

‘PUTNAM ON BIVS AND RADICAL SCEPTICISM’

DUNCAN PRITCHARD & CHRIS RANALLI

University of Edinburgh & National Autonomous University of Mexico (UNAM)

ABSTRACT. The aim of this paper is to explore Putnam’s influential ‘BIV’ argument against radical scepticism, both as he presents this argument and as it has subsequently been reconstructed. §1 explores the BIV argument as Putnam presents it and the anti-sceptical morals that he extracts from this argument. §2 examines a core critique of the argument, so conceived, from Anthony Brueckner. §3 then critically evaluates an influential reconstruction of Putnam’s argument, due to Crispin Wright. §4-5 explores the idea that Putnam’s argument is best thought of as a transcendental response to radical scepticism, and accordingly applies Stroud’s challenge to transcendental arguments to this proposal. Finally, §6 examines an influential criticism of Putnam’s argument which is due to Nagel.

1. PUTNAM’S BIV ARGUMENT

A familiar way of arguing for radical scepticism is by appeal to radical sceptical hypotheses, such as the hypothesis that one might be a brain-in-vat (BIV) which is being radically, and undetectably, deceived about its environment. Roughly, the sceptical argument goes that since such sceptical hypotheses are by their nature indistinguishable from normal experience, so one cannot know that they are false. Furthermore, if one cannot know that they are false, then it follows that one can’t know much of what one believes, most of which is inconsistent with radical sceptical hypotheses.

This last step will almost certainly require some sort of closure-style principle, whereby knowledge is closed under known entailments. Thus, if one does have knowledge of the ‘everyday’ propositions which one takes oneself to know (e.g., that one has hands), and of the fact that these propositions entail the denials of radical sceptical hypotheses (e.g., the BIV hypothesis, because BIVs don’t have hands), then one could come to know that one is not a

BIV. Conversely, insofar as one grants that it is impossible to know that one is not a BIV, then it follows that one cannot know the everyday propositions which are known to be inconsistent with the BIV hypothesis either.

There are many ways of responding to radical scepticism of this form, which we will refer to as *BIV scepticism*. One might deny the relevant closure principle, for example, or one might put forward an epistemology according to which one could know the denials of radical sceptical hypotheses, and so on.¹ On the face of it, Hilary Putnam's (1991, ch. 1) famous argument against BIV scepticism would appear to be a variation on this last anti-sceptical strategy, in that he also seems to be, effectively anyway, claiming that one can know that one is not a BIV.

The parallels between these two anti-sceptical approaches are superficial, however, and the differences significant. In particular, Putnam's overarching goal is not to make the *epistemological* claim that such anti-sceptical knowledge is possible, but rather to motivate the *semantic* claim that one cannot truly think the thought that one is a BIV. But these claims are logically distinct, in both directions. That one can know the denial of the BIV hypothesis is obviously consistent with the possibility that one can truly think the thought that one is a BIV. And while admittedly not so obvious, that one cannot truly think the thought that one is a BIV is consistent with being unable to know that one is not a BIV (we will return to consider this particular non-entailment claim below).

In order to understand Putnam's semantic argument we first need to get to grips with the variety of content externalism that he advances. For the starting point of Putnam's BIV argument is the idea that in order to be able to even think thoughts with a particular content, then certain external conditions—in particular, causal conditions—need to be met. For example, if one is a Martian who has never encountered trees, or been part of a linguistic community which has an established practice for referring to trees, then one cannot even think thoughts about trees. One's thoughts would instead be about something altogether different, insofar as they had content at all.

If we grant the truth of this content externalist thesis to Putnam, then one can immediately see how it can create a problem for BIV scepticism, at least so long as this form of radical scepticism is formulated in a particular way (the reason for this *caveat* will become apparent in a moment). For what goes for our Martians when it comes to having thoughts about trees will also apply to a BIV. In particular, a BIV can no more think thoughts about trees than a Martian can, since neither have enjoyed the right kind of causal connections to trees to make this possible. Insofar as one is able to think thoughts about trees, then, it follows that one is not a

BIV. More generally, what Putnam is disputing is the idea, common to presentations of radical scepticism, that BIVs and their non-envatted counterparts are thinking exactly the same thoughts, the only difference being that while the former thoughts are massively false, the latter thoughts are mostly true.

