How You Know You Are Not a Brain In a Vat

Alexander Jackson

Draft, 3rd July 2013

Abstract: A sensible epistemologist may not see *how* she could know that she is not a Brain In a Vat (BIV); but she doesn't panic. She endorses her empirical beliefs, and as that requires, believes that she is not a BIV. (She does not inferentially base her belief that she is not a BIV on her empirical knowledge—she rejects that 'Moorean' response to skepticism.) I propose that she *thereby knows* that she is not a BIV. I flesh out the proposal, drawing on the empirical literature on metacognition, and explain why it satisfactorily resolves the skeptical puzzle.

1. The puzzle facing a sensible epistemologist.

C.S. Peirce wisely commanded, "Let us not pretend to doubt in our philosophy what we do not doubt in our hearts." (Peirce 1992 p. 29) With that in mind, please answer the following questions.

> Might you be the plaything of Descartes' evil demon, systematically misled by the experiences the demon induces in you?
> Might you be a massively deceived Brain In a Vat (BIV), whose misleading experiences are produced by a computer simulation run by an evil scientist?

Hopefully you did not answer that you might be a BIV, and are hence currently suspending judgment on all empirical matters. Freaking out in that manner is not a sensible response to considering the skeptical scenario. Hopefully you did not answer that you might be a BIV, incoherently combining that attitude with insisting that you do have hands, it is a sunny day, etc. Hopefully you kept your empirical beliefs, and added the belief that you are not a BIV.

Every sensible person who considers the issue believes they are not a massively deceived BIV; I assume the reader is in the club. The philosopher who thinks they might be a BIV is like a patient suffering from Obsessive Compulsive Disorder who cannot leave

their house because they keep returning to check they really did turn the stove off.[1] The patient with OCD doubts their memories of checking the stove; their doubts are pathological, not epistemically virtuous. Similarly, it is pathological to really suspend judgement on whether the skeptical scenario obtains, and hence on whether you are in loving relationships with anyone, or interact with real people who should be treated with moral respect; such attitudes have crazy consequences for decision-making. If a contextualist about 'knows' and related words holds that there are special contexts in which one should accept that one "might be a BIV", they are recommending temporary insanity. Some people find it easy to *say* they might be a BIV; their attention should be directed towards the consequences for their attitude towards what appear to be other people, and then to Peirce's maxim with which we began. (Sometimes one needs to pull oneself together, and not let a feeling of anxiety overwhelm one. On the way to the airport, I have to tell myself that I don't need to check again that I have my passport. The same goes if one starts to worry that one might be a BIV.)

Every sensible epistemologist has already settled that they are not a BIV. Yet the skeptical puzzle continues to tax them. So the puzzle does not concern whether one is a BIV. I suggest it is to say *how* one can know one is not a BIV, or rationally believe it. It can seem that there is *no way* to know that one is not a BIV—all the suggestions in the literature as to how one can properly believe it seem wrong. But that result is unacceptable: believing that <one cannot know that one is not a BIV> rationally commits one to accepting that one might be a BIV, and hence to skepticism.

This paper aims to resolve the puzzle here for sensible epistemologists. We all take it as settled that one is not a BIV, that one has reliable perceptual faculties, and so on. I will use those materials to give an account we can contentedly endorse of *how* one knows that one is not a BIV. That solves the skeptical puzzle. The problem with existing accounts is that they seem implausible *to we sensible epistemologists*; the problem is not that they seem implausible to someone who thinks they might be a BIV.[2] The paper does not aim to

[1] de Sousa (2008 pp. 197–8) discusses this example.

[2] Pryor (2000 pp. 517–8) puts the point by saying that we should abandon the 'ambitious anti-skeptical project' of only using premises a skeptic would grant, and focus on the 'modest' one, using any premise *we* accept.

change the mind of someone who believes they might be a BIV (and is not resisting a strong inclination to believe they are not). Such a person is not to be persuaded by argument; they are to be treated in the same way as a patient with OCD—medically. The paper should change the mind of someone who is strongly inclined to believe they are not a BIV, but feels forced not to because they see no way they could know it.

The skeptical puzzle is so hard because it seems illegitimate to use one's empirical knowledge as evidence that one is not a BIV (which does not leave much to appeal to). In particular, the 'Moorean' solution seems unacceptable. We all agree that one has perceptual knowledge that there is a yellow surface over there, for example. For all p, p logically entails that <it is not the case that: p is false and one is deceived on that score because one is a BIV>. The Moorean tells us to *infer* that logical consequence from what one knows by perception.[3] But we find the recommended inference repugnant. We are told that one should first settle by perception that there is a yellow surface over there, and that matter having been settled satisfactorily, then infer that <it is not the case that: one is a BIV and there isn't a yellow surface over there>. But we think that one has not settled satisfactorily that there is a yellow surface over there, if it is still left open at that point in inquiry that one is a BIV deceived on that score. It would not be appropriate to use one's perceptual beliefs as an inferential basis for other beliefs until one has answered the skeptical worry. In other words, the Moorean says that one can know one is not a BIV *epistemically posterior* to one's perceptual knowledge; but that seems false to us sensible epistemologists. (This must not be confused with the claim that the inference will strike *a skeptic* as illegitimate, because they won't accept the premise. That's true, but is no reason to think Mooreanism does not solve the puzzle this paper addresses. Saying the Moorean solution 'begs the question' strikes me as ambiguous between the good objection and the bad one.)[4]

---

[3] E.g. Pryor (2004, forthcoming); Davies (2004); Tucker (2010).

[4] Pryor (2004 §§5–7, forthcoming §V) and Davies (2004) try to explain the mistake we allegedly make in rejecting the Moorean inference. Pryor observes that if one is unreasonably suspicious that one is a BIV, or that certain epistemologies of perception are correct, then one won't be rationally able to form perceptual knowledge, and so the Moorean inference won't be available. But a sensible epistemologist harbors no such

A second problem for Mooreanism is that <there is a yellow surface over there> *does not* entail that one is not a BIV. That there is a yellow surface over there doesn't even make it likely that one is not a BIV. So the Moorean inference can't be a way for one to know that one is not a BIV (as opposed to knowing that <it is not the case that: one is a BIV and there isn't a yellow surface over there>). But one needs a way to know one is not a BIV. It would be incoherent to form a perceptual belief that there is a yellow surface over there while thinking, "Maybe I'm a BIV, but at least I'm not a BIV who is deceived about whether there is a yellow surface over there!" So Mooreanism is, at best, a radically incomplete theory of how to answer skeptical worries.[5] There are possible replies to this objection; to my mind they are refuted by Matthew McGrath (forthcoming). This is not the place to explore the details. I just want the objection on the table, as it won't even be a prima facie problem for my proposal (§3).

