1. Introduction

An important strand in contemporary discussions of self-knowledge draws from the following remark by Gareth Evans (1982, 225):

[I]n making a self-ascription of belief, one’s eyes are, so to speak, or occasionally literally, directed outward — upon the world.

On a more traditional picture, our capacity to know about the world is supposed to depend upon a distinct and privileged capacity to know our own minds. Evans proposes that we turn this picture on its head: on his view, knowing our own minds — at least insofar as our beliefs are concerned — involves only our world-directed cognitive faculties. Nothing along the lines of inner observation or perception is necessary.

My aim in this paper is to develop and defend an account of our knowledge of our conscious beliefs which follows this general approach. The basic idea is this. A subject’s conscious beliefs partly constitute her outlook on the world, or, as we might put it, they determine what the world is like for her. But a rational subject who is competent with the concept of belief grasps the distinction between her own take on the facts and the facts themselves: she realizes that it is one thing for \( p \) to be part of her own view of the world, and another for \( p \) to be the case. Thus, a rational subject with the concept of belief can step back and recognize her beliefs for what they are: she can self-ascribe them as her own beliefs. As I will argue, this procedure accounts for at least one fundamental way we have of knowing about our own conscious beliefs. Moreover, this procedure involves no inner observation (and further, as I will argue below, neither is it properly understood as inferential).

I claim no originality for this general strategy. Indeed, part of its appeal is that it seems intuitively so straightforward. The way I will develop the strategy, however, is importantly different from other such attempts. In particular, Evans’s followers have by and large failed to acknowledge the epistemological importance of the phenomenology of conscious belief — André Gallois (1996, 18), for instance, goes so far as to deny that propositional attitudes have any distinctive
phenomenology. My account, by contrast, rests on a specific view of
the phenomenology of conscious belief.¹

A different group of recent authors have proposed accounts of self-
knowledge that do rest on the phenomenology of conscious belief. In
my view, however, these accounts tend to go wrong either in their
account of what that phenomenology consists in, or in failing to give
a satisfactory explanation of the link between belief’s distinctive pheno-
menology and the epistemology of self-knowledge.² In my view,
belief is phenomenologically transparent, in the sense that in having
a conscious belief you are only subjectively aware that the world is a
certain way—namely, the way specified by the content of your belief.³

You need have no awareness of the belief itself as a mental particular.
In consciously believing that p you are aware of the content that p; but
you need not have any awareness of your belief as such.⁴

1. A suggestion similar to my own is floated by Martin (1998, 117), but it is not
developed or defended. Gallois (1996; 2010) and Byrne (2005; 2011) present
views motivated by similar considerations of rationality, but they ignore
the importance of phenomenology. Dretske (2012) sketches a view, similar
to my own but, somewhat surprisingly, presents it as a form of ‘conciliatory
skepticism’ about self-knowledge. As will become clear in what follows, I
think there is nothing skeptical about the view. A rather different view, also
inspired by Evans, is developed by Moran (2001). I will not be able to discuss
Moran’s views in this paper.

2. BonJour (2003) and Pitt (2004) argue that the phenomenology of conscious
belief plays a crucial role in our capacity to self-ascribe them, but, as I will
argue below, their account of that phenomenology is problematic. Pea-
ocke (1998), Silins (2012), and Smithies (2012) develop accounts on which
conscious judgment plays a justificatory role for higher-order beliefs, but in
my view their accounts are not satisfying because they do not explain how
the distinctive phenomenology of conscious belief or judgment—i.e., what
it is like to judge or consciously believe that p—underpins our capacity for
self-knowledge.

3. Or, at least, given that your belief might be false, it is for you as if you really
were aware that the world is that way. I will mostly avoid this cumbersome
qualification in what follows, but it should be taken as implied whenever I
use the factive locution ‘aware that...’ and its cognates.

4. We need to be careful here, because talk of awareness of content can mis-
lead by suggesting an analogy with familiar cases of awareness of an object.
It seems plausible that being aware of an object entails that you are aware
of some of its qualitative properties—in the familiar case of perceptual

There are certain events in consciousness, such as pains, whose
phenomenology seems at least prima facie to consist in certain intro-
spectable qualitative properties—for example, it seems at least prima
facie plausible to construe the phenomenology of pains in terms of their
quality of painfulness.⁵ Sometimes this appears to be taken as a model
for phenomenal consciousness across the board. Part of my aim in this
paper is to suggest that this way of thinking about phenomenal con-
sciousness is too narrow. Being aware of something as a fact about the
world can make a subjective phenomenal difference, or a difference to
“what it is like” for one, even if that state of awareness does not instanti-
ate any qualitative properties that one is aware of. This rather minimal
sort of phenomenology suffices, as I will argue, to ground a satisfying
account of how we are able to know what we consciously believe.

Now, if my account succeeds in the case of our knowledge of our
conscious beliefs, one might wonder whether it might be extended
to other conscious mental states (Evans himself suggests that his ac-
count extends at least to perceptual states). I think that this is a line of
research that is well worth pursuing. However, for the purposes of the
present paper I will restrict my attention to belief alone.⁶

awareness, properties such as its color, size, or shape. Nothing of the sort
is relevant here, however. In having the conscious belief that snow is white
you are not aware of any qualitative properties of the content of that belief
(whatever those might be!)—you are, rather, aware that snow is white. I discuss
this issue in more detail in Section 2.2.

5. For present purposes, the qualitative properties or qualities of a thing are
those of its properties that characterize what that thing is like in itself and in
actuality, without “pointing beyond” that thing. So the qualitative properties
of, say, a page of written text would include things like the color, texture,
and chemical composition of the paper, but not the semantic properties of the
text, as characterizing those properties would involve mentioning things other
than this piece of paper. Of course some proponents of a representationalist
approach to consciousness (such as Tye 1995) deny that consciousness ever
consists in the instantiation of such qualitative properties, even in the case
of pains. I do not intend to take sides on this issue here: my point is just that
there is a strong intuition that bodily sensations involve purely qualitative
phenomenal properties, not that they really do.