In order to bring this anti-sceptical line of argument into sharper relief, consider Putnam's claim that one cannot truly think that one is a BIV, such that this claim is in a sense necessarily false. Of course, if one is not a BIV, then this thought trivially expresses a falsehood. But what if one is a BIV? This is where Putnam's content externalism comes into play, since he maintains that such a thought still expresses a falsehood. After all, if one is a BIV, then one cannot even think the thought that one is a BIV, but must inevitably think a thought with a different content instead. Given Putnam's brand of content externalism, it is thus *impossible* to truly think the thought that one is a BIV, and one can, it seems, derive this conclusion on purely *a priori* grounds.

In order for this argument to go through, of course, it is important that being a BIV is analogous to being a Martian in the relevant respects. This is where it becomes important to Putnam's argument—and he is quite explicit about this point—that the BIV hypothesis is construed a certain way, such the brain in question has always been envatted, and such that the stimulation of the BIV is undertaken by supercomputers (and always has been). So construed, there is no route, whether direct or indirect, whereby the BIV could become causally related in the appropriate way to the kinds of things that normal non-envatted agents have thoughts about, such as trees. In contrast, had the envattment been recent, for example, then it would have been in principle possible for the BIV to have had the right kind of causal connections to trees prior to succumbing to this predicament.²

One consequence of this point is that Putnam's response to BIV scepticism is essentially limited, in that it will not apply to all forms of BIV scepticism, but only those formulated in terms of a specific rendering of the BIV hypothesis. Even so, if Putnam's argument goes through it is still a startling result. For is it really possible to argue on purely *a priori* grounds that even this specific form of BIV scepticism is false? Henceforth, we will grant to Putnam that the BIV hypothesis is construed in the way that he stipulates in order to further evaluate this claim.

2. BRUECKNER'S CRITIQUE OF THE BIV ARGUMENT

We can divide the criticisms of Putnam's anti-sceptical argument into two broad categories. The first type says that the argument is in some way *question begging*, while the second type says that the argument is in some way *self-refuting*.

Anthony Brueckner's (1986) core criticism of Putnam's anti-sceptical argument falls into the first category. In particular, he argues that Putnam's anti-sceptical argument begs-the-question against the sceptic in the following sense: the argument seeks to show that we're not BIVs, but establishing that conclusion requires us to know a principle which is alone sufficient for that anti-sceptical conclusion. In order to see how Brueckner's criticism works, let's consider his reconstruction of Putnam's argument:

- (P1) Either I am a BIV (speaking vat-English) or I am not a BIV (and, instead, I'm an English-speaking human).
 - (P2) If I am a BIV, then my utterances of "I'm a BIV" are true iff I'm a BIV*.
 - (P3) If I am a BIV, then I'm not a BIV*.
- So, from (P2) and (P3):
- (C1) If I am a BIV, then my utterances of "I'm a BIV" are false.
 - (P4) If I am not a BIV, then my utterances of "I'm a BIV" are true iff I'm a BIV.
- So, from (P4):
- (C3) If I am not a BIV, then my utterances of "I'm a BIV" are false.
- Therefore, from (P1), (C1), and (C3):
- (C) My utterances of "I'm a BIV" are false.

How can (C) enable one to know that one is not a BIV? First, knowledge is closed under competent deduction. So, if I know that the premises of the argument are true, and I competently deduce (C) from the premises, I'm thereby in a position to know that (C). And (C) says that when I say "I'm a BIV", I express something false. So, I know that when I say "I'm a BIV", I say something that's false. However, as Brueckner points out, the argument for (C) has not established the crucial anti-sceptical premise, namely:

- (*) I know that I'm not a BIV.

According to Brueckner, in order to show that I'm not a BIV, and thereby establish (*), the following principles have to be invoked:

- (A) My utterances of "I'm not a BIV" are true.
- (B) My utterances of "I'm not a BIV" are true iff I'm not a BIV.

From (A) and (B), that I'm not a BIV follows. From the principle that knowledge is closed under competent deduction, it follows that I'm in a position to *know* that I'm not a BIV. However, (B) seems dubious here. After all, as (P2) and (C1) highlights, not all languages are *disquotational*.

Indeed, vat-English is one such non-disquotational language. So this raises the question of whether the language in which I utter “I’m not a BIV” is a disquotational language (like English, for normal non-envatted speakers) or not (like vat-English, for envatted speakers). But according to Brueckner:

I cannot claim to know that I am not a BIV until I can claim to know that [*the anti-sceptical argument*] is a sound argument and that it somehow allows me to know that I am not a BIV. [...] Can I claim to know this without assuming that I am speaking English? (Brueckner 1986, 159)

Brueckner issues the proponent of Putnam’s anti-sceptical argument a challenge here: can one claim to know that their utterances of “I’m not a BIV” are true without *presupposing* that their utterances have disquotational truth-conditions? On Brueckner’s view, it can’t be done because “this must be shown by an anti-skeptical argument, not assumed in advance.” (Brueckner 1986, 160) In other words, vat-English is not disquotational, while English is. BIVs can’t speak English. But since the Putnamian anti-sceptical argument cannot establish that I’m not a BIV without presupposing that I’m speaking English—and therefore not vat-English—it thereby presupposes what it seeks to prove: that I’m not a BIV.