I've described two problems for the Moorean account of how one can know one is not a BIV. There are other accounts in the literature, including 'abductivism'[6] and Crispin Wright's 'entitlement' theory (2004); I don't find them plausible either. I worry that all the existing accounts are unsatisfying. That is, they bite a bullet: they all rest on some claim that still seems false after the theorist has had a go at explaining our mistake to us. We need a new account—one that makes the initial mistaken intuition go away. I think the proposal presented in this paper has the desired effect.

---

unreasonable suspicions; so the source of our repulsion from the Moorean inference has not been identified. Davies observes that one can't make the Moorean inference when one is trying to 'settle the question' of whether one is deceived. But if one assumes that one knows one is not a deceived BIV, and wonders merely *how* one does, then Davies endorses the Moorean inference. But a sensible epistemologist is certain she is not a BIV; so Davies has not identified the source of our anti-Moorean reaction.

[5] I don't think the Moorean can respond that one must inspect one's body, and infer that one is not a BIV, before one can form any other perceptual knowledge. A principled Mooreanism must be grounded in a 'liberal' epistemology of perception, i.e. one which holds that perceptual knowledge does not automatically rest on any independent knowledge (Pryor 2000, 2004, forthcoming).

[6] Beebe (2009) includes an overview and references.

I will focus on articulating the proposal; I won't attempt to compare it in detail to the alternatives in the literature. My strategy can be outlined in four steps.

1. Every sensible person who considers the skeptical scenario comes to believe—in the same kind of way—that they are not a BIV.
2. Everyone who has considered and dismissed the skeptical scenario in the normal way thereby *knows* that they are not a BIV.

Certain ways of coming to believe that one is not a BIV, such as on the basis of a coin toss, are not legitimate. The way sensible epistemologists come to believe that they are not a BIV *is* legitimate—it produces knowledge.

To convince you of (2), I will describe the normal way of coming to believe one is not a BIV, and then defend the claim it is knowledge-producing. I will be guided by the following assumption.

3. The psychological and epistemological account of how sensible epistemologists know they are not a BIV should be an instance of a general account of how humans know they are not going wrong in some way, e.g. misremembering a particular fact.

I don't think (3) is obviously true. I think a theory that respects (3) will thereby be more plausible, and that the results of working under that assumption turn out attractively. The working hypothesis that (3) is true gives us a big head start in saying how sensible epistemologists form the belief that they are not a BIV, by allowing us to draw on the empirical work in psychology on how humans form beliefs that they are not going wrong in some way. That is:

4. Our account of how humans come to believe they are not going wrong in some way should be guided by the relevant psychology literature.

I develop the psychological account in §2, layering the epistemological account on top in §3. §4 explains why the proposal should resolve a sensible epistemologist's philosophical puzzlement. §5 examines how the proposal maps onto the question of whether one must know that one is not a BIV *epistemically prior* to knowing things by perception, or *epistemically posterior*. I reject both standard options. One endorses one's empirical beliefs, and knows that one is not a BIV, *at the same point epistemically speaking*, i.e. 'epistemically parallel'. §6 raises some questions for future work.

Some epistemologists assimilate BIV skepticism to the so-called 'lottery puzzle' (most obviously contextualists such as DeRose (1995), Lewis (1996) and Cohen (1999), but also Nagel (2011)). If I consider whether I might be ill next Tuesday, I will drop my belief that I will teach on that day; the lottery puzzle is generated by the ubiquity of such examples of dropping beliefs and retracting knowledge attributions. (See Vogel 1990, Hawthorne 2004.) I do not intend my solution to BIV skepticism to extend to the lottery puzzle. We don't react to considering the BIV scenario by dropping our empirical beliefs; so the two problems seem different. (Vogel 1990 gives further reasons for thinking the problems are distinct.) We cannot presume that a solution to the puzzle of BIV skepticism would also solve the lottery puzzle; each puzzle should be investigated independently. I think the two problems turn out to have different solutions—I favour a contextualist or relativist solution of the lottery puzzle.

2. The psychology of answering worries.

This section gives a psychological theory of how humans generally come to believe that they are not mistaken about a particular matter. I claim that the way a sensible epistemologist comes to believe she is not a BIV is an instance of that general theory.

One need not believe things to be the way they appear to one to be. I will refer to an appearance or seeming as a 'verdict'. A verdict is a personal-level token representation. If it is used to guide action and further belief-formation, playing the belief-role, it constitutes a belief (cf. Lyons 2009, pp. 70–2, 90–2). Hence it makes sense to say that a verdict is produced by a *belief*-forming faculty. A verdict can be perceptual, mnemonic, inferential, or of some other kind. (I say more about what makes a representation a verdict below.)

Humans can consciously ask themselves whether a particular verdict is correct or might be mistaken. Asking that question triggers a distinctive kind of cognitive procedure, which I call "self-conscious metacognition".[7] One conceives of the question as having mutually exclusive answers: is the verdict right, or might I be mistaken? To arrive at the answer that one is right is to "endorse" the verdict of one's belief-forming faculty. Arriving at the answer that one might be mistaken commits one to rejecting the verdict.

Let me describe my working example. Factual memory is a belief-forming faculty in the relevant sense: it produces occurrent beliefs, and we can assess its verdicts in self-conscious metacognition. I believe that Ulan Bator is the capital of Mongolia, though I don't recall how I learnt that fact, nor any supporting evidence. I can ask myself whether I am right, or whether I might be misremembering on that score. The process of asking and answering that question is an instance of self-conscious metacognition. To end up occurrently believing that Ulan Bator is the capital of Mongolia, I must answer that I am right, "endorsing" the verdict of memory. If I answer that I might be misremembering, I am committed to rejecting the verdict of memory. (Misremembering, i.e. the degradation of stored information, is one kind of worry about memory. Another is that things went wrong in the initial formation of the belief; such as that you formed the belief on the basis of a misprint in a magazine, or on the say-so of an unreliable friend.)

