

Introspection and Inference*

Forthcoming in *Philosophical Studies*

Nicholas Silins, Cornell

Abstract

In this paper I develop the idea that, by answering the question whether p , you can answer the question whether you believe that p . In particular, I argue that judging that p is a fallible yet basic guide to whether one believes that p . I go on to defend my view from an important skeptical challenge, according to which my view would make it too easy to reject skeptical hypotheses about our access to our minds. I close by responding to the opposing view on which our beliefs themselves constitute our only source of first-person access to our beliefs.

Introduction

The thought that we have perceptual access to our mental states is historically common, and perhaps perennially tempting. As John Locke puts it,

“the other fountain, from which experience furnishes the understanding with *ideas*, is the *perception of the operations of our own minds* within us . . . This source of *ideas*, every man has wholly in himself: and though it be not sense, as having nothing to do with external objects; yet it is very like it, and might properly enough be called internal sense (*Essay*, bk. 2, ch. 1).”

Nevertheless, current work on self-knowledge often recoils from the idea that we have something like perceptual access to our mental states. For example, Richard Moran comments on

“the original embarrassment of the ‘inner eye’ and the concern that it cannot be cashed out as anything other than a misleading metaphor (2001: 13).”

To defend this perspective on self-knowledge, philosophers describe the metaphysical nature of perceptual states, and highlight the absence of counterpart states in the case of self-knowledge (Shoemaker 1996 is a classic discussion here).

By focusing on metaphysics, I believe philosophers have overlooked connections between the epistemology of perception and the epistemology of introspection. If we consider the structure of the justification of our introspective beliefs, we will see that in many cases it parallels the structure of the justification of our perceptual beliefs. Moreover, we will see that challenges which arise for the justification of perceptual beliefs arise in the introspective case as well.

The perspective about our access to our mental states I will defend has the following contours.

According to one component of the view, our access to our own mental states is sometimes non-inferential or basic. This is presumably part of what distinguishes the kind of access we have to our own mental states from the kind of access we have to the mental states of others, on the assumption that our access to the mental states of others is always inferential or non-basic.¹

According to another component of the view, our first-person access to our own mental states is nevertheless fallible. To put the point roughly, we can make justified mistakes about what mental states we are in. For example, if you are hallucinating that a sphere is present, with no indication that anything is going wrong, you can still be justified in believing that you *see that* a sphere is present, even though you don't genuinely see that a sphere is present.

According to the **moderate position** I will defend, our access to our mental states is sometimes both fallible and basic. This is not to say that our access is never infallible---perhaps your pain itself sometimes justifies you in believing that you are in pain. And this is not to say that our access to our mental states is never inferential---I will describe an example of inferential introspective access in section 4.

I will argue for the moderate position in more detail in what follows. For now, let me just observe that the view nicely allows that we can make mistakes about our own minds, while still

respecting the point that there is something special about our access to our minds as opposed to the minds of others.

On the moderate position, the justification of some of our introspective beliefs is structurally similar to the justification of our perceptual beliefs, at least as understood by philosophers such as Pollock 1974, Pryor 2000, and Peacocke 2004. According to these philosophers, we can have basic justification to believe that something is red when it looks red to us, even though it could look red to us in a case of illusion where the thing is not red.

There is an important challenge for the view that we have basic yet fallible justification for perceptual beliefs. In brief, the threat is that if we have basic yet fallible justification for beliefs regarding a subject matter, it will be too easy for us to reject skeptical hypotheses regarding our access to that subject matter (Cohen 2002, 2005, Wright 2002, Markie 2005). Given that it is presumably not so easy for us to reject skeptical hypotheses, it may look like we don't have basic yet fallible justification for perceptual beliefs after all. Following Cohen's dubbing of the "easy knowledge" problem, we can call this the "easy justification" problem.

The same problem will arise in the case of introspection, if we have basic yet fallible access to our mental states. I will develop the easy justification problem about introspection in much more detail in what follows, showing what the relevant skeptical hypotheses are. Let me now simply comment on why the problem is important.

The easy justification problem shows that much current literature on self-knowledge has overreacted against perceptual models, by overlooking important parallels with the case of the epistemology of perception. There is also current work in the tradition of inner sense theories, such as that by Gordon (1995), Nichols and Stich (2003) or Goldman (2006). This work will itself face a new challenge from the easy justification problem.

Solving the problem is of explanatory value as well. Ultimately we need an explanation of how we actually access our own minds in the way discussed by the moderate view. Prior to this explanatory question is the question of how such access to our own minds is so much as possible. Our solution to the puzzle will answer the fundamental explanatory question.ⁱⁱ

To make our discussion more concrete, I will focus on the test case of our access to our beliefs, and I will clarify the so-called "transparency of belief" discussed by Edgley (1969),

Evans (1982), Moran (2001), and others. The key idea here is that, in order to answer the question whether you believe that *p*, you can do so by answering the question whether *p*. As I will develop this idea, judgment is a guide to belief.

In section 1, I will begin by clarifying the so-called “transparency of belief”, which I believe to be misconstrued by Gareth Evans and others. This section will argue that the moderate view applies to the case of our introspective access to our beliefs. In section 2, I will formulate the challenge to the moderate view about introspection, as it arises in the case of our access to our beliefs. In section 3 I will respond to the challenge. Finally, in section 4 I will address the most important objections to my solution. The upshot will be an explanation of how it is possible for us to have fallible yet basic access to our beliefs.

Before we proceed any further, however, I should introduce some key terminology and clarify the key concepts for our discussion.

First, when I speak of judgments in what follows, I will focus on conscious judgments, mental acts which modify what it is like to be you at the time you perform them. For a paradigm case of judgment, consider the case of sincere assertion that *p*. When I speak of beliefs in what follows, I will focus on standing or dispositional beliefs, which need not modify what it is like to be you at the time you have them. For example, most likely a minute ago you had a standing belief that Paris is in France, even though you weren’t judging at that time that Paris is in France.

Second, let me fix on what I have in mind when I speak of “basic” or “non-inferential” access to one’s mental states. I will say that you have immediate justification to believe that *p* just in case you have justification to believe that *p*, and you have it in a way which does not rely on your justification to hold any other belief (see Pryor 2004). I prefer the term “immediate” to the terms “basic” and “non-inferential”, even though the latter terms may be more familiar to the reader. I prefer “immediate” because the term “non-inferential” places too much of an emphasis on psychological matters rather than on the underlying structure of justification, and because the term “basic” suggests too close a connection with foundationalism in epistemology.

Immediate justification is not defined in terms of how strong it is, but instead in terms of *how* one gets to have it. The crucial point is that one is not made to have immediate justification by one’s having justification for any other belief.

Third, let me clarify what I have in mind by “first-person” access to mental states. My focus will be on justification rather than knowledge. That’s because I am interested in what your epistemic position can be for a mistake about your own mind, and almost everyone agrees you can’t know a false proposition.ⁱⁱⁱ The term “access” is thus unhelpful, given that one can access only what is there. In what follows, I’ll say that you have **introspective** justification to believe that you are in a mental state M just in case you have justification to believe that you are in M, and you have it in a way such that no one else can have justification to believe that you are in M in that way.^{iv}

I characterize introspective justification by its “peculiarity” (Byrne 2005), rather than its strength, or the exact way in which it is acquired. The definition does not presuppose a perceptual model, and does not even presuppose that there is a unique source of introspective justification. It may well be that there are multiple ways of acquiring introspective justification, with no one account of them all (Prinz 2004, Boyle 2009).

Finally, let me turn to “fallible” access. I’ll say that a state j gives you fallible justification to believe that p just in case (i) j gives you justification to believe that p and (ii) it is possible for you to be in j while it is not the case that p. For example, your visual experience of a red tomato gives you fallible justification to believe that the tomato is red if it gives you justification to believe the tomato is red, while being such that you can have the experience even if the tomato isn’t red.^v

We can now state our key question as follows:

(Question): Does one ever have justification to believe that one is in a mental state where one’s justification is simultaneously introspective, immediate, and fallible?

According to the moderate view, the answer to this question is “yes”. At stake in the debate are the role of consciousness in introspection, and the relation between the epistemology of introspection and the epistemology of perception.