3. WRIGHT’S RECONSTRUCTION OF THE BIV ARGUMENT

In the last section we saw that Putnam’s anti-sceptical argument seems to depend on two core assumptions: that the language in which the anti-sceptical argument is framed is disquotational, and that in the BIV’s language, “BIV” does not refer to BIVs. We also saw how the first assumption might be problematic. As Brueckner argued, since the anti-sceptical argument seems to require the assumption that the speaker’s language is disquotational, and this assumption presupposes that the speaker’s language is *not* the BIV’s language, it follows that the speaker presupposes that the language in which their argument is framed is not the BIV’s language. But isn’t this a crucial step in Putnam’s anti-sceptical argument—one that is meant to be shown, rather than presupposed?

At this stage, one might wonder whether Brueckner’s reconstruction of Putnam’s anti-sceptical argument is the most charitable reconstruction. In addition, one might wonder whether, even if Brueckner’s reconstruction is correct, these two core assumptions are as problematic as Brueckner suggests. Crispin Wright (1992) follows Brueckner in treating those two assumptions as core, but he argues that the disquotation assumption is far less problematic than Brueckner suggests. On Wright’s view, it is the second assumption that is the more problematic of the two.

In order to understand Wright's objection against the second assumption, let's start with his reconstruction of Putnam's anti-sceptical argument.³ First, we begin with the *prima facie* plausible thought that "BIV" is meaningful. After all, it didn't seem to us that Putnam's explanation of the BIV hypothesis was incoherent, in the same way that a set of nonsense sentences immediately strike us as incoherent. So, phrased in the first-person:

- (1) In my language, "BIV" is meaningful.

Now, from (1) and the principle that my language is disquotational, it follows that:

- (2) In my language, "BIV" refers to BIVs.

From (2) and the principle that, in the BIV's language, "BIV" does not refer to BIVs, it follows that:

- (3) My language is not the BIV's language.

However:

- (4) If I am a BIV, then my language *is* the BIV's language.

Therefore, from (3) and (4):

- (5) I am *not* a BIV.

Let's focus on the step from (2) to (3). As we said earlier, Wright maintains that the second assumption of Putnam's anti-sceptical argument is the more problematic assumption. What makes it problematic, on his view, is how it enables the transition from (2) to (3). For while it might allow the speaker to conclude that their language is not the BIV's language, it is far from clear that it should allow the speaker to conclude that other speakers do not speak the BIV's language. But according to Wright, if the argument is valid, then it ought to make no difference whether the argument is framed in the first-person or the third-person.

We might imagine Putnam running his anti-sceptical argument past us, where the first-person pronoun featured in its premises refers to him. Imagine further that Putnam then asks us to consider whether the argument is valid and sound. Reflecting on his argument, we reformulate it so that the first-person pronoun is replaced with a co-referential third-person pronoun. We then state the argument as follows:

- (1*) In Putnam's language, "BIV" is meaningful.

From (1*) and the assumption that Putnam's language is disquotational, it follows that:

(2*) In Putnam's language, "BIV" refers to BIVs.

Therefore, from (2*) and the assumption that, in the BIV's language, "BIV" does not refer to BIVs:

(3*) Putnam's language is not the BIV's language.

However:

(4*) If Putnam is a BIV, then his language is the BIV's language.

And with an application of *modus tollens* on (3*) and (4*), it follows that:

(5*) Putnam is not a BIV.

But isn't it altogether incredible that we should be able to prove, *a priori*, that *Putnam* is not a BIV?

According to Wright, it is. As he puts the objection:

Without supplementary information, you cannot validly infer anything from [*the assumption that Putnam's language is diquotational*] and [*the assumption that in Putnam's language, "BIV" is meaningful*] about how to *specify* what is the reference of 'brain-in-a-vat', as used in [*Putnam's*] language. All you can infer is that a specification in [*Putnam's*] language would be homophonic. That is the same thing as [*premise (2*)*] only if it is presupposed that your language—the language in which the argument is presented—is [*Putnam's*]. (Wright 1992, 77)

Wright thinks that the move from (1*) and (2*) to (3*) is invalid, unless of course we presuppose that our formulation of the argument above is framed in the same language as Putnam's. But this might be contentious. After all, from the fact that our languages are homophonic—that we sound the same, and have the same phonetic and lithographic representations—it doesn't follow that the meanings of our terms are the same. Indeed, this point flows from semantic externalism. Wright's point is that, for all we know as the assessors of the above argument, "BIV" in Putnam's language means *Frenchman*. We're not in a position to know that "BIV" as used in argument means what it does in our language without additional information. What kind of additional information? Wright suggests the following:

(AI) The previous anti-sceptical argument was formulated in Putnam's language, and Putnam's language coincides with the language of the assessor of the argument, inasmuch as the meaning assigned to all of the referring expressions therein are the same.