6. Sosa (2003, 135–6) suggests that failure to account for how we know that we
don’t believe something is a fatal difficulty for transparency-based accounts.
The paper proceeds as follows. I introduce my account in Section 2, by comparing it to some influential alternatives in the literature. Then in Section 3 I develop and defend my view on the phenomenology of belief, on which everything else rests.

2. The Rationality of Self-Ascriptions of Belief

The view I will defend in this paper is this ("KB" for "knowledge of belief"):

(KB) If S is rational and has the concept of belief, then she has a distinctive (non-observational and non-inferential) first-personal way of knowing what she consciously believes.

To begin unpacking and assessing this claim we need at least a preliminary characterization of conscious belief. This is an expression that has currency in everyday speech, and my usage of it is meant to resonate with this. Roughly, your conscious beliefs are those that are available to you as your own reasons for action or belief: if you Φ for the reason that p, then you must consciously believe that p. These are, intuitively, the beliefs that you can draw upon in deliberating about what to do or what to believe (although not all instances of acting or believing for a reason involve explicit deliberation, I take it that at least the possibility of deliberation is essential to our concept of a reason). Such beliefs might be *occurent*, if they are in the moment actually involved in action or deliberation, or *standing*, if they are merely available for such involvement. *Unconscious* beliefs, by contrast, are beliefs that play a role in motivating or guiding action or thought, but are not, in an intuitive sense, available to the subject in deliberation, or as her own reasons.\(^7\)

Notice that, in order to avoid circularity, I have not built any claims about phenomenology in this characterization of conscious belief. As I will argue in Section 3, occurrent conscious beliefs also make a distinctive difference to your phenomenology, or to what it is like for you at the time. Moreover, as it will turn out, this is what underpins both their reason-giving role and our capacity to know what we consciously believe. But this is something to be argued for, not to be built into our characterization of conscious belief.

Notice, moreover, that in characterizing conscious beliefs in terms of availability to deliberation we do not trivialize the claim to self-knowledge: it is one thing to be able to use the content of a belief in deliberation, and another to be able to knowledgeably self-ascribe that belief.\(^8\) Deliberation can be — and typically is — a first-order activity, whose subject matter is the world, not the subject’s own states. If there is a necessary connection between a belief’s being available for

---

\(^7\) The distinction between the conscious and the unconscious intended here is not the same as the distinction between the personal and the sub-personal. Unconscious beliefs are states of the *whole subject*, rather than of some proper part of her. Both common sense and empirical psychology recognize unconscious (or “implicit”) attitudes that play the relevant role in our cognitive economy, and it seems to me natural to call them “beliefs.” Some philosophers prefer to avoid using the term “belief” for such attitudes (*e.g.*, Smithies 2012, 275). I doubt that ordinary usage is determinate enough to settle this question definitively, but I don’t think anything of substance hinges on it.

\(^8\) Note that, as I will be using the term, a self-ascription need not be an overt linguistic act. To self-ascribe the belief that p is simply to come to believe that you believe that p.
deliberation and for self-knowledge, as KB implies, then that connection is a substantive one, not the product of our definitions.

Now it is true that, on some views, conscious beliefs are *ipso facto* the object of some higher-order propositional attitude, such as belief. And it might seem that on such views KB would be trivial, or at least not in need of much separate argument. But I think this is not so. Suppose that every conscious belief is accompanied by a corresponding higher-order belief. Still, KB does not follow, unless such higher-order beliefs (at least normally) qualify as knowledge. But why should we suppose that they do?

Of course, the relevant higher-order beliefs will be true, and reliably so. On some externalist accounts of knowledge this suffices for them to qualify as knowledge; in particular, we do not need to worry about how such higher-order beliefs are justified or rational from the point of view of the subject herself. Regardless of the prospects of such a view in epistemology at large, however, it would seem to falsify the special character of self-knowledge. It is hard to accept that the epistemic status of our beliefs about our own conscious mental states might depend entirely on factors outside our ken. This is why, traditionally, our knowledge of our own conscious mental states has been considered as the natural source of internalist intuitions about knowledge. For the purposes of this paper I will assume that such internalist intuitions do, indeed, deserve respect. Thus, even were we to accept a higher-order view of consciousness for beliefs, we would still need an account of how these higher-order beliefs are justified from the subject’s point of view.

So KB is a substantive philosophical thesis. Is it plausible? Note that KB does not claim that we are either infallible or omniscient about our conscious beliefs. Thus KB does not seem to be overambitious in its scope: the amount of self-knowledge it predicts intuitively seems to be available to us. The difficult question is how such self-knowledge

9. Although there is a large literature on higher-order theories of consciousness in general, this literature has paid relatively little attention to conscious beliefs in particular. But see, e.g., Mellor (1977).

is available to us: is it really, as KB suggests, available in a way that is distinct from observation and inference? This is the main question I will attempt to answer in this paper.

2.2 The Quasi-Perceptual Model and the Transparency of Belief
The thought that self-ascriptions of beliefs and other conscious mental states are often not only reliable but also *rational or justified* from a first-personal point of view is widely shared. This thought, however, might make it seem as if we need to look for something to play the role of evidence for self-ascriptions of belief; and states of *introspective awareness* of one’s conscious beliefs might appear as tempting candidates for that role. Introspective awareness itself is, on such views, analogous to perceptual awareness of objects: it is awareness of events in consciousness as mental particulars.10

It is clear that this approach conflicts with the idea Evans expresses in the passage quoted in Section 1: part of Evans’s point is clearly that our self-ascriptions of belief are not based on internal quasi-perceptual states of awareness of belief. Evans, however, provides no argument for his view. So why should we reject the quasi-perceptual model?

The main reason, I suggest, is phenomenological. Being perceptually aware of an object entails being aware of some of its qualitative properties. For example, if you are visually aware of the vase on the sideboard, then that must be because you are aware of its shape, color, size, etc. Thus, if having a conscious belief entailed anything analogous to perceptual awareness of objects, then in having a conscious belief you would have to be aware of some of its qualitative properties. But this condition, I submit, is not satisfied: in having a conscious belief that *p* you need not be aware of any such mental qualitative properties.11

10. The most prominent recent defender of this view has been Pitt (2004). Pitt (2004, 210) is careful to qualify the analogy with perception: while an apple is clearly distinct from your state of perceiving the apple, your introspective awareness of a conscious belief of yours is not similarly distinct from that belief (see also BonJour 2003, 62). This distinguishes such views from the ‘inner sense’ views Shoemaker (1996) criticizes.