Let me now develop the moderate view and consider whether it is so much as coherent.

1. Transparency and Belief

I start with a famous passage by Gareth Evans:

in making a self-ascription of a belief, one's eyes are, so to speak, or occasionally literally, directed outward--upon the world. If someone ask me 'Do you think there is going to be a third world war?', I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question 'Will there be a third world war?' I get myself in a position to answer the question whether I believe that p by putting into operation whatever procedure I have for answering the question whether p . . . If a judging subject applies this procedure, then necessarily he will gain knowledge of one of his own mental state, even the most determined sceptic cannot find here a gap in which to insert his knife (Evans 1982, 225).

We can orient ourselves with the following claim:

(Slogan): You can answer the question whether you believe that p by answering the question whether p.

This is a helpful place to start, but we need much more detail.

Let's first consider *when* you can answer the question whether you believe that p by answering the question whether p.

It would be a mistake to say that your answer to the question whether p will always supply your answer to the question whether you believe that p. Consider that in some cases your answer to the question whether p will simply be "maybe", as when you ask yourself whether the number of stars is even. Here your answer to the question whether you believe that p will itself tend to be "no" rather than "maybe".

I will focus instead on cases where you answer the question whether p with a "yes" or a "no", and more generally on all situations where you judge that something is the case.

In particular, I'll focus on the position you are in with respect to your present beliefs once you judge that p, whether or not your judgment that p is the result of new enquiry into whether it is the case that p. After all, when your judgment is the result of new enquiry, it might well fail to be a good guide to your previous beliefs, but still can be a good guide to your current beliefs. And when your judgment is not the result of new enquiry, it is still a good candidate to be a guide to your current beliefs.^{vi}

I take myself to have departed here from Evans. His own emphasis is on “putting into operation whatever procedure I have for answering the question whether p”, which sounds like a matter of opening enquiry into whether p, but I think we should leave that emphasis behind.

Let me now turn to another important question: when you can answer the question whether you believe that p by answering the question whether p, *which* epistemic position are you in with respect to whether you believe that p? Evans makes the strong claim that his procedure is “necessarily” a source of knowledge. In section 2 we will see in detail why this is too quick. It is better in any case to work with the more cautious claim that one gets justification for a second-order belief through the procedure. If we only think about knowledge, we won’t even be able to ask whether answering the question whether p can put one in a good epistemic position for a false second-order belief.^{vii}

We need to make two further points about what kind of justification you acquire when you judge that p. First, we should allow that your justification to believe you believe that p might well sometimes be defeated, say by opposing testimony from your therapist. So we should focus on *prima facie* justification. Second, we should consider what epistemic position you are in when you answer the question whether p, whether or not you have taken advantage of that position. You can have justification by answering the question whether p even if you don’t yet have a second-order belief on the basis of the justification. In other words, I am focusing on what is known as “propositional justification” rather than “doxastic justification”. To illustrate the distinction, Holmes and Watson might each have propositional justification to believe that the butler did it, even though only Holmes believes with doxastic justification that the butler did it. Even if they have the same evidence, it can be that only Holmes has made proper use of the evidence in forming the belief that the butler did it, whereas Watson has simply jumped to the conclusion that the butler did it.

We can sum up the points made so far with the following formulation:

(Transparency, First Pass): When you judge that p, your judgment gives you (*prima facie*, propositional) justification to believe you believe that p.

Evans does not argue in favor of Transparency, or in favor of any claim in the vicinity. We can improve on his discussion by considering how to explain the defectiveness of “Moore paradoxical” judgments.^{viii} Consider judgments of contents of the following form:

(MP): P and I don’t believe that P.

Other things being equal, a judgment of a content of that form is irrational. To use the excellent terminology of Goldstein (2000), such judgments tend to be “Mooronic”.^{ix}

Why do judgments of MP tend to be irrational? The Transparency thesis provides a good explanation. The basic idea is that, when you judge a conjunction of the form MP, you flout the evidence provided by your judgment that p.

Let me now give the explanation in more detail. First consider that, when you judge [p and I don’t believe that p], you judge that p and judge that you do not believe that p. Next, if the Transparency thesis is true and you judge that p, you will then have *prima facie* justification to believe you believe that p. Therefore, if the Transparency thesis is true, whenever you make a judgment of the form MP, you will have *prima facie* justification to believe you believe that p, and yet you will judge that you do *not* believe that p. In other words, when you judge the MP conjunction, you judge the second conjunct while having *prima facie* justification to reject the second conjunct.

It need not be irrational to judge something you have *prima facie* justification to reject---you might have all things considered justification instead to believe the content you judge. In particular, there may well be cases where your *prima facie* justification from your judgment is defeated. But such cases are highly unusual---it is exceptional for you to get opposing testimony about what you believe from your therapist, or other defeating evidence bearing on what you believe. Your standard epistemic position is instead such that, if you were to judge in that position that [p and and I do not believe that p], your *prima facie* justification from your judgment that p would not be defeated. So, if the Transparency thesis is true, it is standardly irrational to judge contents of the form MP.^x

Since the Transparency thesis gives a good explanation why it is standardly irrational to judge contents of the form MP, we have some reason to believe the Transparency thesis. Here we have taken a valuable step beyond Evans' remarks.

We still need to say more about the kind of justification supplied by one's judgments. First let me discuss *introspective* justification. Here it is especially important to improve on the slogan with which we started. That's because, in certain circumstances, I can answer the question whether you believe that p by answering the question whether p. For example, I have the background information that, if your pants are on fire, you believe your pants are on fire. This background information permits me to answer the question whether you believe that your pants are on fire by answering the question whether your pants are on fire. But I still do not have introspective access to your beliefs. To further improve on the initial slogan, we need to make it explicit that one's judgments are a source of introspective justification regarding one's own beliefs:

(Transparency, Second Pass): When you judge that p, your judgment gives you introspective justification to believe you believe that p.

Next, I will argue that one's judgments are a source of *immediate* justification regarding one's beliefs. We should believe this hypothesis because it best explains how one's judgments are a source of introspective justification regarding one's beliefs. In the third-person case, when I answered the question whether you believed your pants are on fire by answering the question whether your pants are on fire, I relied on further background information. In the first-person case, when I answer the question whether I believe my pants are on fire by answering the question whether my pants are on fire, I do not rely on further background information. This contrast in terms of immediate justification nicely explains why I have introspective justification in one case although not the other.

The best rival explanation available is not as good. One might say that I rely in my own case on my background belief that, if I judge that p, then I believe that p, whereas in the case of others I rely on my different background belief that if p, then the other person believes that p. This rival explanation is less good because it "over-intellectualizes" the way in which judgment

is a guide to belief. In particular, judgment is a guide to belief in a wider range of cases than those handled by the alternative explanation. Consider people who have the concept of belief and the concept of judgment, but who are (for whatever bizarre reason) thorough eliminativists about judgment although not eliminativists about belief. Such people's judgments still put them in a good position to self-ascribe beliefs, whether or not they rely on a background belief that if they judge that p then they believe that p. Compare the way in which experiences can justify beliefs for a subject, even if the subject is unable to conceptualize that she has the experiences, or even if the subject is a thorough eliminativist with respect to experience. One need not rely on beliefs involving the concept of a state in order to gain justification from the state. The background beliefs brought in by the rival explanation are not needed.^{xi}

The hypothesis that your judgments give you immediate justification to self-ascribe beliefs gives the best explanation of how your judgments give you *introspective* justification to self-ascribe beliefs. So I will now give the following formulation, leaving reference to (*prima facie*, propositional, introspective) justification as understood:

(Transparency): When you judge that p, your judgment gives you immediate justification to believe you believe that p.

This is the lesson of reflection on Evans' discussion, whether or not he would have accepted it itself.^{xixiii}

2. The Easy Justification Problem

I will now address the central challenge to the Transparency thesis.^{xiv} To formulate the challenge, we need to consider how judgment is a fallible guide to belief, and we need to consider how skeptical hypotheses about introspection are relevant and coherent. It is perhaps unfamiliar to consider skeptical hypotheses concerning our beliefs about the internal world, we are much more used to doing so with respect to our beliefs about the external world. Be that as it may, there is a challenge to the possibility of some immediately justified beliefs about the internal world, one which parallels a challenge in the external world case.