Now (AI) would be quite useful, in that we would now be in a position to make a valid inference from (1*) and (2*) to (3*). But Wright's issue with (AI) is with whether we are in a position to

know that our use of “BIV” means for us, the assessors of the argument, what it means for Putnam, the proponent of the argument. For example, with (AI), we can now see that the sub-conclusion (3*) follows from the premises. But what’s left open is how we should understand (2*). After all, we could claim that “BIV” means whatever it does in our language, but with (AI) that’s just trivial. This raises the issue of what “BIV” *means* in our language, and how we know what it means. In short, Wright worries that semantic externalism problematizes one’s *knowledge of content* in the following sense: the kind of knowledge of content that externalism seems to forbid (or at least problematize) is the same kind of knowledge that is needed to make Putnam’s anti-sceptical argument work.⁴ In particular, semantic externalism calls into question the kind of knowledge of content that is needed to secure our knowledge that “BIV” means what we think it means.⁵

4. PUTNAM’S ARGUMENT *QUA* TRANSCENDENTAL ARGUMENT

According to Wright (1992, 85), the most charitable interpretation of Putnam’s anti-sceptical argument is that it’s a “transcendental argument”, and Putnam (1981, 16) himself also describes the argument in these terms.⁶ What makes the argument ‘transcendental’? Put generally, an anti-sceptical argument is transcendental when its conclusion is about non-psychological reality, and its premises are about our psychology, and the non-psychological, necessary conditions for having that type of psychology.

So construed, Putnam’s anti-sceptical argument does seem to fit the bill. Consider this reconstruction of the argument:

- (T1) A necessary condition of the possibility of thinking that *I’m not a BIV* is that I’m not a BIV.
- (T2) I can think that *I’m not a BIV*.
- (TC) I’m not a BIV.

I can think that “I’m not a BIV” says something that’s true.⁷

But the sentence is ambiguous. After all, recall the point that the sentence “I’m not a BIV” in the BIV’s language is true iff I’m not a BIV-in-the-image, while the sentence “I’m not a BIV” in English is true iff I’m not a BIV. In either case, whatever proposition we express with “I’m not a BIV”, it’s a true proposition. The pressing question is *which* true proposition are *we* expressing? That is to say, when we say or think *I’m not a BIV*, *which* true proposition are we expressing? We could say that it expresses the proposition that *I’m not a BIV* rather than the proposition that *I’m not a BIV-in-the-image*. But how are we supposed to know this?

Brueckner puts this objection to the Putnam transcendental argument as follows:

If I am in a vat world, then my sentences do not have disquotational truth-conditions and my beliefs do not have disquotational contents. If I am instead in a normal world, then my sentences do have disquotational truth-conditions and my beliefs do have disquotational contents. If I do not yet know which sort of world I am in (this is what the Putnamian argument is supposed to settle), then it appears that I do not know which sorts of truth-conditions and contents are mine. So it appears that the Putnamian argument is epistemically circular [...] That is, in order to know, or have justification for believing, the argument's self-knowledge premise [*That I can think that I'm not a BIV*], I need to know, or have justification for believing, its conclusion [*that I'm not a BIV*]. (Brueckner 2010, 95)

Brueckner's criticism puts considerable pressure on a proponent of the anti-sceptical efficacy of the Putnamian transcendental argument. The benefit of using an anti-sceptical *transcendental* argument over any other kind of anti-sceptical argument is that the self-knowledge premise of the transcendental argument is supposed to be safe from sceptical criticism. That is, it is supposed to be a premise that even a sceptic should accept. However, as Brueckner highlights, a sceptical predicament easily arises for the self-knowledge premise of the Putnamian transcendental argument. It arises because the kind of epistemic support one would need in order to know that *I am thinking that I'm not a BIV* is the same type of epistemic support one would need in order to know that *I'm not a BIV*.