How do humans normally come to endorse a verdict, when they do so? How do I come to believe I am right that Ulan Bator is the capital of Mongolia, and that I am not misremembering? I start my answer by explaining the role accorded to a Feeling of Rightness (FOR) in the psychology literature on metacognition.[8] That literature focuses on the case of memory; Asher Koriat (2007) provides an excellent survey. Koriat (2011) extends the framework to cover certain perceptual judgements, such as which of two objects is larger. Valerie Thompson (2009) makes a compelling case that 'dual-process'

---

[7] Hilary Kornblith (2012) calls this process 'reflection'. He explains that the distinction between System 1 and System 2 processing cuts across the reflective–unreflective distinction (chapter 5).

[8] Joelle Proust (2010) relates empirical work on FORs to philosophical questions, but not to skepticism.

theories of *reasoning* should make use of the metacognitive framework developed for the case of memory.

Psychologists agree that verdicts are accompanied by a FOR of differing strengths. The idea is introspectively compelling. My memory that London is the capital of the UK is accompanied by a very strong Feeling of Rightness; my memory that Lima is the capital of Peru is accompanied by a weaker Feeling of Rightness. The notion of a FOR gives us a better grasp of what a verdict is: it is something that Feels Right to a certain degree.

Typically, the subject has no introspective access to what's responsible for the strength of the FOR. Psychologists come up with clever experiments that confirm theories about what's responsible. The standard view is that a FOR of a particular strength is largely the product of sub-personal monitoring of the process that generates the verdict. For example, the fluency of the processing responsible for the verdict affects the strength of the accompanying FOR. The details don't matter for our purposes; what matters is that the subject typically can't say why their verdict Feels Right. (Note the empirical literature often refers to the sub-personal processing which produces a FOR as 'metacognition'; that processing must be distinguished from what I've called self-conscious metacognition.)

The strength of a FOR affects behaviour. For example, it affects the subject's willingness to assert the content of their verdict (Koriat 2007 p. 309). More importantly for our purposes, the strength of the FOR affects whether the subject will continue to investigate the matter. A weak FOR causes subjects to continue to investigate. A verdict rendered with a strong FOR causes subjects to close investigation into the matter, treating the question as settled. For example, I might ask a friend with a smart phone to google the capital of Peru to check I'm right; I'm not going to ask someone to check that London is the capital of the UK—that matter is settled. Thompson remarks the same goes in the case of reasoning (2009 p. 175): "It is this … Feeling of Rightness (FOR), that is the reasoner's cue to look no further afield for the answer."

Sometimes the appreciation that the matter is settled takes the form of a Judgement of Rightness (JOR). A JOR is what I called the 'endorsement' of a belief. A JOR is a belief, a conceptual representation, with the content that (for example) one is

right that Ulan Bator is the capital of Mongolia. A FOR is not a belief, but rather "an affective response" (Thompson 2009 p. 181). If one asks oneself the question, a strong FOR accompanying one's verdict will typically cause one to judge that one is right, i.e. form a JOR.

> In most cases, it is assumed that the FOR will be a sufficient basis for judgement, such that one's JOR is completely determined by the strength of the FOR with little, if any, conscious effort. (Thompson 2009 p. 181)

Admittedly, a strong FOR does not always cause a JOR; beliefs about one's competence can also cause a subject to make or withhold a judgement that they are right about a given matter. Koriat calls a JOR based on a FOR an "experience-based metacogntive judgement", and a JOR based on prior beliefs "information-based" or "theory-based" (Koriat 2007 pp. 295–301, 313–4); I will use this terminology below. Presumably, some metacognitive judgments are based on both kinds of cue.

> In sum, it is proposed that the JOR, like other memory-based metacognitive judgments, is multiply determined by both implicit and explicit cues. Implicit cues [i.e. FORs] are based on properties of the retrieval experience, such as its fluency; explicit cues are derived from beliefs that are accessible to conscious introspection. Note that, as is the case with decisions based on heuristic outputs, metacognitive decisions may be based on implicit cues, even when a more accurate judgment could be derived from explicit sources. (Thompson 2009 p. 183)

The important point for our concerns is that it is perfectly normal for someone to judge that they are right about some matter, based solely on a strong FOR.

Let us apply the above standard view to cases in which a subject engages in self-conscious metacognition, asking themselves whether they are right about a given matter, or might be mistaken. I claim that a strong FOR accompanying the questioned verdict will typically cause the subject to believe that they are right about the matter, endorsing their belief. The strong FOR causes the subject to close investigation on the question,

treating it as settled. For example, the strong FOR accompanying the verdict of memory that Ulan Bator is the capital of Mongolia causes me to believe that I am right about the matter, closing investigation. That seems exactly the job a strong FOR is meant to do. (As the case of OCD shows, it is bad to lack strong FORs that close investigation; see de Sousa 2008 pp. 193–4.)

The above claim goes a little beyond what is standard in the empirical literature, for the psychologists do not explicitly talk about cases in which the subject consciously asks themselves whether they are misremembering. I am assuming that considering the possibility that one is making a certain kind of mistake does not rob a strong FOR of its power to produce assent. That assumption is introspectively compelling: I can consider whether I am misremembering, and still endorse my verdict that Ulan Bator is the capital of Mongolia, when all that's consciously available is the strong Feeling of Rightness accompanying my answer. The assumption is also architecturally compelling: it would be absurd for humans to drop any belief the moment they consider the possibility that it is mistaken. (Of course, if it seems to a subject that their verdict is in danger of being misleading, they won't endorse it. But in the cases we are talking about, the subject merely considers the possibility that their verdict is misleading, without it seeming to them that it is in danger of being so.)

I've given a story about how a subject engaged in self-conscious metacognition ends up believing that they are right about the relevant matter. The next step is to extend the story to say how the subject believes that they are not misremembering (or mistaken is some other specified way). Unfortunately, the psychologists don't address this question, so we will have to go off-piste. Hopefully it will be clear how to maintain the spirit of the preceding account, according to which it is the function of the accompanying strong FOR to close investigation into the status of a verdict.