This point, I think, sometimes gets obscured by an ambiguity surrounding the notion of *awareness of content*. Consider, for instance, how David Pitt (2004, 22) argues against the following line of thought, which he attributes to Sara Worley:

To have a conscious thought is to be consciously aware of the contents of the thought. There is no qualitative experience over and above the awareness of the content of the thought. [...] The qualitative difference between thinking that today is Wednesday and thinking that it is Thursday is that one is aware of different contents in the two cases. That difference exhausts the [phenomenological] difference between the two thoughts.

Worley’s suggestion is that having conscious thoughts does not involve being aware of the qualitative properties of the thoughts as mental particulars, but only being aware of the contents of our thoughts. Thus Worley seems to be advocating something like the transparency view I will be defending here. In his response, however, Pitt (2004, 22), construed awareness of content in a very different way:

When we introspect we turn out attention inward, toward the contents of our minds—which are mental if anything is. Thus, to be a direct realist about introspectable properties is to recognize subjective phenomenal characters.

I think Pitt goes wrong concerning what turning our attention toward the content of our minds involves. Turning our attention toward the content of our beliefs does not involve turning our attention inward—on the contrary, it involves turning our attention outward, upon the world (as it is according to us, of course). Attending to the content of a book does not involve attending to what is literally inside the book, *i.e.*, the paper marked with ink; it involves attending to, *e.g.*, the story told by the book, or the characters involved in it (the analogy is borrowed from Dretske 1993). Similarly, attending to the content of your mind does not require attending to anything distinctively inner, but rather to what the world is like according to you.

The point might become clearer if we ask what turning one’s attention outward would involve, on Pitt’s view. If I attend to, say, the geopolitical situation in the Middle East, do I do anything other than attend to the content of my relevant beliefs? Of course, if my interest is in the Middle East rather than my own mental states, I would not naturally describe what I do in terms of attending to content, but the actual procedure seems to be the same in both cases. What might attending to the world involve, if not attending to the content of one’s world-directed mental states?²

Thus, in contrast to the quasi-perceptual account, I believe that attention to the phenomenology of belief suggests that the only type of conscious awareness that is essentially involved in having a conscious belief is awareness that the world is a certain way — namely, the way that the content of the belief specifies. Just as Worley suggests in the passage quoted above, this type of awareness is very naturally described as awareness of content. But it is not awareness of any distinctively mental properties, and it is not quasi-perceptual awareness of the belief itself as a particular mental event. Beliefs are *transparent*, in the sense that they do not register in our internal gaze. It follows that the quasi-perceptual account, which casts states of awareness of beliefs as mental particulars in an evidential role, should be rejected.

2.3 Judgment-Based Accounts
A number of recent authors — including Christopher Peacocke (1998), Nico Silins (2012), and Declan Smithies (2012) — have suggested that our capacity to know what we believe is grounded in a connection between “internal” and “external” evidence, which she uses to argue against some Evans-inspired accounts of self-knowledge. Gertler’s distinction is between old evidence and evidence which one has just acquired. But Gertler gives no good reason for thinking that attending to old evidence involves looking inward in a sense that attending to evidence newly acquired does not: in both cases one attends to the world, in the only way one can do such a thing — *i.e.*, by attending to the content of one’s relevant mental states.
between standing beliefs, which are not phenomenally conscious, and occurrent judgments, which are. These views are significantly different from the quasi-perceptual views discussed above, insofar as occurrent thoughts or judgments are supposed to justify self-ascriptions of belief directly, without intervening states of awareness of oneself judging (Peacocke 1998, 71–2; Silins 2012, 305–06). In denying such intervening states an epistemological role these views are similar to my own. But, I believe, these views are ultimately unsatisfying, because they fail to give a plausible explanation of the link between the phenomenology of judgment and the epistemology of self-knowledge. If it really is true that phenomenally conscious events of judging play an epistemological role in justifying our self-ascriptions of belief, then it should be possible to give an account that shows how the distinctive phenomenology of judgment — what it is like to judge — makes it rational to self-ascribe a corresponding belief.

This demand is just an instance of a plausible broader principle: if you claim that a phenomenally conscious state M justifies a certain response, then you should be able to tell a story about what it is like to be in M that shows why it is rational for a subject who enjoys the relevant phenomenology to respond in that way. As we shall see, however, existing judgment-based accounts fail to meet this demand.

Now, on the views under consideration, judgment is a reliable indication or symptom of belief, since a judgment that p typically either initiates or expresses one’s belief that p (Peacocke 1998, 88; Silins 2012, 306–09; Smithies 2012, 273–74).13 Thus, if one judges that p, then that is a fairly good sign that one also believes that p. This, however, cannot explain why consciously judging that p gives one reason to self-ascribe a belief that p, unless we also assume that one is already aware that one has judged that p, and that awareness plays a justificatory role in one’s self-ascription. But, as we have seen, the views currently under consideration deny that such states of awareness of oneself judging are involved in the justification of self-ascriptions of belief. Justification is supposed to accrue directly from the judgment that p to the self-ascription that p.14 But why might this be so?

Peacocke (1998, 73) argues that, in reasoning, subjects must be sensitive not just to the contents of their attitudes but also to the attitudes themselves. Thus, for example, one’s belief that p will dispose one to use p as a premise for an inference, but one’s wish that p will not. This sensitivity, Peacocke seems to suggest, is what explains why a judgment that p can provide one with a reason for a self-ascription of the belief that p. Now Peacocke is not fully explicit as to whether this sensitivity is supposed to be a conscious capacity: do you need to be aware that you believe (rather than, e.g., wish) that p in order to use p as a premise, or does an unconscious disposition suffice? As we have seen, Peacocke rejects the idea that self-ascriptions of belief are justified by one’s awareness of oneself judging. Thus, we should take Peacocke’s suggestion to be that subjects have an unconscious sensitivity to their own judgments, which disposes them to self-ascribe beliefs with the same contents. But such an unconscious sensitivity clearly does not help us with our question, which concerns the connection between the phenomenology of judgment and its purported epistemological role in self-ascriptions of belief.