I will start by articulating the challenge in both cases, and then will look more closely at the underlying motivation of the challenge.

In the external world case, the challenge concerns whether our experiences can provide immediate justification for beliefs about the external world (Cohen 2002, 2005; Wright 2002). The core worry is that, if our experiences do provide immediate justification for external world beliefs, then it will be too easy for us to reject skeptical hypotheses. In particular, perhaps we will be able to justifiedly reject skeptical hypotheses simply by performing inferences which correspond to the following argument:

I have hands.
If I have hands, then I am not a handless brain in a vat.
So,
I am not a handless brain in vat.

Such an inference is arguably not a way of becoming justified in rejecting the skeptical hypothesis. The skeptical hypothesis after all concerns the unreliability of experience, and we arguably cannot rely on experience itself to answer questions about its own reliability.

The attack on basic justification is that it would give us too much easy justification. According to the challenge, since we do not gain justification to reject skeptical hypotheses by performing such inferences, we do not gain immediate justification from our experiences for external world beliefs either.^{xv}

The challenge to the Transparency thesis is at root the same. The core worry is that, if judgment is a source of immediate justification regarding belief, then it will be too easy for one gain justification to reject certain skeptical hypotheses. Since it is in fact not so easy for one to gain justification to reject those skeptical hypotheses, judgment is not a source of immediate justification after all. In both cases, the idea is that we lack basic justification since we lack easy justification to reject skeptical hypotheses.

The outline of the challenge should now be clear. It will take some work to formulate the challenge to the Transparency thesis more thoroughly.

One question is whether judgment is a source of fallible justification (as I assume experience to be). Here the crucial point is that it is possible for one to judge that p while failing

to believe that p.^{xvi} In the simplest sort of case, one might judge that p while being in the process of acquiring a belief that p, without yet having acquired a belief that p. For a less simple sort of case, consider the following example: I judge that the flight leaves at 1:20, and then realize that I do not and did not believe that the flight leaves at 1:20. In such cases, your judgment that p is a kind of performance error which fails to reflect an underlying belief---“what was I thinking?”, you might go on to say.

One still might object in various ways to my description of the case.

To see that judgment that p is compatible with the absence of belief that p, it is crucial to remember that our focus is on conscious judgment as a guide to standing or dispositional belief. If one were to object by insisting that judgment that p is itself a species of conscious belief that p, one would simply speak past the key claim I am making. I am happy to grant that one could speak truly by saying “one believes that the flight leaves at 1:20” is satisfied in the example, provided that the use of “belief” is understood in terms of conscious belief.

Still, one might insist that when I judge that the flight leaves at 1:20, I at least briefly had a standing belief that the flight leaves at 1:20. I have no iron-clad argument against this position, but an advantage of my own is that it is more economical. To have a standing belief that p, I take it, one must have various behavioral and inferential dispositions. But there is nothing in the case to explain by attributing fleeting dispositions to the subject. So we need not attribute a brief standing belief that the flight leaves at 1:20.

Finally, one might grant I did not believe that the flight leaves at 1:20, but then add that I also did not genuinely judge that the flight leaves at 1:20. Here it is actually enough for my purposes if some state sufficiently similar to judging that p---in terms of what it is like to be in the state---can occur in the absence of belief. That’s because, in the cases in which one is in a judgment-like state without believing that p, I would instead hold that the judgment-like state gives one immediate fallible justification to believe one believes that p. It will be much simpler however to work with the case of judgment in what follows.^{xvii}

Judgment then is a fallible source of justification regarding belief. Let’s now consider skeptical hypotheses about particular cases. Roughly speaking, the skeptical hypotheses will state that one has judged that p, but will deny that one believes that p, and will sometimes go on

to explain the mismatch between judgment and belief. A detailed formulation of a relevant skeptical hypothesis will take more care. When one judges that *p* at a time *t*, and goes on a later time *t'* to self-ascribe a belief that *p*, one's later self-ascription will in effect say that [one believes that *p* at *t'*]. So if a skeptical hypothesis holds that one judged that *p* at a time while failing to believe that *p* at that time, the hypothesis will leave open that one's later self-ascription is true. In particular, the skeptical hypothesis will leave open the important possibility that, when one judged that *p*, one was in the process of acquiring a belief that *p*, without yet having acquired the belief that *p*. For a hypothesis to be a proper skeptical hypothesis, then, it will need to say that one judged that *p* at a time while failing to believe that *p* at that time as well as at the later time of the self-ascription.

Contrary to what Evans said, there is a place here for the skeptic to insert his knife.^{xviii}

Now that we have a skeptical hypothesis in mind, we can sharpen the challenge to the Transparency thesis. Consider the inference corresponding to the following argument, where the material in italics specifies the state which justifies belief in the first premise of the argument:

[JUDGMENT: I judge that p]

(BELIEF): I believe that *p*.

(TRANSITION): If I believe that *p*, then I didn't just [judge that *p* while failing to believe that *p* and while failing to be about to believe that *p*].

So,

(NO INTROSPECTIVE ERROR): I didn't just [judge that *p* while failing to believe that *p* and while failing to be about to believe that *p*].

Adapting some terms from White (2006), we can call the inference in question the "No Error" inference. The subject performing the inference is supposed to be justified by her judgment that *p* in believing that she believes that *p*. The central question is then whether she is able to become justified in believing the conclusion by performing the inference.

Now that we have our key example on the table, we can state the challenge to the Transparency thesis as follows:

(A): If judgment is a source of immediate justification regarding belief, then performing the No Error inference is a way of gaining justification to believe NO INTROSPECTIVE ERROR.

(B): Performing the No Error inference is not a way of gaining justification to believe NO INTROSPECTIVE ERROR.

So,

(C): Judgment is not a source of immediate justification regarding belief.^{xix}

Before developing my response to the challenge, let me discuss some of the considerations in favor of its key premises.

Let's start with the case for premise (A). To see a motivation for the premise, consider what might stand in the way of gaining justification through the No Error inference. A plausible barrier would arise if one were justified in believing the BELIEF premise only in virtue of having independent reason to believe the NO INTROSPECTIVE ERROR conclusion. In such a scenario the inference would be circular in a particular, suspect way. Your judgment would justify you in believing the BELIEF premise, to be sure, but only in virtue of your already having separate justification to believe the NO INTROSPECTIVE ERROR conclusion. Your inference would then "beg the question", in the sense that your justification to believe the premise would presuppose the conclusion. Since you would be in a position to perform the inference from a justified belief in the premise only if you already had independent justification to believe the conclusion, the inference arguably wouldn't be a way of acquiring new justification to believe the conclusion.

Now consider what happens if judgment is a source of immediate justification regarding belief. If judgment is a source of immediate justification regarding belief, there is no way that the No Error "begs the question" in the sense just introduced. After all, you would have justification to believe the BELIEF premise in a way which does not rely on independent reason to believe any other claim, let alone NO INTROSPECTIVE ERROR. Therefore, if judgment is a source of immediate justification regarding belief, then the barrier to the inference in question is removed. Now it's not clear what else might stand in the way of one's gaining justification through the No Error inference. If the barrier we have discussed is the only barrier to the success of the inference, as it so far seems to be, premise A must be correct.^{xx}

As far as premise (B) is concerned, I will set out a detailed argument for it in just a moment. For now we can rely on the plausible thought that, when x gives you justification to believe that p, x does not itself give you justification to believe that x is not a misleading source

of justification to believe that p. For a standard example, consider someone's testimony that p. This testimony might give you justification to believe that p, but presumably it does not give you justification to believe that the person did not give false testimony that p. To acquire justification to believe that the person did not give false testimony that p, when your sole justification to believe that p is his testimony, it does not seem adequate to reason as follows: "p and if p, then there's no way he gave false testimony that p. So, he didn't give false testimony that p." Similarly, when you rely on your judgment to answer a question concerning your beliefs, your judgment presumably does not certify that it itself is not a misleading source of justification regarding your beliefs.

3. The Solution

How might one respond to the challenge?