Let me start by arguing against the analogue of the Moorean response to skepticism. On that view, the subject *infers* that they are not misremembering, from the fact that they are right. I think that such a view is epistemologically implausible: it gets the epistemic order wrong, as Mooreanism does (§1). You can't endorse a mnemonic belief while leaving it till later to settle that you are not misremembering. The view under consideration is also psychologically implausible. It predicts that subjects will have a

strong inferential FOR that they are not mistaken. The subject will feel they have proved their conclusion, just as I do if I infer that Ulan Bator is not in Africa, from the fact that it is the capital of Mongolia. But, I suggest, that's not how it feels to believe that one is not mistaken. To endorse my belief that Ulan Bator is the capital of Mongolia, I must believe that I was not deceived by an unreliable testifier when I originally formed it. I do adopt the belief I was not thus lead astray; but it does not feel as if I have proved it. I do not have the strong inferential FOR predicted by the view under consideration. Consider what I will say if challenged as to how I know that I was not deceived by an unreliable source when I initially formed the belief. The view under consideration predicts that I will say that I am right that Ulan Bator is the capital of Mongolia, from which I can infer that I was not deceived when I formed the belief. That's not what I will say. I will be baffled as to how I know I was not deceived. I will try to ignore the question, insisting that I'm right that Ulan Bator is the capital, and that I was not deceived when I acquired that information.

In my view, that <I was not deceived when I formed the belief> does not itself Feel Right; rather it is a belief that I must adopt as part of endorsing the verdict that does Feel Right. The one FOR, which accompanies the verdict of memory, is enough to close investigation on the question, treating it as settled that Ulan Bator is the capital of Mongolia. One neither has nor needs another FOR to accompany the belief that one is not mistaken about the matter. (A verdict is a representation that is accompanied by a FOR of a certain strength; so one does not have a *verdict* that one is not mistaken.)

It will be important for explaining away the pull of skepticism (§4) that the belief that I am not deceived in some way does not itself Feel Right. While the details won't matter for our purposes, it would be nice to describe a plausible mechanism by which a subject comes to believe that they are not mistaken without inferring it (and hence generating an inferential FOR). I start with an analogy.

Decision-making is a matter of choosing between options, and at least some of the time, one picks one option and rejects the others *in one go*. Suppose one has difficulty picking between two attractive job offers. Eventually one comes to prefer the first job to the second. It would be strange to think the preference can only directly cause one to intend to accept the first job, which then causes one to intend to reject the second.

Rather, the output of the decision-task is both the intention to accept the first job and the intention to reject the second. The two intentions are outputted in one go. Analogously, I suggest that the task of answering the self-conscious metacognitive question outputs two beliefs in one go: that one is right, and that one is not misremembering.

The process of self-conscious metacognition is triggered by asking a question. I wonder whether I am right that Ulan Bator is the capital of Mongolia, or whether I might be misremembering, conceiving of the question as having those mutually exclusive answers. Conceiving of the issue in that way defines the decision-problem as a choice between two options, one of which must be accepted and the other rejected. (It is in setting up the decision-task in this way that I am sensitive to the incoherence of maintaining a belief while suspending judgement on whether it is mistaken.) The FOR accompanying the verdict of memory pushes me towards one answer to the metacognitive question, and nothing pushes me towards the other answer. The strong FOR causes the decision-task to be resolved by outputting two beliefs: the belief that I am right that Ulan Bator is the capital of Mongolia, and the belief that I am not misremembering. (I am happy for the reader to picture the decision-task being performed by a connectionist network, consisting of input units and two accumulator units that represent the two options to choose between (e.g. Bogacz et al. 2007). The strong FOR activates one accumulator, which then laterally inhibits the other. The system thus reaches a state in which it renders an output, namely two beliefs. The model should not be interpreted as taking the FOR to cause only the JOR, and the JOR to cause the belief that one is not mistaken.)

To summarize: I start by wondering whether I am right that Ulan Bator is the capital of Mongolia, or whether I might be misremembering, conceiving those possible answers as mutually exclusive. The strong FOR of the memory settles the matter: it causes the decision-task to be resolved by endorsing the verdict of memory *and* believing that I am not misremembering. Let's introduce some terminology for this cognitive procedure: I *believe* that I am not misremembering *in concert with endorsing my belief* that Ulan Bator is the capital of Mongolia. I conjecture that for any belief-forming faculty, such as

perception or inference[9], humans will usually form beliefs that they are not mistaken in this way.[10]

This psychological theory makes a prediction. We should not expect the subject to be able to articulate much of a defence of their belief that they are not mistaken, when they believe it in concert with endorsing the questioned verdict. They might believe introspectively that the verdict feels obviously right to them; but such a belief cannot do the causal (or epistemic) work of the feeling itself. We can only expect the following kind of insistence from the subject: "I *am* right that Ulan Bator is the capital of Mongolia; I am not misremembering." The prediction seems correct. (The fact that it seems correct for perception and inference too suggests there is a common type of cause for the judgement that one is right—a FOR.)

I claim that a sensible epistemologist *believes* she is not a BIV *in concert with endorsing* her empirical beliefs. The account is more complicated than for memory, because an indefinitely large plurality of beliefs is questioned, not just a single belief. The sensible epistemologist understands that all her empirical beliefs stand or fall together to the BIV worry. She understands the skeptical challenge as posing a question to which endorsing her empirical beliefs, and believing she might be a BIV, are mutually exclusive answers. She can't bring all her empirical (dispositional) beliefs all to mind at once, so she does not

---

[9] The case of inference requires a tweak. I might be significantly more confident that I am not inferring fallaciously than in my premises and hence my conclusion. So I should believe that I am not inferring fallaciously in concert with endorsing the conclusion reasoned to by *supposing* the premises true, rather than believing them so.

[10] Believing one is not mistaken in concert with believing that one is right is quite different from other cases of choosing between two beliefs, such as whether Obama is Republican or a Democrat. Firstly, one could suspend judgement on both options for Obama's affiliation; but one cannot stably suspend judgement on whether one might be mistaken—either one leaves the possibility open or one doesn't. That is, the metacognitive question presents one with a forced choice. Secondly, it is not necessary that one settle 'in one go' that Obama is a Democrat and not Republican: one can reason from one belief to the other. By contrast, one cannot infer between a JOR and the belief that one is not mistaken; see §5 below.

token FORs for all of them. Rather, she explicitly considers certain empirical beliefs, which are accompanied by strong FORs, understanding that they are arbitrary examples of the questioned class. Nothing significantly speaks in favour of her being in a danger of being a BIV. The strong FORs of the beliefs she explicitly considers cause her to: endorse those particular empirical beliefs while appreciating that such examples can be multiplied indefinitely, and to believe that she is not a BIV. She might express the resulting mental state by saying: "I am right that I have two hands, I am right that it is sunny outside, *and so on*; I am not a BIV." Note that the account of endorsing 'her empirical beliefs' is more subtle than one might expect. (I considersome variants of this account in §6.)