5. STROUD'S DILEMMA FOR TRANSCENDENTAL ARGUMENTS

In the previous section, we construed Putnam's argument as a transcendental argument. The basic objection to the argument was that it doesn't seem to put us in a position to know that we are not BIV's. Instead, it seems only to establish that the sentence "I am not a BIV" expresses a true proposition, but the arguer is not in a position to know which true proposition that sentence expresses. As Brueckner (1986, 2010) expressed the point, it seems that the transcendental arguer would have to know that s/he wasn't a BIV in order to know that their utterance or representation of the sentence "I'm not a BIV" expresses the proposition *that I'm not a BIV*.

We can sharpen this objection as follows. One can grant that the Putnamian transcendental argument establishes the conclusion that our utterance or representation of *the sentence* "I am not a BIV" *is true*. The problem, however, is whether this helps us to undercut scepticism, since it is compatible with the truth of the sentence "I am not a BIV" that we *are* BIV's, so that the proposition our sentence expresses is that *I am not a BIV-in-the-image*. Putnam's transcendental argument from content externalism establishes that the sentence has to be true,

because it would be true whether we are in the normal, non-envatted world, or instead in the non-normal, envatted world. But what we want to know is which world we are in; or which true proposition *our* use of “I’m not a BIV” expresses.

Even if one is unmoved by Brueckner’s criticism of the Putnamian transcendental anti-sceptical strategy, Barry Stroud famously argued that transcendental arguments *in general*, despite appearances to the contrary, aren’t effective as anti-sceptical arguments. In brief, Stroud’s argument goes like this. In order for a transcendental argument to have anti-sceptical consequences, it has to presuppose idealism or verificationism (or some sort of anti-realist principle). But these doctrines are themselves *sufficient* to block scepticism, thus rendering the transcendental argument *superfluous*. The dilemma, then, is that either we dispense with idealism or verificationism, and thereby lose the anti-sceptical consequences of the transcendental argument, or else we keep idealism or verificationism, and then render the transcendental argument superfluous. In either case, the transcendental argument is itself ineffective.

Stroud’s criticism of transcendental arguments comes in three parts. The first part is the “hedging” move; the second part is the “bridge of necessity” move; and the third part is the “sufficiency” move. The transcendental arguer begins with a mind-to-world thesis, such as that:

- (T) A necessary condition for the possibility of having psychological states P with content S is that *S is true*.

For example, in Putnam’s argument, (T) would be construed as: a necessary condition for the possibility of thinking that *I’m not a BIV* is that *I’m not a BIV*. However, Stroud insists that the transcendental thesis can always be ‘hedged’ into a weaker mind-to-mind thesis:

The sceptic can always very plausibly insist that it is enough to make language possible [*or thoughts, beliefs, or representations of certain types*] if we believe that S is true, or if it looks for all the world as if it is, but that S needn’t actually be true. (Stroud 1968, 255)⁸

So, applying the hedging move to the transcendental premise of Putnam’s argument, we get:

- (T1*) A necessary condition of the possibility of thinking that *I am not a brain in a vat* is that *I must not seem to be a brain in a vat*.

With (T1*) replacing (T1), the most that a proponent of Putnam’s transcendental argument could infer is (TC*): that I must not *seem* to be a BIV. But it’s not clear how that conclusion could have any anti-sceptical import. For the sceptic can plausibly agree with (TC*), and argue that: “it’s true that you can think that you’re not a BIV, and it’s true that a necessary condition of your thinking that you’re not a BIV is that you must not seem to be a BIV. And I

grant that you don't seem to be a BIV. What I'm suggesting is that, although you don't seem to be a BIV, you nevertheless *don't know* that you're not a BIV".

Now the bridge of necessity move comes into play on (T1*). The argument here would be that in order for the proponent of the transcendental argument to bridge the explanatory gap from *seeming not to be a BIV* to *not being a BIV* one would need to invoke an idealist principle or a verificationist principle (or more broadly, an anti-realist principle). Why? *Prima facie* at least, it's hard to see how there could be *necessary* connection between *seeming not to be BIV* and actually *not being a BIV* that wasn't anti-realist friendly. After all, the BIV hypothesis is set-up so that how everything perceptually and introspectively seems to us to be is radically different from how things actually are. Because the world exists independently of us—*independently* of our conceptual schemes, language, and thought—it is possible for the world to be radically different from how it perceptually seems to us to be.