One response is to reject premise B. This is to maintain that one can acquire justification to reject a skeptical hypothesis by performing the No Error inference. Call this the *Moorean* response. In the epistemology of perception, philosophers such as Pryor (2004) and Markie (2005) pursue such a line. One problem for the approach is accounting for the suspect character of inferences like the No Error inference. To respond to this problem, the approaches try to explain away the appearance that the relevant inference fails to be a way of forming a justified belief, sometimes by focusing on the conceded fact that it is not a way of resolving justified doubts about its conclusion. It's not clear to me how this approach will succeed. One need not even be trying to resolve a justified doubt by performing the inference. Yet even when we acknowledge that the inference is not being performed in order to remove justified doubts, the inference still looks bad. So it's not clear how the appearance of failure is explained away.

There may be some further debunking explanation of why the No Error inference seems to fail. The main problem for the Moorean response is that there is no reason to look for one. The response is entirely unnecessary. We actually can hold on to the view that one has immediate (yet fallible) justification for one's introspective beliefs, without granting that the No Error inference succeeds.

The best response to the problem instead focuses on premise A. The solution is to separate issues about ordinary introspective beliefs from issues about anti-skeptical beliefs. In particular, we can acquire basic justification from our judgments for introspective beliefs without acquiring easy justification to reject skeptical hypotheses.

More specifically, the problem for (A) is that there is more than one way for an inference to fail to be a way of gaining justification to believe a conclusion. In particular, the No Error inference fails for broadly probabilistic reasons, rather than because of any considerations about “begging the question”.^{xxi}

A good way into the diagnosis is to note the following surprising fact about evidence:

(Fact): It’s possible that *e* is evidence for the proposition that *P*, the proposition that *P* clearly entails the (contingent) proposition that *Q*, and *e* fails to be evidence for the proposition that *Q*.

In a type of example used by Pryor (2004), that Jones’s pet is hairless is evidence that Jones has a Manx cat (a kind of hairless cat), and the proposition that Jones has a Manx cat clearly entails that Jones has a cat, even though that Jones’s pet is hairless is not evidence that Jones has a cat. Suppose you started out somewhat confident that Jones has a cat. In response to the evidence that Jones’ pet is hairless, you should decrease your confidence that Jones has a cat, yet still increase your confidence that Jones has a Manx cat.^{xxii}

A similar structure is present in the case of the No Error inference as well. To see why, consider the negation of No Introspective Error:

(INTROSPECTIVE ERROR): I just judged that *p* while failing to believe that *p* and while failing to be about to believe that *p*.

If Introspective Error is true, then one just judged that *p*. So Introspective Error predicts that one has judged that *p*. Now, other things being equal, if a hypothesis predicts that one will be in a certain state, one should increase one’s confidence in the hypothesis when one comes to be in that state. There doesn’t seem to be any exception to the principle in the case of the No Error inference. So it looks like one should increase one’s confidence in Introspective Error upon judging that *p*. Given this fact, it looks like performing the No Error inference won’t be a way to

gain justification to believe No Introspective Error, given that one's justification for the relevant premise requires one in fact to decrease one's confidence in the conclusion.

The argument can be set out in further detail, but I have reserved the discussion for Appendix 1.

Notice that, in giving the diagnosis of what's wrong with the inference, I am making an assumption about one's strongest evidence that one believes that p. I am assuming that one's strongest evidence that one believes that p is that one judges that p. In section 4, I will consider the point that this assumption does not hold for all cases. For now what is crucial is that the assumption holds for some cases. Let's focus on those cases for now.

Another complication is that our diagnosis concerns how one should respond to the evidential proposition that one judges that p. The diagnosis does not speak directly about how one should revise one's confidence in response to the state of judging that p. We need a way to generalize our story. To make the needed generalization, we may note that reflecting on what state you are in should not worsen your overall epistemic situation. Therefore, if you should raise your confidence in INTROSPECTIVE ERROR in response to the evidence that you judge that p, you should also raise your confidence in INTROSPECTIVE ERROR in response to your judgment itself.

3.3. Ramifications

We have a diagnosis of what's wrong with the No Error inference which does not involve the accusation that it "begs the question". The diagnosis focuses on the fallibility of one's justification to self-ascribe a belief, and in particular on the fact that one's having the justification is entailed by the negation of the No Error conclusion. No consideration about "begging the question" is invoked.

It looks like the culprit in the case of the No Error inference is fallibility rather than immediacy. In particular, notice that supposing one lacks immediate justification to believe the premises does not get the inference off the hook, so long as the crucial aspects of fallibility are preserved.

According to premise (A) of our puzzle, the inference does succeed in being a way to form a justified belief in its conclusion if one has immediate justification to believe the introspective premise. At a minimum, we have undermined the case in favor of this claim. According to the case for (A), the only explanation of why the inference fails is that it “begs the question”. But we have seen an alternative explanation of why the inference fails.

We also have good reason to doubt that (A) is true. It looks like we can give the probabilistic diagnosis of why the inference fails, without compromising the view that one has immediate justification for the introspective premise. Again, the diagnosis turns on the prediction of one’s evidence for the premise by the negation of the conclusion, rather than on any consideration about what makes the evidence for the premise work. So it looks like the No Error inference fails even if one does have immediate justification to believe the introspective premise.

Basic justification does not imply easy justification to reject skeptical hypotheses after all.

In the course of solving our puzzle, we have seen that a judgment that *p* is not apt to give one justification to deny that one is making an introspective error. This raises a difficult question: how then does one have justification to reject skeptical hypotheses about introspection? I won’t answer this question here. My concern is in particular with a challenge to the Transparency thesis, rather than with any more general skeptical challenge to introspective justification. My sympathies are with the approach on which we enjoy default entitlement to reject skeptical hypothesis about introspection---we may take it for granted that such hypotheses are false unless we are given significant reason to think they are true (see Burge 1996 for relevant discussion). But there is a further question about exactly how we should enjoy such a default entitlement. My main interest here is in how immediate introspective justification is so much as possible.

The main point for now is that the status of judgment as a guide to belief is one thing, and the status of the No Error inference quite another. We can accept that judgment is a source of immediate justification without licensing the No Error inference. So the central challenge to the Transparency thesis fails.

More generally, we can accept that a source of introspective justification is fallible, while still maintaining that it is a source of immediate justification. We can hold on to the attractive moderate view about introspective justification.

4. The Constitutivist Critique

One might object to my treatment of the No Error inference in various ways. Many of the objections arise with equal force in the case of introspective justification and in the case of perceptual justification.^{xxiii} I will focus on objections with special force in the case of introspective justification.

My own approach has focused on the role of judgments in the epistemology of second-order beliefs. By focusing on judgments I have developed a fallibilist epistemology of some of our introspective beliefs, an epistemology which parallels that of ordinary perceptual beliefs, at least as understood by philosophers such as Pollock 1974 and Pryor 2000. One might however insist that beliefs themselves justify second-order beliefs, to try to break my parallel between the epistemology of perception and the epistemology of introspection. “Constitutivist” views of introspection take such a line, by proposing very tight connections between first-order beliefs and second-order beliefs (see e.g. Shoemaker 1996, forthcoming and Zimmermann 2006). This line of objection is important. It can be used to challenge my claims about the status of judgment as a guide to belief, as well as my diagnosis of what’s wrong with the No Error inference.^{xxiv}

The line of objection can be developed in quite different ways. One way to develop the constitutivist line of objection focuses on entailments going from first-order beliefs to epistemic states such as that of having justification to believe you have a given first-order belief. This is to emphasize “self-intimation”, to use Sydney Shoemaker’s term. A very different way to develop constitutivism focuses on entailments from epistemic states to beliefs. This is to emphasize infallibility.

I will first address issues about self-intimation, and then turn to issues about infallibility. I will argue that no constitutivist thesis is both well-motivated and incompatible with my own position.

Let's start by considering the following very strong claim:

(Strong Self-Intimation): If you believe that p, then the fact that you believe that p gives you introspective justification to believe you believe that p.

Strong Self-Intimation focuses on an entailment from one's having a belief to its giving one introspective justification for a next-order belief. Given the emphasis on what happens when you do have a given belief, the thesis leaves open what happens when you don't have a given belief. The thesis thus leaves open what sorts of justification one can enjoy for a false second-order belief.

The thesis leaves too much open to be a threat. To include beliefs as sources of introspective justification is not to exclude judgments. So the thesis allows that judgment provides immediate justification for a second-order belief.