To summarize in Koriat's terms: the belief that one is not a BIV is an 'experience-based' metacognitive judgement, not an 'information-based' one. As predicted by the general theory, the sensible epistemologist will not be able to say much in defence of her belief that she is not a BIV. She will only be able to insist, "I am right about these empirical matters; I am not a BIV."

3. The epistemology of answering worries.

§2 gave a psychological theory of how humans generally come to believe that they are not mistaken about a particular matter, and applied that theory to the case of believing you are not a BIV. This section layers an epistemological theory on top of those psychological claims. §4 explains why the resulting view yields a satisfying resolution of the sensible epistemologist's philosophical puzzlement.

I propose that believing that one is not mistaken *in concert with endorsing* a questioned verdict can constitute knowledge that one is not mistaken. For example, consider the following underlying facts: I originally acquired the belief that Ulan Bator is the capital of Mongolia in a good way; I stored it normally; I now lack significant reason to think I am in danger of misremembering (i.e. I lack 'defeating' beliefs); and I believe I am not misremembering in concert with endorsing that mnemonic verdict. In virtue of those facts (and maybe some others), two epistemic facts hold: my belief that <I am right that Ulan Bator is the capital of Mongolia> constitutes knowledge; and my belief that <I am not misremembering> constitutes knowledge. Call this "consort knowledge" that I am not misremembering. It comes for free with my knowledge that Ulan Bator is the capital

of Mongolia. Of course, one can know that <one is not misremembering> in this way without knowing in virtue of what one knows it.

I endorse a related account of what makes the relevant beliefs rational. In virtue of having a verdict accompanied by a strong FOR, and lacking significant reason to think you in danger of going wrong, it is rational to believe that you are not mistaken in concert with endorsing the verdict.[11] It obscures things to suppress reference to the approved *way* of believing, simply saying that a verdict with a strong FOR makes it rational to believe that one is not mistaken. That formulation invites the question of *how* a verdict with a strong FOR makes *that* belief rational. By contrast, there is no pressing question as to *how* a strong FOR makes it rational to: believe one is not mistaken in concert with endorsing a belief. The normative principle gives the thumbs-up to what is, from the point of view of personal-level description, a basic and *sui generis* psychological operation. There is no intermediary whereby the FOR makes it rational to believe in concert. Maybe the normative principle can be explained by appeal to the aims of inquiry, such as knowing important truths and having a coherent basis for action; but the plausibility of the normative principle does not depend on giving such an explanation.

Let me deal with a distracting objection. Of course one could press the lottery puzzle (§1) at this point. Everyone knows they misremember from time to time; so it seems I cannot know that this is not such an occasion. So, allegedly, we should not be giving a theory of how I do know I am not misremembering. Let me make three points in response. Firstly, there is no counterexample here to the proposal for 'consort knowledge'. To press the lottery puzzle is to argue that normal adults *do* have significant reason to think they are in danger of misremembering, i.e. they do have 'defeating' beliefs. If the point is granted, then the account of consort knowledge correctly predicts that I do not know I am not misremembering. Secondly, the lottery puzzle simply distracts us from

[11] An exception: my (2011) argues that inferential verdicts are evaluable for rationality, unlike the verdicts of perception and memory. So when one infers fallaciously, the inferred belief, and the belief that one is not inferring fallaciously, are *not* all-things-considered rational. As I explain in that paper, it is still true that having an inferential verdict with a strong FOR, and no 'defeating' beliefs, rationally requires believing the fallacious conclusion.

what's at issue with BIV skepticism; for no such argument can be made that adults have significant reason to think they are in danger of being a BIV. (This might change if people were regularly envatted.) Thirdly, suppose that widespread skepticism is not the right response to the lottery puzzle, but rather we need a contextualist or relativist account of knowledge. Then pressing the lottery puzzle changes the context in a way that prevents us from addressing legitimate and interesting questions about how one knows various things, such as that one is not misremembering.

Back to the main thread: I claim a sensible epistemologist has consort knowledge that she is not a BIV. She lacks significant reason to think she is in danger of being a BIV; many of her empirical beliefs are well-formed in all respects other than the proper ruling out of the skeptical scenario; and she now believes that she is not a BIV in concert with endorsing her empirical beliefs (explicitly considering some of them). In virtue of those underlying facts, two epistemic facts hold: the explicit endorsements of beliefs that are otherwise well-formed constitute knowledge that she is right about the relevant empirical matter; and her belief that she is not a BIV constitutes knowledge. (The explicitly considered empirical beliefs need not constitute knowledge for the belief the one is not a BIV to do so; it is enough that one (dispositionally) knows something one would be happy to explicitly endorse as part of rejecting the skeptical scenario.) One has consort knowledge that one is not a BIV. Believing that you are not a BIV is just a side-effect of holding the empirical beliefs you should. Its epistemic standing derives from its connection to those empirical beliefs.[12]

One might object that strong perceptual FORs are not evidence that one is not a BIV, because a BIV would have them too. Luckily, I deny that a strong FOR functions as evidence that you are not mistaken, or is your reason for believing you aren't. It's a familiar 'externalist' claim that knowledge and justification do not require the possession of evidence; I say the same goes for rational belief. We should take seriously the distinction between inferential and non-inferential justification, reserving the claim that there is 'evidence or reasons for belief' for the inferential case. (I defend this position in

[12] Christopher Hookway (2003, 2008) argues that strong FORs indicate that skeptical scenarios "can be ignored" (2008 p. 63). I say that if they are not ignored, they can be easily ruled out.

my 2011.) Further, we can see directly that the objection under consideration cannot be re-stated so as to hit my proposal. What makes it rational for me to believe I am not a BIV *should* be something I have in common with a BIV; for it is also rational for a BIV to believe it isn't one.

'Consort knowledge' avoids both objections to Mooreanism described in §1. Firstly, it does not countenance *inferring* that one is not a BIV from any empirical knowledge. Secondly, consort knowledge that one is not a BIV is still made available by empirical knowledge the content of which does not make it likely that one is not a BIV, such as when one sees that there is a yellow surface over there. Consort knowledge is available because being a BIV is a skeptical worry, i.e. because there is a particular kind of incoherence in endorsing the questioned verdict without believing one is not a BIV.

