

Contra Kim on Naturalized Epistemology

December 11, 2009

Abstract

In ‘What is Naturalized Epistemology?’ Jaegwon Kim famously argues against Quine that, without normative epistemic notions such as justification and/or rationality one can’t make sense of the notion of having a belief. Thus, the ‘naturalized epistemology’ project of rejecting the normative notions of justification and knowledge, in favor of the merely descriptively study of how we manage to get true beliefs, which methods of inference are truth preserving and so forth, is incoherent.

In this paper, I will argue that, if there are such things as reasons, Kim hasn’t given us any good ones for believing in epistemic normativity. In a nutshell: Kim argues that we need the normative notions like justification in order to make sense of belief. But, I will argue that attributing beliefs does not require the presence of any person-level beliefs about justifications. And neither does the acceptance of the notion of belief commit us to giving a non-trivial conceptual analysis of believing, in the course of which we might have to appeal to notions like justification. Thus, whether or not the denier of epistemic normativity is irrational to deny that there are reasons, he’s not (in any intuitive sense) being irrational in remaining unmoved by Kim’s arguments.

1 Introduction

In ‘What is Naturalized Epistemology?’ Jaegwon Kim famously argues against Quine that, without normative epistemic notions such as justification and/or rationality one can’t make sense of the notion of having a belief. Thus, the ‘naturalized epistemology’ project of rejecting the normative notions of justification and knowledge, in favor of the merely descriptively study of how we manage to get true beliefs, which methods of inference are truth preserving and so forth, is incoherent.

In this paper, I will argue that, if there are such things as reasons, Kim hasn't given us any good ones for believing in epistemic normativity. In a nutshell: Kim argues that we need the normative notions like justification in order to make sense of belief. This argument can be understood in two ways - either as a) a claim that we need to assume the existence of epistemic norms in order to give an adequate conceptual analysis of the notion of belief (a notion which Quine and other deniers of epistemic normativity do accept) or b) a claim that in attributing beliefs the denier of epistemic normativity implicitly assumes those norms for belief which he denies in the philosophy classroom.

But, whichever of these arguments Kim intends to be making, (I claim) his argument fails. Accepting the notion of belief, does not commit one to giving a non-trivial conceptual analysis of believing, in the course of which we might have to appeal to notions like justification. And attributing beliefs does not require the presence of any *person-level* beliefs about justifications. Thus, whether or not the denier of epistemic normativity is, in some way, irrational to deny that there are reasons, he's not (in any intuitive sense) being irrational in remaining unmoved by Kim's arguments.

2 Denying Epistemic Normativity

First, let me say a bit about how I understand the position which I will be defending (at least for the sake of argument) from Kim. By 'denying epistemic normativity' I mean the position that there are facts about what's true, but no facts about what you should believe, or which people are justified in believing which propositions.

We can make this position vivid to ourselves by focusing on cases where the truth of all propositions in question is uncontroversial, but substantial questions about justification remain.

So, for example, the denier of epistemic normativity (henceforth The Denier) will agree that certain mathematical propositions are true like $2+2=4$ or the four color theorem. However, if you ask them which of these necessary mathematical truths one ought to believe, they will reject the presumption that there are any facts about what one ought to believe. Similarly, if we ask The Denier to consider two people, one of whom has normal psychology which causes him to find $2+2=4$ obvious without appeal to any further claims for proof, while the other has abnormal psychology which causes him to find the four color theorem equally obvious, without appeal to any further claims, he will say there is no fact of the matter about which of these people are justified in assuming the propositions that they do.

Note that The Denier may well accept almost all the same reasoning as someone who accepts epistemic normativity. Philosophers who deny that there are moral facts may well volunteer for the same organizations and avoid the same crimes as those who accept them. They just do this without making the claim that it is good to so volunteer, or one ought to avoid such crimes. Similarly, The Denier will deploy modus ponens and mathematical induction and so forth, like everyone else, (and believe these methods are reliable like everyone else) but reject the further claim that one ought to reason in accordance with these methods. The only inferences he won't make are those that involve ascriptions of epistemically normative properties like justification and knowledge.

Also, note that The Denier can consider various descriptive properties. Some of them will have sharp boundaries, like the property N of being a necessarily truth preserving method of inference. Others will have more vague boundaries, such as the property M of being a method of inference that most people would accept, even after spending a few days locked in a room with a working logician, or the property of being both N and M (i.e. both truth preserving and popular).