Strong Self-Intimation does affect how one should regard the No Error inference. In fact, the following much weaker claim does as well:

(Weak Self-Intimation): It's sometimes the case that (i) you believe that p and (ii) the fact that you believe that p gives you introspective justification to believe you believe that p.

If either Strong or Weak Self-Intimation obtains, we arguably need to give a non-uniform treatment of the No Error inference. Contrast the following two cases:

Case 1: You judge that p but don't believe that p.

Case 2: You judge that p and believe that p.

As far Case 1 is concerned, the probabilistic diagnosis we gave goes through. However, Case 2 will require a different treatment. In such cases, one will sometimes or always enjoy a source of justification such that one has it only if one does believe that p. The justifier in question will then be incompatible with the INTROSPECTIVE ERROR hypothesis, as opposed to being predicted by the INTROSPECTIVE ERROR hypothesis. So then the probabilistic treatment of

the No Error inference would be blocked, at least as far as some instances of Case 2 are concerned. This does not guarantee that the No Error inference will succeed in such cases--- perhaps there will be some other barrier to its success.^{xxv} However, it will much more plausible to hold that the inference is a way of gaining justification to believe NO INTROSPECTIVE ERROR in such cases.

Since my main concern is to defend judgment as a source of immediate justification, I am not concerned to reject the current proposals. They are actually compatible with my main claims. There will still be cases of type 1, and we will need the probabilistic diagnosis to account for them. It is no matter if there is a different story for cases of type 2.

I will now turn to a much more threatening line of objection. Here the focus is on infallibility rather than self-intimation.

The most extreme objection in the vicinity makes the following claim:

(Constitutive Justification): If you have introspective justification to believe you believe that p, the source of your justification for the higher-order belief is your belief that p.^{xxvi}

CJ says that only beliefs provide introspective justification for second-order beliefs. This is an infallibilist conception of introspective justification. If CJ is true, one can't have a second-order belief which is both false and introspectively justified. Moreover, if CJ is true, judgments aren't a source of introspective justification, at least as according to the Transparency thesis defended in section 1.

I will argue that CJ is false, by considering a case of inferential justification.

Consider a runner who regularly confronts his Nemesis at track meets, and who almost always believes he will lose when he believes Nemesis will be in a given race. Being concerned with maintaining a winning attitude, he has learned by introspection that he has this tendency. Given this background, it is open to him to reason in the following roundabout way about what he believes:

(D) I believe that Nemesis is in the race.

(E) If I believe that Nemesis is in the race, then I believe that I will lose.

So,

(F) I believe that I will lose the race.

Through such reasoning, our subject can sometimes acquire an introspectively justified but false second-order belief. Consider an isolated case in which he does believe that Nemesis will be in the race, yet does not believe that he will lose, and broaches the question of whether he believes he will lose in the roundabout way outlined above. In such a case he can still have an introspectively justified belief in (D) and a justified (yet false) belief in (E). Since (D) and (E) jointly entail (F), and his belief in (F) is formed on the basis of reasoning from those justified beliefs, he plausibly will be introspectively justified in believing (F), despite the fact that (F) is false. CJ then looks to be wrong.

In response, one might insist that the subject is not *introspectively* justified in believing the conclusion, given that she is inferentially justified in believing the conclusion. This objection overgeneralizes from cases like the case of pain. When you are in pain, your introspective justification to believe that you are in pain most likely is immediate. But consider cases like that of hope or expectation. To know that you expect that p, rather than hope that p, you might need to reason about counterfactual situations in which you learn that p (Williamson 2000). In such a case your justified belief that you expect that p might inferentially rest on your justified belief that, were you to learn that p, you would not be surprised, but you would be disappointed. Here you would still have introspective justification to believe you expect that p. Indeed, such a case is arguably still a paradigm of introspective justification. Not all introspective justification is immediate justification.

We can strengthen the point that our subject is introspectively justified in her false belief. If you are justified in a peculiar way at the starting point of your inference, in a way unavailable to others, then you will be justified in a peculiar way (if at all) at the ending point of your inference. Given that your starting point is unavailable to others, no one else will be able to reach your ending point in the way you do. Since peculiar justification is sufficient for introspective justification, the “logic” of introspective justification has the following feature:

(Transmissibility): If you are introspectively justified in believing that p_1 , and you have a justified belief that q on the basis of competent reasoning from $p_1 \dots p_n$, then you are introspectively justified in believing that q .^{xxvii}

The case just given only concerns inferential justification. So the objector might naturally fall back to a claim restricted to immediate justification:

(Constitutive Justification 2): If you have immediate introspective justification to believe you believe that p , then what gives you justification for the second-order belief is your belief that p .

I take judgment to provide counterexamples to CJ2, but these cases will not be convincing to the opponent. The point I will like to press is that CJ2 needs a defense, one it does not yet have.

One option is to say that immediate justification is in general infallible. This view is difficult to defend. It might be tempting to think that, if a state gives one fallible justification to believe that p , then there is a “gap” between the obtaining of the state and the truth of the proposition that p , so that some intermediate belief will be needed so as to “bridge the gap” between them. But this line of thought is risky. Unless the intermediate belief is infallibly justified, the line of thought can be repeated concerning the justification of the intermediate belief. And since infallibly justified beliefs are in very short supply, the line of thought threatens to show that very few beliefs enjoy fallible justification. Since many of our beliefs are fallibly justified if justified at all, the reasoning will have skeptical consequences. We should avoid such overkill.

CJ2 needs an argument tailored to the case of introspection, no general considerations about immediate justification will do.

The best argument I know of starts by invoking the idea that *absences* of belief are self-intimating.^{xxviii}

(H) If you do not believe that p , then you have immediate introspective justification to believe you do not believe that p .

The argument then appeals to the idea that you can't have justification to believe both of two contradictory propositions:

(I) If you have justification to believe that p, then you do not have justification to believe it's not the case that p.

If both these claims are true, we get the following conclusion:

(J) If you do not believe that p, then you do not have immediate introspective justification to believe you believe that p.

(J) entails that, if you do have immediate introspective justification to believe you believe that p, then you do believe that p. This is not quite the same claim as CJ2. CJ2 specifies the source of your introspective justification, J does not.

I set aside the task of getting from J to CJ2, as well as the important question of how to defend the premises of the argument. I do so because the argument has a very serious problem: there is no uniform reading of its premises on which they are all true.

(H) is true only if it is read in terms of *prima facie* justification. When you don't believe that p, you might well still have all things considered justification to believe you do believe that p, for example from opposing testimony from your therapist. If we try to read (I) in terms of *prima facie* justification, however, we get a falsehood. You can easily have *prima facie* justification to believe each of two contradictory claims, the difficulty is just in having all things considered justification to believe both of them.

When we put together the readings of (H) and (I) which have a chance of being true, we only get the following invalid argument:

(H*) If you do not believe that p, then you have immediate introspective *prima facie* justification to believe you do not believe that p.

(I*) If you have *all things considered* justification to believe that p, then you do not have *all things considered* justification to believe it's not the case that p.

So,

(J) If you do not believe that p, then you do not have immediate introspective *prima facie* justification to believe you believe that p.

In sum, out of the various constitutivist claims one might advance against my position, they are all either false, or unsupported, or compatible with my position.

Conclusion

Whether or not there is an “inner eye”, there are important parallels between the epistemology of perception and the epistemology of introspection. In both cases we can gain a kind of justification which is immediate and fallible. In particular, a judgment can give one immediate justification to believe one believes that p, even though one can judge that p without believing that p. That is the proper understanding of the poorly understood “transparency of belief”.

I would argue that there is such a basic yet fallible structure in many other cases of introspection as well (which is not to say that all cases involve it). For example, consider factive perceptual states such as that of seeing that p, where of necessity one sees that p only if p. Someone can have introspective justification to believe that they are seeing that p even when they are not. Suppose someone is looking at the Muller-Lyer illusion for the first time, who has no indication that it is an illusion. She can still have introspective justification to believe that she is seeing that the two lines are different in length (to believe she’s seeing that L, let’s say). That’s because she will have justification to believe she is seeing that L, and she will do so in a way such that no one else can have justification, in that way, to believe she is seeing that L. Given that the lines are not different in length, however, she in fact is not seeing that L.