Silins’s discussion of immediate justification suggests a parallel between his view and certain views in the epistemology of perception. On such views, having an experience as of p can give you justification to believe that p which is immediate in just the same sense as the justification that, on Silins’s view, judging that p gives you to self-ascribe the belief that p. But this analogy cannot take us very far (not that Silins thinks that it does; I am here extrapolating beyond anything he explicitly says). According to the relevant views on perception, the content

13. To my mind, putting matters this way is actually misleading. Judgment does not express or initiate belief (at least if ‘initiate’ means something like cause), but rather entails or constitutes belief (I discuss this point in more detail below, in n. 19). For the purposes of the present discussion, however, I will follow my interlocutors’ way of speaking.

14. It is also worth noting that, according to Peacocke (1998) and Smithies (2012), the connection to conscious judgment plays an epistemological role in justifying self-ascriptions of belief even if one does not actually make the judgment. But the fact that a judgment that never gets made would be a reliable indicator of belief surely can play no justificatory role.
of the perceptual state is just the same as the content of the belief it justifies; indeed we might expect that this sameness of content would play a central role in an explanation of why perception plays the epistemic role claimed for it (as is suggested by, for example, Jim Pryor [2001]). But the content of a self-ascription of belief is obviously not the same as the content of the judgment that is supposed to immediately justify it. The fact (if it is a fact) that a perception with the content that \( p \) provides immediate justification for a belief that \( p \) provides no reason to think that a judgment that \( p \) might provide immediate justification for the belief that one believes that \( p \).

Note that to say as much is not to accept what Silins (2012, 310) calls the “Face Value Constraint”, i.e., the claim that a mental state can provide immediate justification for another only if they share the same content. The point is just that sameness of content is an essential part of one plausible explanation of why perceptual states provide immediate justification for beliefs. Since the content of a judgment is not the same as the content of the self-ascription it is supposed to provide immediate justification for, a different explanation is needed. But we have yet to see such an explanation.

2.4 The “Stepping Back” Account

If, in agreement with the judgment-based accounts discussed above, we deny an epistemological role to states of awareness of one’s judgments as mental particulars, what is it about the phenomenology of judgment (or occurrence belief, as I prefer to think of it) that grounds the rationality of self-ascriptions of belief?

I have already introduced the idea that belief is transparent, in Section 2.2. The context there was polemical: my point was only that introspection does not support the idea that in having a conscious belief you must be aware of anything distinctively mental. Now, however, I want to suggest that we can also put the idea of transparency to positive use: even if the phenomenology of belief is transparent, it can still ground the rationality of self-ascriptions of belief. The crucial point is to recognize that accepting the transparency of belief does not entail denying that belief has a distinctive phenomenology; it does not entail denying that there is something that it is like for you to have an occurring, conscious belief that \( p \). Having an occurring, conscious belief involves being aware that the world is a certain way — the way that the content of the belief specifies. Moreover, this is a claim about conscious phenomenology — it is a partial characterization of what it is like for you at the subjective, phenomenal level.

Now, the claim that belief has any kind of distinctive phenomenology is of course notoriously controversial, and I do not expect what I just said to carry much power to convince. Section 3 of this paper is devoted in its entirety to defending my view on the phenomenology of belief. For the rest of the present section I want to argue that if this claim is granted, then we can put it to good use in an account of self-knowledge.

Consider a subject who has a conscious, occurring belief that \( p \). Thus she is aware — in the phenomenal sense — that \( p \). Suppose, moreover, that our subject is rational and in possession of the concept of belief. If this is correct, then she must grasp the distinction between her own take on the facts and the facts themselves: she must know that her taking the world to be a certain way is a different matter from the world’s really being that way. This piece of knowledge, I suggest,
enables our subject to step back from her awareness that \( p \), and self-ascribe the belief that \( p \).

Notice that the claim is not that the subject infers from the premise that she is aware that \( p \) to the conclusion that she believes that \( p \) (indeed, as I will explain below, I think it is misleading to describe the transition here as any kind of inference at all). Rather, from the subject’s own point of view, the transition is from the way the world is to a self-ascription of the belief that the world is that way. Moreover, for the stepping back procedure to be applied, the subject need not suspect that her beliefs fail to represent the world as it really is. It would be absurd to suggest that you can self-ascribe the belief that \( p \) only if you suspect that \( p \) fails to truly represent the world — for if you do suspect this, you should rationally not believe that \( p \). Indeed, the stepping back procedure would in principle be available even to an extremely epistemically arrogant subject, who presumes that none of her beliefs are false. The claim is that a rational subject with the concept of belief must have a single piece of general knowledge — namely, that it is one thing for her to take the world to be a certain way and another for the world to be that way — which enables her to apply the stepping back procedure in every case in which she takes the world to be a certain way.

Obviously, self-ascriptions made via the stepping back procedure will be true, and reliably so. Moreover, it seems clear that they are rational, from the point of view of the subject herself. So they have a very good claim to count as knowledge. This stepping back procedure, therefore, provides rational subjects who possess the concept of belief with a distinctive way to know what they believe.

Now the account just provided applies directly only to one’s occurring conscious beliefs. What about one’s standing conscious beliefs? Normally, one’s standing conscious beliefs can become occurring to one just as a result of one’s turning one’s attention to the relevant topic. (Lapses of memory, either random or motivated by various forms of cognitive bias, are of course a live danger here — and can undermine one’s capacity for self-knowledge in any given case.) Although there is such a thing as consciously searching one’s memory for information on a particular topic, typically the process through which one’s standing beliefs become occurring for one is itself unconscious and involuntary. But as soon as the belief occurs to one, one can self-ascribe it by stepping back. It follows, therefore, that, although the present account does not entail that one actually knows that one has the standing beliefs that one has, it does entail that for each such belief one has a straightforward procedure for knowing that one has it. In order to self-ascribe the belief that \( p \), one simply has to consider whether \( p \) — whereupon it will occur to one that \( p \), and one will be in a position to self-ascribe the belief via standing back.