According to Stroud, an idealist principle or a verificationist principle is sufficiently anti-sceptical already, rendering the need for a transcendental argument superfluous. In particular, such an argument doesn't do any special anti-sceptical work that couldn't have been done by the idealist principle or the verification principle on its own. As Stroud puts the point more generally:

Even if we can allow that we can come to see how our thinking in certain ways necessarily requires that we also think in certain other ways [...] and we can appreciate how rich and complicated the relations among those ways of thinking must be, how can truths about the world which appear to say or imply nothing about human thought or experience be shown to be genuinely necessary conditions of such psychological facts as that we think and experience things in certain ways, from which the proofs begin? It would seem that we must find, and cross, a bridge of necessity from the one to the other. That would be a truly remarkable feat, and some convincing explanation would surely be needed of how the whole thing is possible. (Stroud 2000, 158-59)

Stroud's complaint here is that it's hard to see how the truth of certain contingent, non-psychological propositions about the world could be necessary conditions of the truth of certain psychological-propositions.

However, a proponent of the Putnamian transcendental argument can argue that content externalism provides the explanation of how that "bridge of necessity" is to be crossed without invoking an idealist or verificationist principle. After all, for the content externalist, the contents of some of our thoughts are in part *world-dependent*, whereas idealism entails that the world is *mind-dependent*. For the idealist, for it to be possible for anything to be a BIV is for it to be possible to represent something as being a BIV. The former metaphysical possibility depends on

the latter conceptual possibility. Content externalism allows for the combination of world-dependence *and* mind-independence.^{9,10}

6. INEFFABLE SCEPTICAL HYPOTHESES

Even if Putnam's transcendental argument doesn't fall victim to Stroud's general critique of the anti-sceptical import of transcendental arguments, it might fall victim to another general concern. Thomas Nagel (1986) presents a number of criticisms against Putnam's BIV argument,¹¹ but his core complaint is that the conclusion of this argument constitutes a *reductio* of Putnam's semantic externalism.

The crux of Nagel's argument is that it is manifestly obvious that it is possible that we're BIVs, so that any theory which implies that it is not possible is false. According to Nagel:

Such theories [*as semantic externalism*] are refuted by the evident possibility and intelligibility of skepticism, which reveals that by "tree" I don't mean just anything that is causally responsible for my impressions of trees, or anything that looks and feels like a tree, or even anything of the sort that I and other have traditionally called trees. Since those things could conceivably not be trees, any theory that says they have to be is wrong. (Nagel 1986, 73)

One reading of Nagel's criticism of the semantic externalist response to scepticism is that the possibility that scepticism is true provides at least *prima facie* grounds for thinking that what fixes the reference of our referring-expressions is not whatever is "causally responsible" for our typical use of those expressions.

More generally, Nagel claims that, if anything, Putnam's BIV argument actually *exacerbates* the sceptical difficulty, for all it actually shows is that one can't truly think that one is a BIV *even if one is a BIV*:

If I accept the argument, I must conclude that a brain in a vat can't think truly that it is a brain in a vat, even though others can think this about it. What follows? Only that I cannot express my skepticism by saying "Perhaps I am a brain in a vat." Instead I must say "Perhaps I can't even think the truth about what I am, because I lack the necessary concepts and my circumstances make it impossible for me to acquire them!" If this doesn't qualify as skepticism, I don't know what does. (Nagel 1986, 73)

Nagel's claim is thus that, far from resolving the sceptical problem, Putnam in fact simply trades a familiar form of scepticism with one that is new and mysterious, according to which one can't even be sure that one is able to think true thoughts about the nature of one's epistemic predicament.¹²