I have not avoided Mooreanism by adopting the standard 'conservative' alternative, associated with Crispin Wright (2002, 2004, 2007). On that view, that one is not a BIV is a presumption that must be in place epistemically prior to forming perceptual knowledge. As I explain in §5, my view is that one settles that one is not a BIV at the same point epistemically speaking as one endorses one's empirical beliefs. (I think my proposal is more similar to Mooreanism than conservatism: it holds that the verdicts of our belief-forming faculties are inherently credible.) I suspect that Wright's view fails to plausibly avoid skepticism; but let me sketch another line of criticism. Wright holds that warrant to believe anything on the basis of perception is dependent upon the subject *already* having warrant to acccept that their perceptual verdict will be correct (2002 pp. 335–8, 2007 pp. 25–8). That fails to take 'experience-based' JORs seriously: it is the proper role of strong FORs, not some prior warrant, to cause us to take the senses to settle questions about the external world.

The existence of consort knowledge is inconsistent with traditional foundationalism, according to which the foundations do not derive their status from that of other beliefs, and the rest of our beliefs are inferentially supported by the foundations. On my proposal, one knows that one is not a BIV non-inferentially—it isn't Mooreanism. Yet the belief that one is not a BIV derives its status from its relation to one's empirical beliefs, so it would be misleading to say it is foundationally justified. It is because some

empirical beliefs are in all other respects fit to constitute knowledge (or be rational) that the belief that one is not a BIV gets the relevant positive status.

Nor does consort knowledge have much in common with traditional coherentism. According to traditional coherentism, inferential support flows in both directions between mutually supporting beliefs. On my proposal, inferential support flows in neither direction between the relevant beliefs. According to traditional coherentism, membership of the coherent set is what makes every belief in the set justified; there is no relevant asymmetry between the beliefs that cohere. On my proposal, the belief that one is not mistaken constitutes knowledge partly in virtue of the fact that the endorsed belief is in all other respects fit to do so, and not vice versa.

4. A satisfying solution to the puzzle of BIV skepticism.

This section explains why accepting that there is consort knowledge that one is not a BIV is a satisfying solution to skeptical puzzle. A satisfying solution identifies the mistaken intuition, makes it go away, and explains why we had it in the first place. I think the solutions in the literature, including Mooreanism, fail the second and third of those tasks. They still seem wrong when all is said and done; they bite a bullet. My proposal does not. A sensible epistemologist might initially assume that the way she came to believe she is not a BIV does not render it knowledge; but there isn't a residual intuition that there is no consort knowledge. I am not claiming that reflection makes it intuitively obvious that there is consort knowledge. I am claiming that reflection removes any intuition that there is no such knowledge. That allows us to decide on theoretical grounds whether there is. Consort knowledge is the only plausible story on the table about how you know you are not a BIV, it seems to me; and it is part of a plausible general story about how humans know they are not mistaken on a particular occasion. So I think we should accept that there is consort knowledge without any unease. All sensible epistemologists have consort knowledge that they are not a BIV; coming to believe that's how they know it should end their puzzlement.

Our initial tacit rejection of consort knowledge is easily explained away. When a belief Feels Right, the question of how one knows is easily dismissed; one insists that one just knows it, and does not doubt that one does. For example, one might not be able to

say how one knows that in Gettier's example (1963), Smith doesn't know that <either Jones owns a Ford or Brown is in Barcelona>; but one will insist it is obvious; one does not start to worry that there is no way one could know it. But suppose you ask yourself how you know you are not a BIV. That belief does not itself Feel Right. That stops you comfortably declaring that it is simply obvious. Instead, you bring your *beliefs* about your knowledge-forming capacities to bear on the question—you do some 'information-based' metacognition. There are a number of ways you believe you can know things, including perception, memory, and inference; but those ways do not provide a way to know you are not a BIV. So it seems to you that there is no way for you to know you are not a BIV. If the belief were accompanied by a strong FOR, you would add the method by which you come to believe you are not a BIV to your list of ways you know things; but in this case there is no FOR. There is no mystery as to why, given the absence of a FOR, people don't spot the possibility that the belief-forming method they are employing produces knowledge. Hopefully, the non-obvious argument of this paper will cause you to add consort knowledge to the list of ways you believe you can know things.

(The above explanation of the skeptical intuition depends on talking about verdicts and FORs, as opposed to appearances of differing strengths. The reader would object if I were just to claim that it does not seem to one that one is not a BIV. How come one believes it, if it does not seem true? I reply that a 'seeming' or verdict is a representation to which a FOR of a certain strength is attached. I gave a story in §2 about what causes one to believe one is not a BIV, without a FOR attaching to *that* belief.)

A philosopher might worry that my response to skepticism is unsatisfying, because it does not respect some 'internalist' constraint on a solution. Let me make three points in response. Firstly, the 'internalist' had better not be asking for evidence that one is not a BIV; that erroneously assumes that self-conscious metacognition must be 'information-based', not 'experienced-based'.[13] Secondly, it is not easy to give a plausible 'externalist' response to skepticism. For example, it seems natural for externalists to be Mooreans (Cohen 2002), but that's not a plausible response (§1). I think my solution is the first

[13] Ernest Sosa says that 'reflective endorsement' of one's beliefs must draw on a web of beliefs about the world and one's place in it (2009 pp. 151–2, 239–40). If he is talking about ordinary JORs, then he concedes too much to internalism (p. 44).

plausible one to be presented. It sure doesn't feel like I've cheated, ducking the hard question about skepticism. Thirdly, my solution is compatible with a popular construal of 'internalism': it says that the rationality of believing you are not a BIV supervenes on your current mental state.

I have now presented the core of my proposal. §5 draws consequences for 'epistemic priority' and the like. §6 raises some issues for future work.

5. The epistemic order.

I have argued that a sensible epistemologist believes she is not a BIV in concert with endorsing her empirical beliefs. This section investigates the upshot of that position for the 'epistemic ordering'. One appeals to the epistemic ordering when one complains that a logically valid inference from known premises is bad because one must believe the conclusion *epistemically prior* to believing the premise. The epistemic order forbids some way of *forming* a belief; it also places restrictions on how one structures one's (dispositional) beliefs at a time. For example, a subject may base one belief on another for a period of time; no inference need take place during that time, but the basing may violate the epistemic order in the same way the corresponding inference would. How one structures one's dispositional beliefs is itself a matter of one's dispositions.