This, incidentally, gives us a way to interpret Evans’s (1982, 225) remark that “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 \( p \).” Evans’s remark has sometimes been interpreted so as to suggest that on each occasion of self-ascription I need to make up my mind anew on the question whether \( p \)
(Moran 2001 seems to read Evans this way). But it is not at all clear that this interpretation is correct: if I already have a standing belief that \( p \), the relevant “procedure for answering the question whether \( p \)” might be just whatever unconscious mechanism is involved in its occurring to me that \( p \).

2.5 Stepping Back versus the Doxastic Schema

A crucial aspect of my stepping back account is the claim that a rational subject with the concept of belief can distinguish between her take on the facts and the facts themselves. André Gallois (1996, 2010) marshals similar considerations of rationality in support of his view that self-ascriptions of belief are based on inference, in accordance with the following “doxastic schema”:

\[
(DS) \text{p, therefore I believe that p}
\]

This transition obviously looks a lot like the one I suggested above. But, in fact, there are significant differences between Gallois’ views and the view proposed here.

To begin with, I think it is misleading to describe the transition that DS tries to capture as an inference. Neither Gallois nor other supporters of DS have provided a fully fleshed out account of inference, but it seems fair to say that if you (non-hypothetically) infer \( A \) from \( B \), then you must be relying upon \( B \). This is what particles such as “therefore” or “thus” indicate. As I have been suggesting, however, far from relying upon \( p \), self-ascribing a belief that \( p \) involves stepping back from, or bracketing, your commitment to \( p \). It involves recognizing that \( p \) is part of your subjective take on the facts, as distinct from the facts themselves.

Recognizing that the stepping back move is not any kind of inference allows us to avoid a number of objections that have been raised against the inferential view. As even its proponents recognize, it is hard to see why it is rational to reason in accordance with DS. DS is neither deductively valid nor inductively strong: as has been widely noted, \( p \)’s being the case does not in general make it more likely that any particular person believes that \( p \) (the fact that it is raining in Jakarta does not make it likely that Mary, who lives 5,000 miles away, believes that it is). So how can it be rational for a subject to self-ascribe the belief that \( p \) on the grounds that \( p \)?

Byrne’s (2011) response to this problem is to argue that DS is a “good” rule, i.e., one that reliably leads to a true conclusion if one believes its premise. It is not at all clear, however, that this answers our question, which concerns why any particular inference in accordance with DS should seem rational from the point of view of the subject herself. Suppose that Mary recognizes that DS has this feature — that is, she knows that if she believes that \( p \), then she can reliably infer from \( p \) that she believes that \( p \). But this piece of knowledge is entirely useless to her in practice: in order for Mary to recognize that any particular situation is one in which application of DS is rational, she would have to already know that she believes that \( p \). But then she has no need for DS in the first place. Thus the fact that DS is a good rule gives us no help in understanding why it is ever rational, from the point of view of the subject, to apply that rule.

From the present point of view, the response to this line of objection is entirely straightforward: the transition involved in moving from

$p$ to a self-ascription of the belief that $p$ is not an inference, and thus it is not a case of forming a belief on the grounds that $p$. It follows that the rationality of the transition is not undermined by pointing out that it is neither deductively valid nor inductively strong. The transition at hand is a case of stepping back from $p$ — a sui generis, and yet intuitively rational, transition. As we saw, we can explain why this transition is rational, partly in terms of the phenomenology of occurrent conscious beliefs and partly in terms of general knowledge that goes together with possession of the concept of belief.

3. The Transparent Phenomenology of Belief

As we have seen, my account of how we know our own conscious beliefs crucially depends on a particular view about the phenomenology of occurrent conscious beliefs. But this view, and indeed the very idea that beliefs have a distinctive phenomenology, is controversial. In this section I want to provide some arguments in their favor.

I have so far been working with a broadly functional characterization of conscious belief, according to which a belief counts as conscious if and only if it is available to the subject as her own reason for action or belief, or in deliberation. In this section I will argue that it is only because of their connection to phenomenology that conscious beliefs can play this role.

If you have a conscious but standing belief that $p$ then, normally, if the question whether $p$ becomes relevant to you that belief will also become occurrent — it will occur to you that $p$. This is the process — typically unconscious and involuntary — whereby the content that $p$ enters deliberation. But, as I will be suggesting, only contents you are phenomenally aware of can figure in deliberation. Non-phenomenally conscious content can causally influence and perhaps even motivate action; but it cannot render it intelligible or rational from the point of view of the agent herself. It follows, then, that the content of occurrent beliefs is phenomenally conscious content. Moreover, although we do not always need to deliberate about what to do or think (many of our actions and inferences are “automatic”), the possibility of deliberation is essential to our concept of a reason: if $p$ is your reason for doing or thinking something, then there must be a piece of reasoning in which $p$ figures as a premise and which you could have gone through, in explicit deliberation, to the same conclusion. Thus, even in cases where deliberation does not actually take place, phenomenal consciousness is still necessary for the content of a belief to be your reason for action or thought.

My argument for this requires drawing a distinction, within the domain of actions, between those that are done for reasons and those that are not. I will then suggest that that distinction is to be understood in terms of phenomenally conscious beliefs: in the former case, but not in the latter, one acts on contents that one is phenomenally conscious of.

Let us start by noting that there is a clear phenomenological difference between actions done for reasons and those that are not: the former are experienced as intelligible by their agents, while the latter are not (despite the fact that they might still count as rational, in light of the agent’s own goals and beliefs). I hope that it will be granted that there is such a difference. My argument is going to be that this difference is often explained by another one: the phenomenological difference between being aware of a content at the phenomenal, subjective level, and being guided by information which is not so available to one. This latter difference is the difference that an occurrent conscious belief makes.