They might even call the latter property “justification*”, and talk about which propositions are justifiable*. But, insofar as the extension of justification* differs from that of intuitive conception of justification (as presumably any descriptive surrogate will), they will say that there are no facts about justification, only the descriptive surrogate justification*. And, obviously, he won’t say that one ought to believe those propositions which are justified*. Just as a moral antirealist can talk about morality in scare-quotes e.g. ‘the kind of actions people tend to admire’[allowing a certain amount of vagueness about actions that are admired by many but not all], the denier of epistemic normativity can talk about the kind of ‘the kind of truth-preserving inferences people which tend to make’.

3 Kim’s argument

In section 5 of “What is Epistemology Naturalized?” Kim aims to argue that “the concept of belief is an essentially normative one”. What does this mean? Kim succinctly states his argument as follows, citing Davidson with regard to the second sentence.

“Belief attribution essentially requires a “radical interpretation” of the cognizer, of his speech and intentional states; that is we must construct an “interpretive theory” that simultaneously assigns meaning to his utterances and attributes him beliefs and other propositional attitudes.

Even a cursory consideration indicates that such an interpretation cannot begin- we cannot get a foothold on our subject’s realm of meanings and intentional states unless we assume his total system of beliefs and other propositional attitudes to be largely and essentially rational and coherent”

I fear I must start by taking this claim obnoxiously literally. When so understood, it says:

- (1) For A to attribute B a belief, A must construct an *explicit* theory which

assigns meanings to B's utterances, and in order to do this A must assume that B is rational.

If (1) were true, it would certainly show that it was impossible to for someone who didn't think there were facts about rationality to interpret others as having beliefs, and the project of naturalized epistemology (which aims to study the engineering question of how people manage to actually form true beliefs) would be scuttled.

However, it is *surely false* that interpreting someone requires consciously constructing a theory of what their words mean. For, some english speakers have interpreted other english speakers. However, no one has ever yet come up with an explicit theory which assigns meaning to even a small fraction of an english speaker's utterances. The task of giving an explicit Davidsonian theory of how English words combine to determine the truth conditions for constituent sentences is a massive open problem in linguistics! Thus, interpreting an english speaker cannot require formulating an *explicit* theory about what their utterances mean as per (1).

We can see the same point even more quickly by introspection. Is it the case that whenever you attribute someone a belief you first consider an explicit theory about what words in their language mean? Surely not, and if perchance you do, you could make a lot of money by remembering what this theory was, and telling some linguists.

This suggests two possible understandings of Kim's argument.

Version A:

Kim thinks interpretation involves somehow using an *implicit* theory about the meaning of the person's utterances, in the construction of which one needs to make *implicit* assumptions about what's rational to believe.

We clearly do have dispositions to interpret people as saying certain things,

and one might consider this to amount to an implicit theory about what they mean. For example, perhaps Kim thinks:

- In order for A to interpret B's "Schnee ist weiss" as meaning that snow is white, A must have an implicit theory of what B's words mean -in the sense of being inclined to translate "Schnee ist weiss" as snow is white, and various other of B's utterances in various ways.
- The only way to "even begin" to get such a theory/acquire these dispositions without assuming that the subject we are translating is rational.

In this case, Kim's argument centers on the following claim: in order to acquire dispositions to translate we need to assume the people we are trying to translate are rational (and hence, that there are facts about rationality).

Version B: We don't need to make assumptions about rationality in the course of attributing beliefs, but we do need to accept this notion in order to properly make sense of what beliefs are, as philosophers. On this version of the argument, Kim's point is that one cannot give an adequate conceptual analysis of belief without appealing to some notion of epistemic normativity, and thus anyone who accepts the notion of belief is forced to accept some notion of epistemic normativity as well.

4 Version A

Let's start with the first interpretation. Is it really the case that when Quine and co. interpret people they are implicitly appealing to beliefs about what it is rational to believe?

4.1 Implicit Beliefs

As noted before, it's surely false that in order to learn its first language a baby needs to *explicitly* say to itself 'I will assume that the people around me are rational'. It's not even clear that a baby could think this thought prior to learning its first language. Also, even for adults, the experience of picking up new slang doesn't seem to involve any conscious postulation of rationality. Indeed (at least in my personal experience), it doesn't even seem to involve any conscious posits. Rather, as I see more and more instances, the word 'blog' (or whatever it may be) just comes to seem more and more informative to me/leads me to make more and more inferences etc.

