rational, then this will at least be a strong indication that they are also justified (even if to concede that beliefs are reasonable and rational is not *ipso facto* to say that they are also justified).

^{xiii} This principle is closely related to what has become known as *Phenomenal Conservatism* in epistemology. For a recent collection of essays debating the merits of such principles, see Tucker 2013.

^{xiv} While nearly all philosophers hold that justification is non-factive, like all philosophical positions, there are of course some dissenters. See for example Sutton (2007); Littlejohn (2012). For critical discussion of Sutton, Madison (2010b); for critical discussion of Littlejohn, see Steglich-Petersen (2014).

^{xv} See Duncan Pritchard (2012), especially pp. 38-45. For critical discussion of Pritchard's treatment of the New Evil Demon case, see Madison (2014).

^{xvi} See Madison (2014) for a sustained defense of this claim.

^{xvii} Here and throughout I have been assuming that the victims of the New Evil Demon have evidence for their beliefs. Someone might object that the demon victims either have no evidence whatsoever, or if they do, they have far less of it than their non-demon world counterparts. One would think this if one holds Timothy Williamson's thesis that one's total evidence is just one's total knowledge (E=K). For Williamson's presentation of this position, see his (2000) ch. 9. Also, John Hyman and Peter Unger have independently defended nearly identical views of evidence; see Hyman (1999) and (2006); Unger (1975). Suffice it to say that Williamson's position has not been without its critics. See for example several of the essays in Greenough and Pritchard (2009).

While detailed engagement of Williamson's position is beyond the scope of the present paper, I shall note that I take the New Evil Demon problem to put pressure *against* the thesis that $E=K$, in the following way: *given that* the victims of the New Evil Demon are justified in believing as they do, and supposing that evidence is that which justifies belief, then there is evidence in the demon world. We must then ask: what must evidence be like in order that it can obtain in such worlds?

Perhaps it will turn out to be true that $E=K$ and therefore that evidence is factive: perhaps then victims in the demon world have different evidence than their non-demon counterparts, but on the face of it, that is still consistent with them having the same justified beliefs. For example, suppose that I am justified in believing that there is a painting before me, on the basis that I see that there is a painting before me. While my demon world counterpart cannot see that he is looking at painting (since that is false), he can *seem* to see that there is a painting before him (which is true). In addition, my counterpart knows that it seems to him that he sees a painting before him, and so he is in possession of evidence, which arguably, justifies him in believing that there is a painting before him.

In short, I take the New Evil Demon case as putting constraints on the theory of evidence, rather than approaching the New Evil Demon case with a theory of evidence already in hand, as a Williamsonian will. For discussion of the dialectical question of how to approach the theory of evidence, see Kelly (2008).

^{xviii} An anonymous referee for this journal suggested another possibility that deserves further investigation. It is worth exploring if the value of justification might be explained in the following way: if a belief is justified, then it is right to hold it; if it is right to hold it, then that is what one ought to do. In general, doing what one ought to do is better than failing to do what one ought to do. Here the thought is that the value of justification can be explained in terms of rightness and oughts, rather than in virtue of being instrumental to some independent epistemic value like true belief.

^{xix} Allan Hazlett suggested to me that my arguments in effect show the flip side of what the so-called 'swamping problem' aims to establish: I maintain that justified false belief can still be valuable from the epistemic point of view, and so the value of justification cannot be accounted for purely instrumentally as a means to true belief, whereas the swamping problem attempts to show that there is *more* of epistemic value than mere *true* belief, and so the value of justification cannot be purely instrumental. See Pritchard (2011) for more on the swamping problem, as well as the idea that it threatens any account of the value of an epistemic status that regards that status as more valuable than mere true belief, but explains that greater value purely instrumentally as a means to true belief.

^{xx} Sosa develops the position in several places, including Sosa 1991; Sosa 1993; Sosa 2001; Sosa 2003, etc.

^{xxi} Which possibilities are relevant here? Something can be instrumentally valuable even if it does not actually bring about a desirable state of affairs. For example, a fire extinguisher can be of instrumental value, even if it is never used to put out a fire. So it might seem that for a fire extinguisher to be instrumentally valuable, it must be *possible* for it to put out fires (in certain circumstances). But it is obviously far too inclusive to hold that an object actually has instrumental value just in case it has instrumental value in *some* possible world. This is because *any* object has instrumental value in *some* world or other. To take a recent example from Dale Dorsey, the body of a dead fly is not of instrumental value in this world just because it is instrumentally valuable in worlds where dead flies are as prized and as rare as diamonds are in our world (Dorsey 2012, p. 145). The present point is simply that specifying the nature of instrumental value is far from straightforward. In fact, instrumental value is a relatively unexplored area of value theory, one whose nature is both elusive and contentious. Some recent relevant discussions of instrumental value include Bradley (1998); Kagan (1998); Ronnow-Rasmussen (2002); Dorsey (2012).

^{xxii} Thanks to audiences at the Universities of Edinburgh, Charles Sturt, Sydney, Kent, and Cape Town. Thanks to an anonymous referee for *Pacific Philosophical Quarterly*, Victoria Madison, Gwen Bradford, and especially Rhiannon James and Allan Hazlett for helpful discussion and written comments on earlier drafts of this paper.