More controversially, I think that our subject will also enjoy immediate introspective justification to believe she is seeing that L. At a minimum, if you grant that her (non-factive) visual experiences can give her immediate justification to believe that L, you should also grant that her (non-factive) visual experiences can give her immediate justification to believe that she is seeing that L. Once you accept that her experience gives her immediate justification to believe some external world proposition that p, her experience should look like an equally good source of immediate justification to believe a mental proposition with implications about how the external world is.^{xxix}

For further examples, consider other factive mental states such as that of remembering that p. Or consider our access to relational mental states such as that of seeing o, where one sees o only if one has an appropriate causal relation to the thing.

To understand the epistemology of introspection, we must not look away from the epistemology of perception.

In each case, a powerful challenge arises to the claim that we can gain immediate yet fallible justification: the threat is that one will be able to gain justification in illegitimate ways to reject skeptical hypotheses. In each case a challenge about “easy justification” arises, so that we need an explanation of how it is so much as possible to gain justification in the way we do. It is this question about the very possibility of fallible introspective justification I have addressed here. In responding to the challenge, I hope to have shown that issues about immediate justification are quite independent from issues about anti-skeptical justification. Even when we lack easy justification for anti-skeptical beliefs, we can still enjoy justification for introspective beliefs which is fallible yet basic.

Appendix 1

Here I draw on White (2006) to give a sharper diagnosis of what’s wrong with the No Error inference. We can start by considering the following consequence of Bayes’ Theorem:

$$P(\text{INTROSPECTIVE ERROR}|\text{JUDGMENT}) = \\ \frac{P(\text{INTROSPECTIVE ERROR}) \times P(\text{JUDGMENT}|\text{INTROSPECTIVE ERROR})}{P(\text{JUDGMENT})}$$

Now, since INTROSPECTIVE ERROR obviously entails JUDGMENT,

$$P(\text{JUDGMENT}|\text{INTROSPECTIVE ERROR}) = 1$$

So,

$$P(\text{INTROSPECTIVE ERROR}|\text{JUDGMENT}) = \frac{P(\text{INTROSPECTIVE ERROR})}{P(\text{JUDGMENT})}$$

So far, we can see that the probability of INTROSPECTIVE ERROR conditional on JUDGMENT is the prior probability of INTROSPECTIVE ERROR divided by the prior probability of JUDGMENT. Thus, provided that the prior probability of INTROSPECTIVE ERROR is greater than 0 and the prior probability of JUDGMENT is less than 1, the probability of INTROSPECTIVE ERROR conditional on JUDGMENT will be greater than the prior probability of INTROSPECTIVE ERROR. After all, if the conditions on the prior probabilities are satisfied, the probability of INTROSPECTIVE ERROR conditional on JUDGMENT will be the result of multiplying the (non-zero) prior probability of INTROSPECTIVE ERROR by a number greater than 1, yielding a product greater than the prior probability of INTROSPECTIVE ERROR.

The needed conditions in fact are met. You weren't certain that INTROSPECTIVE ERROR is false, and you weren't certain that you would judge that p. So

$$\begin{aligned} P(\text{JUDGMENT}) &< 1 \\ P(\text{INTROSPECTIVE ERROR}) &> 0 \end{aligned}$$

Therefore,

$$P(\text{INTROSPECTIVE ERROR}|\text{JUDGMENT}) > P(\text{INTROSPECTIVE ERROR})$$

Here I assume that you should update in this case by conditionalization, and that JUDGMENT is your strongest new evidence in the case. It then looks like you should increase your confidence in INTROSPECTIVE ERROR in response to JUDGMENT, and so decrease your confidence in NO INTROSPECTIVE ERROR in response to JUDGMENT. If you should decrease your confidence in NO INTROSPECTIVE ERROR in response to JUDGMENT, JUDGMENT is not evidence in favor of NO INTROSPECTIVE ERROR.

We now need to turn our attention to the No Error inference itself. I take it that, if your evidence for the claim that you believe that p fails to be evidence for NO INTROSPECTIVE ERROR, you fail to gain justification to believe NO INTROSPECTIVE ERROR by performing

the inference. When JUDGMENT is your strongest evidence that you believe that p, you do not gain justification to believe NO INTROSPECTIVE ERROR by performing the No Error inference.

References

- Bilgrami, A. 2006: *Self-Knowledge and Resentment*. Cambridge: Harvard University Press.
- Boyle, M. 2009: "Two Kinds of Self-Knowledge", *Philosophy and Phenomenological Research*, 78: 133-64.
- Breueckner, A. 1988: "Moore Inferences", *Philosophical Quarterly*, 48: 366-9.
- Burge, T. 1996: Our entitlement to self-knowledge. *Proceedings of the Aristotelian Society* 96:91-116
- Byrne, A. 2005: "Introspection", *Philosophical Topics*, 33: 79-104.
- Cassam, C. 2007: *The Possibility of Knowledge*. Oxford: OUP.
- Cassam, C. 2010: "Judging, Believing and Thinking"
- Cassam, C. 2011: "Knowing What You Believe"
- Chalmers, D. 2003: "The Content and Epistemology of Phenomenal Belief", in Q. Smith and A. Jokic (eds), *Consciousness: New Philosophical Perspectives*. Oxford: OUP.
- Cohen, S. 2002: "Basic Knowledge and the Problem of Easy Knowledge", *Philosophy and Phenomenological Research*, 65: 309-29.
- 2005: "Why Basic Knowledge is Easy Knowledge", *Philosophy and Phenomenological Research*, 70: 417-430.
- Davies, M. 2004: "Epistemic Entitlement, Warrant Transmission and Easy Knowledge", *Aristotelian Society Supplementary Volume*, 78: 213-45.
- Edgley, 1969: *Reason in Theory and Practice*. London: Hutchinson.
- Evans, G. 1982. *The Varieties of Reference*. Oxford: OUP.
- Fernández, J. 2003: "Privileged Access Naturalized", *Philosophical Quarterly*, 53: 352-372.
- 2005: "Self-Knowledge, Rationality, and Moore's Paradox", *Philosophy and Phenomenological Research*, 71: 533-556.
- Gallois, A. 1996. *The World Without the Mind Within*. Cambridge: CUP.
- Gertler, B. forthcoming, "Self-Knowledge and the Transparency of Belief"
- Goldstein, L. (2000), 'Moore's Paradox', in P. Engel (ed.), *Believing and Accepting* (Dordrecht: Kluwer Academic Publishers), 65 – 92.
- Goldman, A. 2006: "Self-Attribution" in *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*, OUP, 223-57.
- Gordon, R. 1995: 'Simulation without Introspection or Inference from Me to You', in *Mental Simulation: Evaluations and Applications*, eds. T. Stone and M. Davies (Oxford: Blackwell, 1995), 53–67.
- 2007: "Ascent Routines for Propositional Attitudes", *Synthese* 159:151-65.
- Green, M. and Williams, J. 2007. *Moore's Paradox: New Essays on Belief, Rationality, and the First Person*. Oxford: OUP.