REFERENCES

- Boghossian, P. (1989). 'Content and Self-Knowledge', *Philosophical Topics* 17, 5-26.
- Brueckner, A. (1986). 'Brains in a Vat', *Journal of Philosophy*, 83, 148-67.
- (1992). 'Semantic Answers to Skepticism', *Pacific Philosophical Quarterly* 73, 200-19
- (1999). 'Transcendental Arguments from Content Externalism', *Transcendental Arguments: Problems and Prospects*, (ed.) R. Stern, Oxford: Clarendon Press.
- (2003). 'Trees, Computer Program Features, and Skeptical Hypotheses', *The Sceptics: Contemporary Essays*, (ed.) S Luper, Aldershot, UK: Ashgate.
- (2010). *Essays on Skepticism*, Oxford: Oxford University Press.
- (2011). 'Skepticism and Semantic Externalism', *Routledge Companion to Epistemology*, (eds.) S. Bernecker & D. H. Pritchard, New York: Routledge.
- (2012). 'Skepticism and Content Externalism', *Stanford Encyclopedia of Philosophy*, (ed.) E. Zalta, <http://plato.stanford.edu/entries/skepticism-content-externalism/>.
- Davidson, D. (1983). 'A Coherence Theory of Truth and Knowledge', *Kant oder Hegel?*, (ed.) D. Henrich, 428-38, Stuttgart: Klett-Cotta.
- (1990). 'Afterthoughts, 1987', *Reading Rorty*, (ed.) A. Malichowski, 136-68, Oxford: Blackwell.
- (1999). 'Reply to Barry Stroud', *The Philosophy of Donald Davidson*, (ed.) L. E. Hahn, 162-66, Chicago: Open Court.
- Davies, M. (1998). 'Externalism, Architecturalism, and Epistemic Warrant', *Knowing Our Own Minds*, (eds.) C. MacDonald, B. Smith & C. Wright, 321-61, Oxford: Oxford University Press.
- Dretske, F. (1970). 'Epistemic Operators', *Journal of Philosophy* 67, 1007-23.
- Forbes, G. (1995). 'Realism and Skepticism: Brains in a Vat Revisited', *Journal of Philosophy* 92, 205-22.
- Kallestrup, J. (2012). *Semantic Externalism*, London: Routledge.
- McKinsey, M. (1991). 'Anti-Individualism and Privileged Access', *Analysis* 51, 9-16.
- Nagel, T. (1986). *The View from Nowhere*. Cambridge: Cambridge University Press.
- Nozick, R. (1981). *Philosophical Explanations*, Oxford: Oxford University Press.
- Nuccetelli, S. (ed.) (2003). *New Essays on Semantic Externalism and Self-Knowledge*, Cambridge, MA: MIT Press.
- Putnam, H. (1981). *Reason, Truth, and History*, Cambridge: Cambridge University Press.
- (1994). 'Comments and Replies', *Reading Putnam*, (eds.) P. Clark & B. Hale, 242-95, Oxford: Blackwell.
- Pritchard, D. H. (2002a). 'McKinsey Paradoxes, Radical Scepticism, and the Transmission of Knowledge across Known Entailments', *Synthese* 130, 279-302.
- (2002b). 'Recent Work on Radical Skepticism', *American Philosophical Quarterly* 39, 215-57.
- (2002c). 'Resurrecting the Moorean Response to the Sceptic', *International Journal of Philosophical Studies* 10, 283-307.
- (2005). *Epistemic Luck*, Oxford: Oxford University Press.
- (2008). 'McDowellian Neo-Mooreanism', *Disjunctivism: Perception, Action, Knowledge*, (eds.) A. Haddock & F. Macpherson, 283-310, Oxford: Oxford University Press.
- (2012). *Epistemological Disjunctivism*, Oxford: Oxford University Press.
- (2013). 'Davidson on Radical Skepticism', *Blackwell Companion to Donald Davidson*, (eds.) E. LePore & K. Ludwig, 521-33, Oxford: Blackwell.
- (2015). *Epistemic Angst: Radical Skepticism and the Groundlessness of Our Believing*, Princeton, NJ: Princeton University Press.

- Pritchard, D. H., & Ranalli, C. (2013). 'Rorty, Williams and Davidson: Skepticism and Metaepistemology', *Humanities* 2, 351-68.
- (Forthcominga). 'On Metaepistemological Scepticism', *Traditional Epistemic Internalism*, (eds.) M. Bergmann & B. Coppenger, Oxford: Oxford University Press.
- (Forthcomingb). 'Scepticism and Disjunctivism', *Skepticism: From Antiquity to the Present*, (eds.) D. Machuca & B. Reed, London: Continuum.
- Stroud, B. (1968). 'Transcendental Arguments', *Journal of Philosophy* 65, 241-56.
- Wright, C. J. G. (1992). 'On Putnam's Proof That We Are Not Brains-in-a-Vat', *Proceedings of the Aristotelian Society* 92, 67-94.
- (2000). 'Cogency and Question-Begging: Some Reflections on McKinsey's Paradox, and Putnam's Proof', *Philosophical Topics* 10, 140-63.
- Zagzebski, L. (2009). *On Epistemology*, Belmont, CA: Wadsworth.

NOTES

¹ The rejection of closure as a response to BIV scepticism was famously advanced by Dretske (1970) and Nozick (1981). For two very different defences of an epistemological proposal according to which it is possible to know the denials of radical sceptical hypotheses, see Pritchard (2002*c*; 2005) and Pritchard (2008; 2012). For a general overview of the contemporary literature on radical scepticism, see Pritchard (2002*b*; 2015, part one).

² There are interesting questions to be asked about such non-permanent envatment cases, such as whether the BIV will eventually lose the capacity to think ordinary thoughts (e.g., about trees), though we will be setting these concerns to one side here. For a classic discussion of issues of this type—albeit focused not on BIV cases but rather parallel concerns which arise in ‘twin-earth’ cases—see Boghossian (1989).