We may describe how a person structures their beliefs at a time, without evaluating how they do so, by saying that: S believes that p epistemically prior to believing that q. We may then say whether it is permissible to do so. On this way of talking, Pryor holds that it is permissible for someone to believe things on the basis of perception epistemically prior to believing that they are not a BIV; and Wright holds that it is impermissible. (A reprehensible violation of the epistemic order need not prevent all the beliefs concerned from constituting knowledge. Someone who follows the Moorean recommendation still knows they have hands.)[14]

[14] Suppose Tim believes that Albany is the capital of New York solely on the basis of a report in an atlas (Cohen 1999 pp. 74–6). It is misleading to say that he *must* believe that the relevant sentence in the atlas is not a misprint epistemically prior to believing the geographical fact. There's nothing wrong with Tim having no (even dispositional) attitude on whether the sentence is a misprint, and believing that Albany is the capital. The

I claim that the sensible epistemologist endorses her empirical beliefs and settles that she is not a BIV *at the same point epistemically speaking*, and does so permissibly. This requires being more sophisticated about the epistemic order than is standard, in two ways.

Firstly, the possibility that some matters could be settled at the same point epistemically speaking has been neglected. Mooreans like Pryor say that one can settle that one is not a BIV epistemically posterior to one's perceptual knowledge, while their opponents (such as Wright) typically object that one must settle it epistemically prior. The possibility that two questions are to be settled at the same point, that they are to be settled 'epistemically parallel', is typically not considered.

When one believes that p epistemically parallel to believing that q, inferential support cannot flow in either direction, even when the relevant inference is logically valid. The sensible epistemologist can't infer in either direction between her belief that she is not a BIV and her empirical beliefs. Plausibly, any belief is epistemically parallel to believing that it constitutes knowledge. For example, I can't reason that since I know I have hands, it must be true that I have hands. Nor can I reason that since I have hands, and I so believe with good justification, and I am not in a weird Gettier situation, I know I have hands. Plausibly, that's because (if I consider both questions) I must settle at the same point epistemically speaking that I have hands and that I know I do. (The suggestion entails that if one considers whether one knows, then believing that p rationally commits one to believing that one knows it. It does not follow that knowing entails knowing that one knows.)

Epistemic parallelism must be distinguished from epistemic independence. I believe that 43+35=78, and that I am wearing black socks. The former belief is neither prior, posterior, nor parallel to the latter belief; they are 'independent'. When two beliefs are epistemically independent, the epistemic order permits using one, possibly with other premises, to inferentially support the other (and vice versa). But when two beliefs are epistemically parallel, it is impermissible to use one to inferentially support the other.

The sensible epistemologist believes she is not a BIV in concert with endorsing her empirical beliefs. The second deviation from orthodoxy about the epistemic order rejects

---

epistemic order does forbid his so believing while being (occurrently or dispositionally) agnostic as to whether the sentence is a misprint.

the assumption that that relations such as epistemic priority and epistemic posteriority only hold between *individual* beliefs. I propose that the sensible epistemologist's belief that she is not a BIV is epistemically parallel to her (dispositional) empirical beliefs *collectively*. (The empirical beliefs she explicitly consider do not get special treatment in this respect.) It is familiar that some relations involving pluralities apply collectively not distributively. For example, the table top is kept level by the table legs collectively; it is not kept level by any one of the legs. Again, the stones form a circle; it is not the case that each stone forms a circle. My proposal is that the sensible epistemologist endorses her (dispositional) empirical beliefs *collectively* at the same point epistemically speaking as she believes that she is not a BIV.

I must deny that for any of her empirical beliefs, the sensible epistemologist endorses it at the same point as she believes that she is not a BIV. The reason is that 'settling at the same point' is a transitive relation. So if two (intuitively independent) perceptual beliefs are each individually held epistemically parallel to the belief that she is not a BIV, then the two perceptual beliefs are held epistemically parallel to each other. But that's absurd: the epistemic order does not prohibit using one perceptual belief (together with other premises) to inferentially support another. We should deny that the sensible epistemologist's empirical beliefs can be individually compared in her epistemic ordering to the belief she is not a BIV. Rather, the only way to understand the structure of her beliefs is to see that she endorses her empirical beliefs *collectively* at the same point epistemically speaking as she believes she is not a BIV. I am not saying that relations such as epistemic posteriority never hold between individual beliefs; sometimes one does simply infer a conclusion. I am just saying that such relations do not exhaust the facts about how one structures one's beliefs, and is permitted to do so. The epistemic order is a partial order on pluralities of beliefs (including single beliefs).

6. Issues for future work.

This section raises two issues for future work. Firstly, let's reconsider whether the sensible epistemologist should believe <she is not a BIV> in concert with endorsing *all* her empirical beliefs.

Let me start by saying that while it is *better* to believe one is not a BIV in concert with endorsing a plurality of empirical beliefs, it is not necessary in order to rationally believe and know that one is not a BIV. Consider a subject who only recognizes the skeptical challenge to her perceptual belief that there is a yellow surface over there. It is not the case that she thinks her other empirical beliefs are immune to the challenge; she just doesn't consider whether the challenge generalizes. That is, she only asks herself whether she is right that there is a yellow surface over there, or whether she might be mistaken because she is a BIV. She endorses her belief that there is a yellow surface over there, and believes in concert with that endorsement that she is not a BIV. I am inclined to think this subject knows she is not a BIV, and her belief to that effect is rational.

However, the sensible epistemologist's collective response to her empirical beliefs is epistemically better. Firstly, she understands the full scope of the skeptical worry that she might be a BIV. Secondly, she appreciates that each of her empirical beliefs can play the same role in rejecting the possibility she is a BIV. She displays a much better understanding of the situation than the person who doesn't see the generality of the skeptical worry. Not all her empirical beliefs are otherwise well-formed; but as long as some of them are (whether explicitly considered or not), she thereby knows that she is not a BIV.

There remains the question of whether the sensible epistemologist should believe she is not a BIV in concert with endorsing *all* her empirical beliefs, or rather with endorsing a subclass of her empirical beliefs. Candidates for the relevant subclass include: her perceptual beliefs; her non-inferential empirical beliefs (including beliefs stored in memory); her non-inferential empirical beliefs plus those that constitute an explanatory understanding of some empirical phenomenon; and all of them. I don't have the space here to give the issue a proper treatment, and remain officially neutral.