- Hájek, A. (2007): “My Philosophical Position Says ‘p’ and I Don’t Believe ‘p’”. In Green, M. and Williams, J. (eds) *Moore’s Paradox: New Essays on Belief, Rationality, and the First Person*. Oxford: OUP.
- Hazlett, A. forthcoming: “The Myth of Factive Verbs”, *Philosophy and Phenomenological Research*
- Heal, J. (2002). *Mind, reason, and imagination*. (Cambridge: Cambridge University Press)
- Horgan, T. and Kriegel, U. 2007: “Phenomenal Epistemology: What is Consciousness that We Know it So Well?” *Philosophical Topics* 17: 123-44.
- Klein, P. 1981: *Certainty: A Refutation of Skepticism*. U. of Minnesota Press.
- Markie, P. 2005: “Easy Knowledge”, *Philosophy and Phenomenological Research* (70): 406-16.
- McDowell, J. 1982: “Criteria, Defeasibility, and Knowledge”, *Proceedings of the British Academy* (68): 455-79. Also in J. Dancy, ed., *Perceptual Knowledge*, Oxford University Press, 1988.
- 1995: “Knowledge and the Internal”, *Philosophy and Phenomenological Research* (55): 877-93.
- Moran, R. 2001: *Authority and Estrangement*. Cambridge: Harvard University Press.
- Nichols, S. and Stich, S. 2003: “Reading One’s Own Mind”, in *Mindreading: An Integrated Account of Pretence, Self-Awareness and Understanding Other Minds*, OUP.
- Peacocke, C. 1999: *Being Known*. Oxford: OUP.
- 2008: *Truly Understood*. Oxford: OUP.
- Pollock, J. 1974: *Knowledge and Justification*. Princeton: Princeton University Press.
- Prinz, J. 2004: “The Fractionation of Introspection”, *Journal of Consciousness Studies* (11): 40-57.
- Pryor, J. 2000: “The Skeptic and the Dogmatist”, *Nous*, 32: 517-49.
- 2004: “What’s Wrong with Moore’s Argument?”, *Philosophical Topics* (14): 349-78.
- 2005: “There is Immediate Justification” in M. Steup and E. Sosa, eds., *Contemporary Debates in Epistemology*, Blackwell.
- forthcominga: “When Warrant Transmits”, available at < <http://www.jimpryor.net/research/papers/Wright.pdf>>.
- forthcomingb: “Uncertainty and Undermining”, available at < www.jimpryor.net/research/papers/Uncertainty.pdf>.
- Ryle, G. 1949: *The Concept of Mind*, London: Hutchinson.
- Schiffer, Stephen. 2004: “Vagaries of Justified Belief,” *Philosophical Studies*, 119: 161-84.
- Schwitzgebel, E. forthcoming: “Acting Contrary to Our Professed Beliefs, or the Gulf Between Dispositional Belief and Occurrent Judgment”
- Searle, J. 1983. *Intentionality: An Essay in the Philosophy of Mind*. CUP.
- Setiya, K. forthcoming: “Knowledge of Intention”
- Shoemaker, S. 1995: (1995), ‘Moore’s Paradox and Self-Knowledge’, *Philosophical Studies*, 77: 211 – 28
- 1996: *The First-Person Perspective and Other Essays*. Cambridge: CUP.
- forthcoming: “Self-Intimation and Second-Order Belief”
- Silins, N. 2008: “Basic Justification and the Moorean Response to the Skeptic”, *Oxford Studies in Epistemology: Volume 2*.

- forthcominga: “Judgment as a Guide to Belief”, in *Introspection and Consciousness*, eds. D. Smithies and D. Stoljar. Oxford: OUP.
- forthcomingb: “Seeing Through the ‘Veil of Perception’”, *Mind*.
- Sosa, E. 2003: “Privileged Access”, in *Consciousness: New Philosophical Perspectives*, eds. Q. Smith and A. Jokic. Oxford: OUP.
- Sutton, J. 2007: *Without Justification*. Cambridge: MIT.
- Tye, M. 2002. “Representationalism and the Transparency of Experience,” *Nous*
- Vahid, H. 2005: “Moore’s Paradox and Evans’s Principle: A Reply to Williams”, *Analysis*, 65: 337-41.
- Way, J. 2007: “Self-Knowledge and the Limits of Transparency”, *Analysis* 295: 223-30.
- Weatherson, B. 2008: “The Bayesian and the Dogmatist”, *Proceedings of the Aristotelian Society*, 107 (2007): 169-85.
- White, Roger. 2006: “Problems for Dogmatism”, *Philosophical Studies*, 131: 525-557.
- Williams, J. 2004: “Moore’s Paradox, Evans’s Principle, and Self-knowledge”, *Analysis*, 64: 348-53.
- Williamson, T. 2000: *Knowledge and its Limits*. Oxford: OUP.
- 2004: “Skepticism”, in F. Jackson and M. Smith (eds.) *The Oxford Companion to Analytical Philosophy*. Oxford: Oxford University Press.
- Wright, C. 2002: “(Anti)-Sceptics Simple and Subtle: Moore and McDowell”, *Philosophy and Phenomenological Research*, 65: 330-48.
- Zimmermann, A. 2004: “Unnatural Access”, *Philosophical Quarterly*, Vol. 54, No. 216 (July 2004), pp. 435-438.
- 2006: “Basic Self-Knowledge: Answering Peacocke’s Criticisms of Constitutivism”, *Philosophical Studies*, 128 (March 2006) pp. 337-379.

* This paper integrates my (2008) and (forthcominga), and inherits the debts of those papers.

ⁱ For defense of the view that we can have non-inferential or basic access to the mental states of others, see McDowell (1982).

ⁱⁱ For a valuable recent discussion of “how-possible” questions, see Cassam (2007).

ⁱⁱⁱ For dissent, Hazlett (2009).

^{iv} At a minimum, you have justification in a way such that no one else has the ability to have justification to ascribe a mental state to you in that way. I leave open whether there is any stronger sense in which no one else can have justification to ascribe mental states to you in that way.

^v According to Sutton (2007), there is no such thing as a justified false belief. His view is actually compatible with the view that we enjoy fallible justification for some beliefs. In order to enjoy fallible justification from a state to believe that p, the crucial point is that one could be in the state when it is not the case that p, never mind whether one could have justification from the state to believe that p when it is not the case that p.

^{vi} For further discussion of related complications, see Shah and Velleman (2003) and Boyle (2009).

^{vii} Given Evan's bold claim that his procedure is necessarily a source of knowledge, one might think that he is really concerned with access to conscious belief [i.e. judgment], rather than to standing or dispositional belief.

One reason why I take Evans to focus on dispositional belief is that he leads with the question "do you think that there will be a third world war?" I take this question to not best be read in terms of conscious judgment, since I take the proper wording for a question about judgment to be "are you *thinking* that there will be a third world war?" The question with which Evans begins is instead more naturally read as being about dispositional belief, and I take his further discussion based on consideration of that question to be concerned with dispositional belief throughout.

Another reason why I take Evans to focus on dispositional belief is his sentence in *Varieties of Reference* immediately following the passage I quoted in the main text:

"We can encapsulate this procedure for answering questions about what one believes in the following simple rule: when you are in a position to assert that p, you are *ipso facto* in a position to assert 'I believe that p' (Varieties of Reference, 225-6)"

I am often in a position to assert that p when I am not judging that p, for example throughout this month I am in a position to assert it is June, even when I am not judging it is June. I believe throughout the month that it is June, but I judge it is June only occasionally if ever. Further, I am in a position throughout this month to assert it is June, but I am in a position to assert I am judging it is June only occasionally. Given these points, the rule Evans proposes is more charitably understood in terms of dispositional belief. It is implausible that, whenever I am in a position to assert that p, I am in a position to assert I judge that p. It is much more plausible that, whenever I am in a position to assert it is June, I am in a position to assert I (dispositionally) believe it is June.

^{viii} Thanks to Alan Hájek for urging me to think about connections between the transparency method and "Moore's paradox".

^{ix} Why think there is anything rationally amiss with judging MP? One important clue is that although MP can be true, brief reflection indicates MP cannot be truly believed (Shoemaker 1995). The crucial point is not that MP can't be truly believed (presumably many beliefs in necessary falsehoods can be rational). The crucial point is that only brief reflection is required to learn that MP can't be truly believed---believing MP is problematic in an easily accessible way. Here is why. If I believe that [p and I don't believe that p], then I do believe that p, yet at the same time believe I don't believe that p (I assume here that believing [p and q] entails believing p and believing q). My belief in the first conjunct is sufficient for the falsehood of my belief in the second conjunct. So only a little reflection reveals that, if I believe that [p and I don't believe that p], I am mistaken in so doing. This suggests it is at least typically irrational to believe MP, and presumably also to judge MP.

For a useful overview of forms of "Moore's Paradox", and descriptions and explanations of their absurdity, see the introduction of Green and Williams (2007). For further discussion of Moore's paradox and issues about self-knowledge, see Shoemaker (1995) and (forthcoming).