Consider that staple of philosophical discussions of consciousness, the phenomenon of blindsight. In this phenomenon, subjects who are partially blind due to damage to the primary visual cortex are nevertheless able, when prompted, to direct appropriate behavior toward target-objects located in their scotoma, or blindfield, which they deny seeing — e.g., they reach out for them at the right place, and with their

21. Note that for $p$ to “enter deliberation” in this sense you don’t have to actually draw any conclusions from $p$—you might instead decide that, although $p$ appeared prima facie to support a particular line of thought or action, it was in fact defeated by other considerations. Many more considerations enter deliberation than you end up drawing conclusions from.
hand pre-shaped in their right way. But their success comes as a surprise to the subjects; they merely find themselves extending their arm in the right direction and pre-shaping their hand for the right sort of grip; they can make no sense of why they act in these ways.\(^{22}\)

This sense of unintelligibility on the part of blindsight subjects has nothing to do with defects in their capacity for awareness of agency as such. It is generally accepted that agents have a special sort of awareness of their actions.\(^{23}\) This agential awareness can be disturbed in certain conditions (e.g., alien hand syndrome), but blindsight is not one of them. The point, rather, is that while blindsighters have genuine agential awareness of their actions, they nevertheless experience their actions as unintelligible. How might we explain this fact?

No doubt the empirical details of the explanation will be complex. But it seems reasonable to expect that, whatever the details might be, the following features of the situation will be central to such an explanation. In the normal case, the content of your agential awareness (e.g., “I am extending my arm to the left”) links up with other contents you are aware of (e.g., “my target is over there on the left”). But in the blindsight case the latter type of awareness is absent: the agent is not aware of the facts that she is responding to—for example, she is not aware that there is a target located in the direction in which she is stretching her arm. Thus, although she is aware that she is performing a particular action, she is not aware of any facts whereby that action might be explained.\(^{24}\) No wonder, then, that the action is unintelligible to her.

22. For a classic discussion of the cognitive neuroscience of blindsight, see Weiskrantz (2009). In addition, my own approach to the issue has been influenced by Campbell (2003) and Smithies (2011).

23. For a thorough discussion of psychological and philosophical issues involved in such awareness of agency, see the essays in Roessler and Eilan (2003). For present purposes we do not need to give an account of what exactly normal awareness of agency comes down to—e.g., whether it is perceptual in nature or not—so long as it is agreed that agents normally possess it.

24. Or, at least, she will not be aware of any facts that might explain the relevant features of that action. Familiarly, action explanations are sensitive to what descriptions we use to pick out the relevant actions. For example, in the blindsight case the subject will, presumably, be aware of both her action and an explanation for it under the description “trying to carry out the task requested by Dr. X.” She will not, however, be aware of anything that explains her doing so by (e.g.) extending her arm to the left rather than to the right.

25. This is not to suggest that blindsight subjects have unconscious beliefs about objects in their blindfields. The relevant contrast here is simply between the presence of occurrent conscious beliefs and their absence. It is plausible that the information about objects in their blindfields that the blindsighters’ motor system exploits is not available to them in any sense; it is merely sub-personal.
of view about perception I favor, perception does not have propositional content. Perception is awareness of objects and their features, not awareness that something is thus-and-so. This is not to deny that perception plays a rational role in our mental economy: it is in virtue of your visual awareness of a red rose that you are justified in believing that there is a red rose in front of you. It is only to deny that perceptions have the sort of propositionally articulated content that is required for a state to give you reasons for action or thought. Reasons are propositions, and perception does not give you that. 26

Of course, this is not the place to defend this view about perception. For present purposes, we can set aside the earlier objection to using blindsight as evidence for my account of the importance of phenomenology because, once we know what we are looking for, we can find the same structure in cases that have no direct link with perception.

Consider Fred, some of whose actions might be explained by ascribing to him the unconscious belief that all of his endeavors are doomed to failure — although Fred, if asked, sincerely denies having such a belief. For instance, this belief might help explain why Fred rejected a job offer which seemed like a perfect fit for him. But notice that in such a case it is very likely that Fred himself will have experienced his rejection of the offer as unintelligible: he might, for example, say that he simply found himself unable to accept the offer. My suggestion is that the explanation of Fred’s sense of unintelligibility lies in his not being aware, at the subjective level, of his endeavors as doomed to failure. This is so despite the fact that, by hypothesis, this belief is what led him to reject the job offer. Once again, we see a distinction between a content’s ability to influence or explain action and the subject’s being aware of that content in a way that enables her to experience her actions as intelligible, or to take it as her own reason.

26. The view that perception does not have propositional content is of course controversial. For recent arguments for that view, see Campbell (2002), Travis (2004), Brewer (2006), Johnston (2006), and McDowell (2009). How exactly perception is linked to perceptual beliefs, and in what sense this link is a rational one, is a matter of dispute.

It is crucial to note that the phenomenological difference between action guided by conscious belief and action guided by unconscious belief can be explained by a phenomenological difference in the first-order beliefs themselves, without having to introduce any sort of awareness of one’s own beliefs as such. Think back to the case of blindsight. In order to explain why the blindsighters’ actions come as a surprise to them, we do not need to invoke the fact that they are not aware that they believe or perceive that there is an object in their blindfield; all we need is the fact that they are not aware that there is an object in their blindfield. Similarly, in Fred’s case, the reason why his action is experienced as unintelligible by him is not that he is not aware that he believes that his endeavors are doomed to failure — after all, even if he did know that on the same sort of explanatory grounds as we do, his experience would plausibly still be one of merely finding himself unable to accept the offer. The explanation, rather, is just that Fred is not aware that his endeavors are doomed to failure — it just doesn’t seem to him to be a fact about the world.

These considerations support the idea that the phenomenology of conscious belief is transparent. It is precisely this transparent phenomenology that allows conscious belief to play its distinctive role in deliberation: in having an occurrent conscious belief that p you are aware of the content that p as a fact about the world, and so you are in a position to take p as your reason for action or thought.

Moreover, as we saw earlier, a rational subject who possesses the concept of belief is also able to use her awareness of the content that p in a different way as well. Such a subject knows that it is one thing for p to be part of her own take on the world, and a different thing for p to really be the case. Thus, the subject is able to step back from her awareness that p, and self-ascribe the belief that p. This is how the rationality of self-ascriptions of occurrent beliefs is grounded in their phenomenology.

I now turn to some objections to this account.
3.2 Is Phenomenal Consciousness Necessary for Reasons?