³ Putnam (1994) has said that Wright put “more clearly than I myself did [...] the premises and the deductive steps involved in my argument.”

⁴ Compare with Wright:

“The worry is whether fully exorcising the sceptical doubt which brain-in-a-vat examples and other similar stories raise would not require precisely the kind of identifying knowledge of content which semantic externalism itself proscribes.” (Wright 1992, 78-79)

⁵ This connects to a general worry which has been raised regarding content externalism, which is the extent to which it is compatible with the idea that we can have *a priori* access to our own mental states. For some of the core literature on this topic, see Boghossian (1989), McKinsey (1991), Davies (1998), Wright (2000), and the articles collected in Nuccetelli (2003). See also Pritchard (2002*a*).

⁶ For example, here’s how Putnam discusses his reasoning:

“[M]y procedure has a close relation to what Kant called a ‘transcendental’ investigation; for it is an investigation, I repeat, of the preconditions of reference and hence of thought—preconditions built in to the nature of our minds themselves, though not (as Kant hoped) wholly independent of empirical assumptions.” (Putnam 1981, 16)

⁷ Compare with the following formulation of argument, which highlights the potential ambiguity:

(T1) A necessary condition of the possibility of thinking that the sentence “I’m not a BIV” is true is that that sentence is true.

(T2) I can think that the sentence “I’m not a BIV” is true.

(TC) The sentence “I’m not a BIV” is true.

⁸ One might wonder whether the hedging move is fair to the proponent of transcendental arguments. After all, part of the Kantian project is show that there are interesting mind-to-world relations which can be known by *a priori* reasoning. So, it seems Stroud’s hedging move challenges a major part of the Kantian project. However, even Stroud’s challenge lacks adequate support, and is in this way unfair, the thrust of Stroud’s concern here is that the proponent of the transcendental argument *has* to reply in ways which remove the anti-sceptical efficacy of the transcendental argument. For example, she’ll have to say that the hedging the mind-to-world thesis to a mind-to-mind thesis isn’t enough to establish the transcendental conclusion, and this is where Stroud challenges the proponent of the transcendental argument to explain why this is so without invoking an idealist or verificationist principle. Cf. Brueckner (2010, 109).

⁹ Consider these remarks by Brueckner:

“Granted, if idealism is assumed, then the existence of various psychological facts will entail that a world of mountains, lions, and cars exists. And the fact that such a world exists will (at least at first blush) seem to be a non-psychological fact. However, the sort of transcendental argument I would like to investigate does not at all depend on idealism of any kind. According to the anti-sceptical content externalist, the existence of content-bearing mental states requires that non-psychological reality be a certain way. But its being that way is not constituted by some sustaining psychological reality.” (Brueckner 1999, 230)

However, one might worry that Stroud’s second concern arises here. The verificationist maintains that the sentence “I’m not a BIV” is true if and only if “I’m not a BIV can be verified. One might think that one implication of Putnam’s anti-sceptical argument is that “I’m not a BIV” is always verifiable, since it’s verifiable in English or in natural language more generally (e.g., verified as true) and verified as true in vat-English (or in envatted form of a natural language) as well. Whether this is a collapse into verificationism is unclear, however, because the proponent of Putnam’s transcendental argument can plausibly suggest that no *general* verificationist principle is being invoked. Instead, they can say that some *special* sentences turn out to be always verifiable as being true (e.g., sentences like “I’m not a BIV”).

¹⁰ Interestingly, treating Putnam’s BIV argument as a transcendental argument also draws it closer to Davidson’s (1983; cf. Davidson 1990; 1999) famous anti-sceptical strategy, particularly since Davidson also (at least in later work) described his anti-scepticism as being transcendental in form. The other obvious parallel, of course, is that both philosophers primarily motivate their anti-scepticism by appeal to a form of content externalism. Although it would take us too far afield to explore these issues here, we think there is much to be gained by examining Putnam’s

BIV argument and Davidson's anti-scepticism in tandem. For a recent discussion of Davidson's response to radical scepticism, see Pritchard (2013).

¹¹ For example, Nagel criticises Putnam's causal theory of reference, from which he argues that BIV's couldn't refer to trees (etc.). According to Nagel (1984, 72), while the sceptic might not be able to provide an explanation of how our terms refer without the referents existing, the sceptic is nevertheless "not refuted unless reason has been given to believe such an account impossible." For a defence of this criticism, see Zagzebski (2009).

¹² Thanks to Jesper Kallestrup and Sandy Goldberg.