However, something like implicit assumptions may well be involved, in the following sense. There are many different possible functions from the sentences in English to the class of propositions. In learning to interpret someone, we acquire dispositions that correspond to only one of these functions. In effect, we rule out a wide range of possible meaning attributions in response to a relatively small sample of experiences with people using english words. Thus, it seems like certain possibilities (e.g. everyone you will ever interact with always being wrong in the way they use the word "blog") are ruled out in advance. Perhaps no possible course of experience would get you to believe this is the case. Thus there's a sense in which whatever sub-personal properties eventually produce my confident identification of "blog" with "collection short casual bits of original writing posted to the web" ignore certain possibilities.

So, if we imagine that the brain processes which produce this identification involve some kind of homunculus which itself has beliefs, then this homunculus would seem to be working on the assumption that there are certain particular relationships between the meaning of a person's sentences and what those sentences mean in their language. Non-metaphorically, we can put the point like

this. There are many more possible translations of a language than there are (distinguishably different) courses of experience which might lead one to adopt one of these translations. So, whatever function describes the way our brains react to hearing sequences of words in our first language cannot be 1-1.

But, between these truths and Kim's conclusion there are a number of serious problems.

4.2 Objections

Firstly, there's the obvious issue of the homunculus metaphor. At best it makes metaphorical sense to say that sub-personal processes have beliefs and make "assumptions", like the assumption of rationality.

Secondly, to the extent this does make sense, (the processes give you meaning attributions that maximize some set of constraints containing rationality), it's much less clear that it makes sense to ask questions about whether your brain processes are assuming people tend to believe what's rational and that inferences on list X are rational, vs. that it's assuming that people tend to believe what has descriptive property D and the inferences on the list have property D.

Thirdly, if we assume THIS makes sense, the empirical evidence seems strongly to support the latter claim (note: we know empirically it's not the case that the list used is literally the same one used to produce assent/dissent to "is this rational" because when people deny epistemic normativity they don't instantly stop ascribe beliefs)

Note that it sounds much cleaner to say 'we interpret so as to maximize rationality' vs. 'we interpret so as to make people tend to make these inferences and those etc.'. So you might *think* there's some argument from information storage or complexity (e.g. there has to be some simple theory about how to translate otherwise how could our finite minds do it?) But notice that in

order to do the translation the homunculus has to somehow get to conclusions about which combinations beliefs are irrational to avoid attributing them. So a system that “assumes” that people are rational, stores that a certain collection of inferences are rational and then tries to attributing them where possible, will have to be at least as complex as one that directly tries to avoid attributing those inferences.

Finally, under the incredible assumption that there’s a definite fact about what your brain is assuming, and what it’s assuming is that rationality is maximized, it’s not clear why you should agree with it about metaphysics. Would the Chinese room be committed to all of Searle’s views which he used to produce it’s behavior?

Thus it would seem that the only person whose assumptions you are committed to upholding (you) doesn’t explicitly formulate a meaning theory, and hence doesn’t make any assumptions in doing so. And only thing who could be understood as explicitly formulating a meaning theory (the language faculties in your brain) and hence possibly needing to assume things about rationality in doing so (the language learning centers in your brain) is almost certainly not something that has beliefs or makes assumptions, and certainly not something whose assumptions you are committed to agreeing with.

5 Reliability

At this point, Kim might point out that you are (plausibly) committed to thinking whatever sub-personal processes produce your ascriptions of belief to people are reliable. If you thought that the language learning centers were ignoring perfectly good possibilities in giving rise to your intuitions about how to translate people, this would presumably give you reason to treat these intuitions with skepticism.

It might seem to follow from this that you are committed to agreeing what the language faculties are assuming - at least if we grant Kim the first three points above, and hence assume this sub-personal faculty is making some assumption about rationality.

But not so! What you are committed to thinking is that *the result of this sub-personal processing is reliable*. All you have to think is that meaning supervenes on use in such a way that assigning meaning by ignoring *those things which your sub-personal processes were inclined to ignore* is reliable. And, note that the latter is a descriptive category!

6 Version B

Let us now turn to the second reading of Kim's argument. Perhaps Kim means to argue as follows: the notion of justification and epistemic normativity is necessary to make sense of the very idea of beliefs. Someone believes that P iff an ideal interpreter would assign them the belief that P. And such an ideal interpreter assigns them beliefs, by interpreting their utterances in such a way as to jointly maximize a) the simplicity of the interpreter's theory and b) the degree to which (on the whole) the subject comes out to have beliefs that are justified. Thus, it doesn't make sense to study the reliability with which someone forms true beliefs, while rejecting the notion of epistemic normativity.