According to what I argued above, a content can figure as your reason for thought or action only if it is phenomenally conscious. Someone might worry that my argument rests on confusing phenomenal consciousness with access consciousness, in the sense of Block (1995). Block’s notion of access consciousness is a matter of content’s being “poised” for use in reasoning and the rational control of action. Thus, one might think that, since what is at issue in my argument is the control of thought and action, all we need is some kind of access consciousness, and discussions of phenomenology are simply irrelevant.

But such criticism would be missing my point. Contents that are access conscious in Block’s sense might still not be available to the subject as her own reasons for acting. Consider Block’s “superblind-sighter”, a subject in whom information derived from the blindfield can spontaneously initiate action. According to Block, such information is access conscious. But, intuitively, there is no sense in which such information could figure as the subject’s own reason for doing anything; after all, the subject herself, as opposed to her sub-personal motor control systems, has no access to that information. Her actions would still be experienced as unintelligible by her (even though, perhaps, after a sustained streak of successes she would no longer be surprised by them). If so, mere access consciousness is not sufficient for a content to be available to you as your reason.

In fact, we don’t need to speculate about superblind sight to make this point. Unconscious attitudes can motivate complex goal-directed behavior. For example, according to standard analyses of self-deception, a self-deceived subject implicitly recognizes that p and is motivated by that recognition to act in ways that allow her to maintain the illusion that not-p. The content that p is thus available for the rational control of action in Block’s sense, and thus has a good claim to being access conscious. But it is not available to the subject herself, as her reason for acting. My suggestion is that phenomenology explains this fact: it is not available to her as a reason for acting because she — at the subjective, phenomenological level — is not aware that p. To the extent that the deception succeeds, things are for the subject as if not-p.

The lesson to draw from this is just that it is one thing for an action to be rational, in the sense of being adequate to the normative demands of the situation, and another for it to be done for reasons. The self-deceived person’s stratagems are instrumentally rational, in the sense that they are likely to promote her goal of keeping herself in the dark about p; but in pursuing those stratagems she is not acting for the reason that p. By contrast, a smoker’s continued smoking might be irrational in light of her own values and beliefs, and yet each time she lights up a cigarette she is acting for reasons. Block’s access consciousness has to do with the rational control of action, rather than with acting for reasons.

A different objection to the claim that phenomenal consciousness is necessary for a state to give one reasons might focus on actions that are performed automatically, and yet intuitively count as done for reasons.

Typically, the self-deceived subject will not be in a position to report the content that p as what she believes on the matter, so it doesn’t meet one of Block’s criteria for access consciousness. However, given that access consciousness is defined in terms of availability for control, it is only to be expected that such dissociations will occur when a content is only partially available for control. For present purposes, settling which (if any) cases of partial availability amount to access consciousness is not crucial. What is crucial, as we shall shortly see, is the distinction between a content’s being available for rational control in Block’s sense, and a content’s being available as the agent’s own reason.

A similar lesson is drawn by Smithies (2011, 267). Smithies puts the point in terms of a distinction between normative and causal accessibility, where in order to use a content as your reason it must be normatively accessible to you. I think the way Smithies frames the distinction is somewhat unfortunate, however, as it is something of an open question whether in taking something as your reason to Φ you must also think that it really justifies Φ-ing, i.e., that it is a reason in the normative sense to Φ (for dissent, see Stocker 1979; Velleman 1992; Setiya 2007). Everyone should agree, however, that taking something as your reason for acting has explanatory import for you: if you Φ for the reason that p, then your Φ-ing is intelligible to you in light of p.

27. It is something of an open question whether this ever happens with actual blindsighters. Given that unconscious contents can surely motivate behavior in other contexts, I see no reason to deny at least the possibility of superblind sight.

28. Typically, the self-deceived subject will not be in a position to report the content that p as what she believes on the matter, so it doesn’t meet one of Block’s criteria for access consciousness. However, given that access consciousness is defined in terms of availability for control, it is only to be expected that such dissociations will occur when a content is only partially available for control. For present purposes, settling which (if any) cases of partial availability amount to access consciousness is not crucial. What is crucial, as we shall shortly see, is the distinction between a content’s being available for rational control in Block’s sense, and a content’s being available as the agent’s own reason.

29. A similar lesson is drawn by Smithies (2011, 267). Smithies puts the point in terms of a distinction between normative and causal accessibility, where in order to use a content as your reason it must be normatively accessible to you. I think the way Smithies frames the distinction is somewhat unfortunate, however, as it is something of an open question whether in taking something as your reason to Φ you must also think that it really justifies Φ-ing, i.e., that it is a reason in the normative sense to Φ (for dissent, see Stocker 1979; Velleman 1992; Setiya 2007). Everyone should agree, however, that taking something as your reason for acting has explanatory import for you: if you Φ for the reason that p, then your Φ-ing is intelligible to you in light of p.
for reasons. Suppose that, while absorbed in her book, Mary picks up a cup of tea from the desk and takes a sip. Intuitively, Mary has acted intentionally and for reasons, even though no deliberation took place. Must it still be the case that her action involved phenomenally conscious occurrent beliefs? On the view I have been defending here, the answer has to be “yes”: if part of Mary’s reason was that there is tea in the cup, for example, then it must have occurred to her that there is tea in the cup. I do not think that this is implausible. For one thing, remember the point about transparency: to say that it occurred to Mary that there is tea in the cup is, on my view, just to say that she was aware that there is tea in the cup. It entails nothing about her being aware of herself thinking about the tea in her cup. Moreover, phenomenal consciousness has both a foreground and a background. While some things are experienced vividly and occupy much of our attention, others we barely notice. Actions performed automatically, like Mary’s picking up her cup of tea, fall toward the unnoticed part of this spectrum. Naturally, the same goes for the occurrent beliefs involved in them.

Finally, it is important to note that, while Mary’s action is intentional under some descriptions, it is not intentional under others. For example, being absorbed in her book, Mary plausibly paid no attention to the exact trajectory her hand took through the air to reach the cup on her desk. Thus nothing was Mary’s reason for that aspect of her action. The precise trajectory of her movement was chosen by sub-personal mechanisms, not Mary herself. It follows that Mary need not have any occurrent conscious beliefs that would make that aspect of her action intelligible. In general, it seems safe to say that automatic actions are typically intentional only under rather broad descriptions. It follows that they need involve only rather broad conscious beliefs as well.