Neither do I have space to address the objection that my proposal is subject to an analogue to the 'Problem of Easy Knowledge' for Mooreanism.[15] The problem is to come up with a principled account that does not legitimize consort knowledge in cases in which it is deeply implausible. For example, suppose one looks at a table that appears red, and one wonders whether it appears so because it is white and illuminated by a red light. The

[15] See Cohen (2002), Wright (2002, 2007), and McGrath (forthcoming), amongst others.

reader might find it implausible that one could simply believe that [the table is not white and illuminated by red lights] in concert with endorsing one's perceptual belief that [the table is red]. But why would that be illegitimate, if there is consort knowledge that one is not a BIV? I think these matters interact with our stances on the lottery puzzle (§1) and on the scope of perceptual content. The Problem of Easy Knowledge requires an extended treatment; I leave it to future work. Let me just say it bodes well that the account of consort knowledge was designed to deal with the common-or-garden case of knowing that one is not misremembering; it was not cooked up specially to deal with BIV skepticism.

7. Conclusion.

I've explained that you, you sensible epistemologist, already knew that you are not a BIV before reading this paper. Now you know *how* you know it. The skeptical puzzle is thus resolved.[16]

References

Beebe, James. 2009. "The Abductivist Reply to Skepticism," *Philosophy and Phenomenological Research* 79, pp. 605–636.

Bogacz, R., Usher, M., Zhang, J., & McClelland, J. L. 2007. "Extending a Biologically Inspired Model of Choice: Multi-Alternatives, Nonlinearity and Value-Based Multidimensional Choice," *Philosophical Transactions of the Royal Society B* 362, pp. 1655–70.

Cohen, Stewart. 1999. "Contextualism, Skepticism, and the Structure of Reasons," *Philosophical Perspectives* 13, pp. 57–89.

---- 2002. "Basic Knowledge and the Problem of Easy Knowledge," *Philosophy and Phenomenological Research* 65:2, pp. 309–329.

[16] I would like to thank the following for helpful feedback on this material: Alan Brinton, Andrew Cortens, John Hawthorne, Jonathan Jenkins Ichikawa, Peter Klein, Asher Koriat, Tania Lombrozo, Conor McHugh, Jennifer Nagel, Ram Neta, Ernest Sosa, and Jonathan Way. Special thanks to Stephen Crowley, Brian Kierland and Matthew McGrath for extensive comments.

Davies, Martin. 2004. "Epistemic Entitlement, Warrant Transmission and Easy Knowledge," *Proceedings of the Aristotelian Society Supplemental* 78, pp. 213–245.

DeRose, Keith. 1995. "Solving the Skeptical Problem," *Philosophical Review* 104, pp. 1–52.

de Sousa, Ronald. 2008. "Epistemic Feelings," in eds. Brun, Doguoglu, & Kuenzle, *Epistemology and Emotion*, Ashgate.

Gettier, Edmund. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23:6, pp. 121–3.

Hawthorne, John. 2004. *Knowledge and Lotteries*, OUP.

Hookway, Christopher. 2003. "Affective States and Epistemic Immediacy," *Metaphilosophy* 34:1/2, pp. 78–96.

---- 2008. "Epistemic Immediacy, Doubt and Anxiety: On a Role for Affective States in Epistemic Evaluation," in eds. Brun, Doguoglu, & Kuenzle, *Epistemology and Emotion*, Ashgate.

Jackson, Alexander. 2011. "Appearances, Rationality, and Justified Belief," *Philosophy and Phenomenological Research* 82:3, pp. 564–593.

Koriat, Asher. 2007. "Metacognition and Consciousness," in eds. Zelazo, Moscovitch & Thompson, *The Cambridge Handbook of Consciousness*, Cambridge University Press.

---- 2011. "Subjective Confidence in Perceptual Judgments: A Test of the Self-Consistency Model," *Journal of Experimental Psychology: General*, 140:1, pp. 117–139.

Kornblith, Hilary. 2012. *On Reflection*, OUP.

Lewis, David. 1996. "Elusive Knowledge," *The Australasian Journal of Philosophy* 74, pp. 549–567.

Lyons, Jack. 2009. *Perception and Basic Beliefs*, OUP.

McGrath, Matthew. Forthcoming. "Dogmatism, Underminers and Skepticism," *Philosophy and Phenomenological Research*.

Nagel, Jennifer. 2011. "The Psychological Basis of the Harman-Vogel Paradox," *Philosophers Imprint* 11:5.

Peirce, Charles S. 1992. *The Essential Peirce Volume 1*, eds. Nathan Hauser & Christian Kloesel, Indiana UP.

Proust, Joelle. 2010. "Metacognition," *Philosophy Compass* 5:11, pp. 989–998.

Pryor, James. 2000. "The Skeptic and the Dogmatist," *Nous* 34:3, pp. 517–549.

---- 2004. "What's Wrong with Moore's Argument?" *Philosophical Issues* 14, pp. 349–378.

---- Forthcoming. "When Warrant Transmits," in ed. Annalisa Coliva, *Wittgenstein, Epistemology and Mind: Themes from the Philosophy of Crispin Wright*, OUP.

Sosa, Ernest. 2009. *Reflective Knowledge: Apt Belief and Reflective Knowledge Volume II*, OUP.

Thompson, Valerie. 2009. "Dual-Process Theories: A Metacognitive Perspective," in eds. Evans & Frankish, *In Two Minds: Dual Processes and Beyond*, Oxford University Press.

Vogel, Jonathan. 1990. "Are There Counterexamples to the Closure Principle?" in eds. Roth & Ross, *Doubting*, Kluwer.

Wright, Crispin. 2002. "(Anti-)Sceptics Simple and Subtle: G.E. Moore and John McDowell," *Philosophy and Phenomenological Research* 65:2, pp. 330–348.

---- 2004. "Warrant For Nothing (And Foundations For Free)?" *Proceedings of the Aristotelian Society Supplemental* 78, pp. 167–212.

---- 2007. "The Perils of Dogmatism," in eds. Susana Nuccetelli & Gary Seay, *Themes from G.E. Moore: New Essays in Epistemology and Ethics*, OUP.