Now, I have two objections to this. Firstly, it is not clear that, in order to understand the notion of belief, we must accept some kind of analysis of it into other terms. You might think: we are just trained in the practice of interpretation, like we are trained to recognize certain things as games. We don't do this by consciously reasoning about justification, and Davidson's maxims or any other thing that one might use to try to define the notion of belief. Maybe there aren't any informative necessary and sufficient conditions for having a

belief that P, or the only conditions are extremely complicated and will only be discovered after years of work by linguists. If this is right, the argument ‘Unless the notion of justification is coherent, there will be no informative analysis of what it takes to count as having a given belief! Therefore, the notion of justification is coherent.’ looks pretty unconvincing. Maybe the notion of belief is primitive.

Secondly, suppose that we do actually attribute people beliefs by first (consciously) thinking of what they would be justified in believing given their situation. Even then I don’t see how this requires anything other than what you might call a ‘Cornell anti-realist’¹ notion of justification. That is, there is a distinction between token beliefs that are justified vs. unjustified *but* this distinction doesn’t turn out to pick out the kind of philosophically interesting/normative distinction, which typical users would naively expect it to. So, people did indeed use words like ‘chastity’ or ‘bearing the white man’s burden’ in a relatively coherent way, and we can pick out the descriptive class of things that these terms were thought to apply to. But, a historian can talk about this class of things (e.g. they might say “despite making these criticisms of her niece, the duchess often engaged in ‘unchaste’ behavior herself”) without, himself, taking the notion of chastity to have any normative force. (So you could also call this descriptive analog, of a term X which normally has normative/philosophical baggage, the notion of ‘X in scare quotes’.)

Cornell anti-realism with regard to justification, says that we have a well-established practice of separating propositions into things that it’s OK to assume vs. not. So, for example, we say it’s OK to assume that $2+2=4$, or that every red ball is a ball, but not OK to assume that water is H_2O , or the contium hypothesis, or some complex logical truth which turns out to be valid, when

¹This name is kind of a joke, about the meta-ethical position ‘Cornell realism’, which people say protests too much in calling itself a kind of realism.

that feels obvious to you, but you don't have a derivation of it from claims that feel obvious to other people. And we *think* that in doing so we are tracking some some special philosophical difference, between true propositions.

But in fact, this distinction in our practice, doesn't track anything philosophically interesting. Rather, it's just that there are certain true propositions, which the vast majority of people feel are obvious, and these are the ones we say it's OK to assume. We say that people who assume these true propositions, and reason from them, count as having knowledge, whereas people who assume other true propositions just have fortunate true beliefs. Practically, it's clearly useful to draw such a distinction, since it tracks the distinction between cases where a person can convince pretty much everybody else of their belief P (they are reasoning from premises that everyone finds obvious) vs. cases where some people may well find all the arguments they can give for P question-begging, since they don't believe some of the true premises. But, all *this* distinction tracks, is a distinction in human psychology: what truths do human beings typically find obvious? Can S justify the belief that P, from premises which (actually) most humans find obvious? If so, they count as having knowledge, and otherwise they don't. Knowing this, one might continue to use the word "justified", just as a romantically adventurous Bloomsbury-set historian might keep using the word "unchaste", but they'd be using it in the merely descriptive, scare-quoted sense.

This kind of deflationary view of what justification amounts to, seems quite compatible with Quine's position, and his new proposal for epistemology. If my expression "knowable a priori" just rigidly designates the class of true propositions which either a) feel obvious all employed people circa 2009 or b) can be derived from these, via a process of reasoning which feels obvious to all employed people circa 2009, this class of propositions may be interesting to psy-

chologists, but it doesn't seem very interesting philosophically. And, I don't see how attributing beliefs, even if it does require thinking about justification, requires accepting a notion of justification that's any more normative than the historian's notion of chastity. Thus, I don't buy that Kim has genuinely posed a problem for Quine.

7 Conclusion

Thus, it would seem, there is no sense in which The Denier is committed to accepting a notion of rationality, by their willingness to assign beliefs. All they are doing is taking their immediate judgements about what people mean at face value. And all this commits them to is (something like) thinking that their immediate inclinations to ascribe beliefs to people and content to utterances are reliable.

Now, it so happens that the possible combinations of beliefs which the Denier ignores when interpreting people, are also ones that are generally considered to be massively irrational. So, the Denier might be inclined to say that it's impossible to count as being in one of *these states*. He might even come up with a (vague at the edges) descriptive category of the kind of inferences people tend to avoid making, and avoid ascribing. But he is not committed to thinking that the inferences falling in this category have any additional normative property of *being such that they ought not be made*. The advocate of epistemic normativity's claim that some kind of inference *ought* not be made, does not follow from the claim that it would be impossible to count as accepting sufficiently many instances of that kind of inference.