3.3 Reductive Accounts of the Phenomenology of Belief

Although my account of the phenomenology of belief is fairly minimal, the idea that belief or other non-sensory cognitive states have any distinctive phenomenology at all is notoriously controversial (for detailed discussion, see the essays in Bayne and Montague 2011). Opposition to this idea does not take the form of a flat denial that there is ever anything that it is like to consciously believe (or think, or question, or doubt…) something, but rather the form of trying to account for that phenomenology in other, supposedly more familiar, terms (see, e.g., Lormand 1996; Prinz 2007). I want to say here why I don’t find such attempts compelling.

The central idea in this reductive approach is that consciously having the content that $p$ in mind consists either in visual or auditory imagery of a sentence that means that $p$, or in a mental picture that stands in for $p$. For example, one might visualize the thought that it will rain this afternoon by picturing dark clouds and rain. I do not wish to deny that occurrent conscious thought is often (or perhaps even always) accompanied by inner verbalization or imagery. My objection to the reductive approach is to the claim that what makes the content of thought phenomenally conscious is our awareness of images or words.

Proponents of the reductive strategy often try to motivate it by appeal to Ockham’s razor: we should not multiply irreducible types of consciousness beyond what is strictly necessary. The problem with this, however, is that awareness of images or words seems incapable of playing the role in our mental economy that I have been arguing conscious thoughts play. Suppose that, as you leave your house in the morning, you take your umbrella with you for the reason that it will rain in the afternoon. As I have argued, taking the content that it will rain this afternoon as your reason requires that you be phenomenally conscious of that content. But if your consciousness of that content consisted merely in your picturing dark clouds and rain, how could that give you a reason for taking your umbrella with you? Presumably it could only do so by suggesting to you that it will rain — but then it is awareness of that (putative) fact about the world, and not of any images, that gives you the reason for your action. Pictures or words are not, as such, reasons for anything; it is your awareness that the world is a certain way that gives you reasons for action or thought.
3.4 Transparency in Belief and in Perception
On a different front, one might wonder whether my claim that belief is phenomenologically transparent might compromise the idea that belief has a distinctive phenomenology. After all, the phenomenology of perception has often been claimed to be transparent as well (see Harman 1990 for a classic statement of the view); and thus one might wonder how my account is able to account for the phenomenological difference between belief and perception.

A detailed answer to this question would depend on a fully fledged account of perception, and this is not the place to pursue this topic. As a starting point, however, it is important to note that what transparency comes down to in the case of perception is very different from what it comes down to in the case of belief. Consciously believing that \( p \) entails taking it that \( p \), or being aware of \( p \) as a fact about the world. Perceptual phenomenology, by contrast, does not determine what one takes the world to be like. Consider, for example, the Müller-Lyer illusion. Familiarly, one’s perception of the Müller-Lyer diagram continues to be as of two lines that differ in length even if one knows that they do not in fact differ in length, and one is in no way tempted to take them to differ in length. This suggests that we should not construe the transparency of perception in terms of taking anything to be the case. Perception is, at most, an invitation to take things to be a certain way. Accepting that invitation involves belief.\(^3\)

How exactly one develops this thought will, of course, depend on the details of one’s views about perception. Still, it might be worthwhile to take a moment to register that the present point fits well with the suggestion made earlier, that perception does not have propositional content. Perception, on this view, is a relation not to abstract entities such as propositions or facts, but rather to concrete particulars in one’s environment. Accordingly, the phenomenology of perception is not a matter of being aware of anything as being the case (since only facts or propositions can be the case), but rather a matter of being aware of those concrete particulars. Being perceptually aware of such particulars, in turn, naturally provides one with the opportunity to become selectively aware of various facts (or purported facts) about them; but this now is a matter of belief, not mere perception. I am, for example, visually aware of the tree outside my window, with its springtime foliage in innumerable shades of green, and this awareness puts me in a position to become aware of a large number of (purported) facts about it as well, depending on my interests and the vagaries of my attention. Becoming aware of such facts about the tree, however, is a matter of forming beliefs about what the tree is like, on the basis of my perceptual awareness of the tree itself.

On this account, perception remains transparent, as it only involves being aware of the environmental objects that one perceives, rather than of anything distinctively mental. But, in contrast to conscious belief, perception does not involve awareness of facts: what one sees when one sees through one’s perceptual states is the concrete particulars populating one’s environment, not facts or propositions. We are thus in a position to recognize that perception and belief are both transparent in some sense, without losing sight of the radical phenomenological differences between the two.

4. Conclusion
My guiding principle throughout this paper has been that we need an account that explains how the phenomenology of belief—what it is like to have an occurrent, conscious belief that \( p \) — explains the rationality of self-ascriptions of belief. As I have tried to show, existing accounts fail in this task, either because they fail to give a plausible account of the phenomenology of conscious belief or because they fail to explain precisely what the link between phenomenology and epistemology is. I have argued that the phenomenology of belief consists in its transparency, and that this suffices to explain the rationality

30. I thank an anonymous referee for urging me to address this point.
31. This point is not refuted by the observation that “\( X \) sees that \( p \)” plausibly entails that \( X \) takes \( p \) to be the case. This is because seeing—that is, plausibly, not a purely perceptual state; rather, it constitutively involves perceptually based belief (and plausibly knowledge) that \( p \). (See also Williamson 2000, 38.)
of self-ascriptions of belief, via the stepping back procedure. Having an occurrent belief entails being aware of the content of the belief as a fact about the world; and a rational subject with the concept of belief is capable of stepping back from that awareness, and self-ascribing the belief.  

References


32. I presented earlier versions of this paper to audiences at the University of Sydney and the University of New South Wales. I am very grateful for the many useful comments I received. I owe special thanks to Melissa Merritt, who read through and commented on several drafts of this paper. Finally, I thank two anonymous referees for Philosophers’ Imprint for their thorough and thoughtful comments.