^x The Transparency thesis says that judgments that p are themselves a source of justification to believe you believe that p, not just that you have justification from some source or the other to BBp when you judge that p. One might object that one does not need so strong a claim as the

Transparency thesis to explain what's wrong with Moore paradoxical judgments, perhaps a mere claim about correlation will do the job. In response, I would say if one accepts that there is a correlation between judging that *p* and having justification to BB*p*, one should explain this correlation. Transparency provides the simplest explanation of the correlation, by saying that the judgment that *p* is itself the source of the justification to BB*p*. So even if one uses a mere correlation thesis to explain the defectiveness of Moore paradoxical judgments, one still in the end should accept the Transparency thesis. (I'm grateful here for discussion with Jim Pryor).

^{xi} Moran himself seems to understand the transparency method as involving one's background beliefs about a suitable connection between judgment and belief. For discussion of this aspect of his account, see Cassam (2010, 2011).

^{xii} For a different take on the transparency method, in terms of considerations about "rule-following", see Byrne (2005) and Setiya (forthcoming). For criticisms of Byrne (2005) see Shoemaker (forthcoming) and Gertler (forthcoming).

A question I leave open is whether the transparency method can somehow be generalized beyond the case of belief. For discussion of the case of absence of belief, see Sosa (2003). For discussion of the case of visual experience, see Evans (1982), Byrne (2005), and Peacocke (2008). For discussion of still further cases, see Gordon (1995, 2007), Byrne (2005), and Way (2007).

^{xiii} On my view, the transparency method is a way of obtaining introspective justification or knowledge about what beliefs you have. My view isn't that the method is the only way of obtaining introspective justification/knowledge about what beliefs you have. Evans seems to think otherwise, given his remarks about what one "must" do to answer the question whether one believes that *p*. To see why there are arguably other sources of introspective justification, consider that you could have introspective justification to believe that *p* when you may never have judged that *p*. Perhaps your standing belief that *p* could give you introspective justification for a second-order belief, without doing so via an intermediary judgment (Zimmerman (2006).

^{xiv} A different challenge appeals to a constraint on immediate justification:

(Face Value Constraint): Necessarily, if a state *M* gives one immediate justification to believe that *p*, then *M* has the content that *p*.

The Face Value Constraint is false. One way to see this is by considering the case of visual experience. Your visual experience can give you immediate justification to believe that you are conscious, whether or not it has the content that you are conscious. In response, one might insist that, according to philosophers such as Searle (1983), experiences have contents which involve references to the experiences themselves. However, this is not yet to say that an experience of mine will have the content that *I* am having the experience. So the counterexample stands even if philosophers such as Searle are right.

One might wonder whether there are non-introspective counterexamples to the Face Value Constraint. I think there are, and I discuss them in my forthcomingb.

^{xv} I use the term "easy justification" to echo Cohen (2002)'s discussion of "basic knowledge and the problem of easy knowledge". It should be said that the talk of "easy justification" and "easy knowledge" is somewhat misleading. On Cohen and Wright's ultimate views, it is not especially

hard to reject skeptical hypotheses. They consider that one might be able to rationally reject skeptical hypotheses simply by taking it for granted they are false, in the absence of significant reason to think they are true. This is easy! The focus of Cohen and Wright is specifically on whether it is legitimate to reject skeptical hypotheses on the basis of Moorean inferences, rather than more generally on whether there are any easy ways to reject skeptical hypotheses.

^{xvi} For the view that judgment is insufficient for belief, see Peacocke 1999, for the denial, see Zimmermann 2006. For further discussion of judgment and belief, with useful further references, see Schwitzgebel (forthcoming).

^{xvii} Zimmerman (2006) considers the fallback position I just outlined, and objects to it as follows:

... if experience with the phenomenal character of genuine judgment is not sufficient for belief, we can have experiences with this phenomenal character that are not real judgments (for they don't initiate, sustain or accompany beliefs). If our second-order introspective beliefs are grounded in such judgment-like experiences, knowledge of our beliefs is not direct, but instead mediated by inconclusive inferential grounds or states of inner perception (367).

I see no reason to accept the dilemma he proposes. Judgment-like experiences arguably can provide inconclusive immediate justification, we need an argument that they can't.

^{xviii} One could simply consider a skeptical hypothesis to the effect that one judged that p while failing to believe that p at the later time of the self-ascription. This skeptical hypothesis would leave open whether one believed that p at the time of the judgment.

An advantage of this skeptical hypothesis is that it sidesteps issues about whether judging that p entails believing that p. We can all agree that it is possible for one to judge that p at a time and to fail to believe that at a nearby later time, simply through forgetting or some cognitive interference.

A disadvantage of the hypothesis is that it does not permit us to address issues of propositional justification in addition to doxastic justification. When we consider whether judging that p gives you immediate justification to believe you believe that p, we actually needn't consider the later time at which you actually form the self-ascription.

^{xix} A separate challenge departs more directly from the claim that, in order to believe that p, you must meet dispositional requirements which need not be satisfied when you judge that p.

According to the challenge, it simply follows that conscious judgment that p can't be an immediate guide to belief that p (see Cassam 2010 for a discussion in this vicinity). Since the challenge focuses on the insufficiency of judgment for belief, it seems to simply assume that a source of immediate justification for Q must suffice for the truth of Q. The assumption requires separate defense before the challenge can succeed (see section 4 for criticism of the assumption that immediate justification must be infallible).

^{xx} For arguments along these lines, focusing on the case of perceptual justification, see Davies (2004) or Pryor (forthcominga).

One might wonder whether "begging the question" really is problematic. Perhaps an inference can beg the question while still being a way for one to acquire justification to believe its conclusion. For present purposes, I will simply take it for granted that begging the question is problematic.

^{xxi} Here I draw on valuable work in White (2006). For related discussion, see also Williamson (2004), Schiffer (2004), or Cohen (2005).

^{xxii} See also the "clever car thief" case in Klein (1981: 33-36).

^{xxiii} According to one line of objection, the probabilistic diagnosis I gave is actually incompatible with the claim that one has immediate justification to believe the premise. I discuss this sort of objection in section 4 of my (2008). According to another, more radical, line of objection, the probabilistic diagnosis is mistaken (see Pryor forthcomingb or Weatherson 2008).

^{xxiv} A similar move can be made in the case of perceptual justification as well. For example, one might insist that the only experiences which justify one in believing that *p* are factive mental states of seeing that *p* (McDowell 1995). These states of seeing that *p* are such that one is of necessity in them only if *p*. Here there would be no question of being in the same state when one is in a skeptical scenario in which one's external world belief is false.

Although this move can indeed be made in the case of perceptual justification, I take the move to be much more plausible in the case of introspective justification.

^{xxv} For discussion of this sort of issue in the case of perception, see Wright (2005).

^{xxvi} For an endorsement of this claim, see Zimmerman (2006, section 8), and especially his remark that "when our second-order introspective beliefs are formed and maintained in a first-person way they are grounded in the very first-order mental states that make them true (370)." However, he sometimes makes a weaker claim: "if we have any false, justified beliefs about what we believe, the grounds for these beliefs will be different in kind from the grounds with which we hold our typical second-order introspective beliefs (371)." Here one could allow for an introspectively justified yet false second-order belief, and simply insist that such a belief would have a different ground from an introspectively justified true second-order belief.

^{xxvii} Compare: if you are *a posteriori* justified in believing that *p*₁, and you have a justified belief that *q* on the basis of competent reasoning from *p*₁ . . . *p*_{*n*}, then you are *a posteriori* justified in believing that *q*.

Contrast: if you are immediately justified in believing that *p*₁, and you have a justified belief that *q* *solely* on the basis of competent reasoning from *p*₁ . . . *p*_{*n*}, then you are NOT immediately justified in believing that *q*.

^{xxviii} I'm grateful here for discussion with Declan Smithies and Daniel Stoljar.

^{xxix} One might be reminded of the much discussed "McKinsey puzzle" here, and deny that one can have introspective justification for a belief which concerns the external world. Given that the source of the introspective justification in question is experience, however, we should cease to be puzzled by the case. It is not a case of somehow *a priori* knowledge of the world. There is nothing paradoxical about the view that experiences can be a source of justification for beliefs about the external world, or for beliefs about the mental world which have implications about the external world.

One might instead protest that an experience can give one immediate justification to believe that *p* only if it has the content that *p*, and insist that visual experiences don't have contents to the effect that one is seeing that *p*. For further discussion of this objection, see my forthcomingb.