Senses of First-Person Authority
Matthew Parrott

If I sincerely tell you that I believe that the Red Sox have a better baseball team than the Yankees, you will usually take what I say to be true and think that I believe that the Red Sox have a better team than the Yankees. In ordinary cases, you do not even consider whether there is additional evidence for what my belief is. You simply trust what I say without knowing anything about how accurate my belief reports have been in the past. What I say about what I believe seems to give you conclusive evidence about what I in fact believe. But why is that?

All of us immediately presume that what someone says about her own psychological states is true; we defer to what a normal psychological subject says about the existence and character of those states. In deferring to her self-ascriptions, we seem to be acknowledging and responding to some kind of special authority intrinsic to the first-person point of view. It seems that we are justified in deferring to a person’s psychological self-ascriptions because she ordinarily has this distinctive kind of authority. This is, at any rate, how things have struck many philosophers who write about first-person authority. For example, Jane Heal notices that "what people say, in the first-person and present tense about their own thoughts is treated as authoritative." And Barry Smith claims that "others do not ask us for justification, because they regard us as authorities on matters of our own psychology." But, although everyday considerations like deference suggest the presence of first-person authority, how should we best understand this phenomenon?

The concept of "authority" has different senses. Sometimes a person has authority in an epistemic sense, by being better situated than anyone else to assess evidence or to make
relevant observations on certain facts. The leading authorities in biology or medicine are epistemic authorities; they either know more or are better placed than others to learn about facts in their respective fields. But a person may also be an authority by being better situated than other people to do something, to exert control over something, to be responsible for something, or to determine something in a unique way. Consider an umpire at a ballgame. The umpire is an authority for whether the pitch just thrown is a strike. Similarly, I am an authority for the grades of papers in the courses I teach. Importantly, although the umpire and I may in fact know better than most people what strikes or passing grades are, our knowledge is not what secures our authority. This is not to say that either of us is ignorant of the relevant facts. The umpire will typically know that the last pitch is a strike and I know the grade for the last paper I read. But neither of us must be in a better epistemic situation than other people in order to have the kind of authority we do. The umpire and I both have authority in an agential sense. We are authorities because we are the only ones responsible for determining facts in a given domain.3

Philosophers traditionally take first-person authority to have an epistemic sense, believing it to consist in each person being in the best position to know about the existence and character of her own psychological states. There are two main problems with this way of understanding the phenomenon. First, there is a great deal of recent psychological research that seems to indicate each of us does not, in fact, enjoy any epistemic advantage over others when it comes knowing about our own psychological states. If things are as this research suggests, there may be no such thing as first-person authority in the epistemic

---

3 I do not intend these two senses to be exhaustive. There are other senses of "authority" but they do not bear on the scope of this paper.
I believe this is a serious difficulty for any epistemic approach to understanding the phenomenon of first-person authority. It is, however, an empirical problem and one might be skeptical about drawing substantive philosophical conclusions from this sort of empirical research. In this paper, I want to focus on a second problem. The problem that I have in mind faces any attempt to account for first-person authority in epistemic terms independently of the results of any psychological.

When we defer to a person’s psychological self-ascriptions we seem to be responding to something exclusive to the first-person point of view; it seems we are acknowledging a special kind of authority that one could possess only from the first-person standpoint. However, an epistemic interpretation of this phenomenon is not faithful to its first-personal character. According to an epistemic view, a person possesses authority in virtue of standing in some kind of privileged epistemic relation to her psychological states. But, as I will argue, it is possible for an individual to stand in this relation to psychological states toward which she adopts a detached third-person perspective. Because such a person can nevertheless enjoy epistemic advantages over others, having epistemic authority does not require that a person to relate to her psychological states in an engaged first-personal way. Therefore, if we understand first-person authority to be something epistemic, if we take the phenomenon to have an epistemic sense, it will not be a special kind of authority exclusive to the first-person point of view. The epistemic sense of first-person authority fundamentally mischaracterizes the phenomenon we are trying to understand.

As an alternative to the traditional epistemic approach, I believe we ought to understand first-person authority in a purely agential sense. An agential account of first-

---

4 This body of research is extensive and seems to grow more every year. A brief overview may be found in Wilson (2002).
5 For recent discussions of this work see Carruthers (2010), Haybron (2007), and Schwitzgebel (2010).
person authority holds that a person has authority because she is better situated to do something with respect to her psychological states. More specifically, I shall argue that when a person relates to them in a first-personal way, she has a capacity to determine what her own psychological states are directly on the basis of her own sense of good reasons for them. Unlike epistemic accounts, the agential account does require a person engage with her psychological states in a first-personal way in order to possess authority.

Before presenting the account, it will be helpful to have a clearer picture of what is distinctive about the first-person point of view. In section one I shall propose two principles that capture essential features of the first-personal way of relating to one’s own psychological states. Notably, it is possible for someone to fail to adhere to these principles. But when that happens she ends up taking a third-person perspective on her psychological states. In section two I demonstrate that a person who relates to her own psychological states in a third-personal way may nevertheless have epistemic authority with respect to them. Because epistemic authority is therefore not exclusively first-personal, I argue that an epistemic approach could not adequately explain the phenomenon of first-person authority.

I begin to develop an agential account in section three. Not only does this account explain a distinctive kind of authority, it also explains the two principles characteristic of the first-person point of view. In the final section of the paper, I contrast my account with other attempts to explain first-person authority in terms of agency. There has been important recent work that has attempted to explain how a person is in an especially good position to know about her own psychological states in virtue of being a cognitive agent. But I argue

---

Bar-On (2004), Ch. 4 makes a similar criticism against epistemic approaches but uses a slightly different example to make her point. She considers the possibility of another person acquiring epistemic expertise on our psychological states and objects that, on the epistemic view, our authority would be transferable. I believe considering a person who takes a third-personal stance toward her own psychological states helps more to bring out the distinctive ways in which our authority is connected to the first-person point of view. Although I am in broad agreement with Bar-On’s criticism of the epistemic approach, I disagree about why authority cannot be transferred.
that even if certain aspects of cognitive agency can plausibly explain why a person has a
distinctive way of knowing her psychological states, they cannot explain how she has an
*epistemically* authoritative or privileged way of knowing about those states.\(^7\) Although we do
have first-person authority because of our agency, this authority is a purely agential
phenomenon.

\[I\]

In order to make an authoritative psychological self-ascription, one that warrants the
deference of others, it seems that a person must relate to her psychological states in an
engaged first-personal way rather than a detached third-personal way. Focusing on
asymmetries between the first- and third-person can help us better understand what is
special about the first-person point of view. In what follows, I will primarily discuss the
psychological state of believing but I think analogous points can be made about other types
of psychological states.\(^8\) Comparing the way I normally relate to my own beliefs to the third-
personal way I relate to another person's reveals characteristic features of the first-personal
way of having a belief that seem to be especially relevant to understanding the phenomenon
of first-person authority.

---

\(^7\) Byrne (2011a) also discusses this distinction. In his terminology, there is a difference between having
privileged access to one’s own psychological states and having peculiar access to them.

\(^8\) As examples, Moran (2001) and Bilgrami (2006) both offer quite elaborate views of first-person authority but
focus on beliefs. Nevertheless, analogous accounts can be given for other types of rationally sensitive
psychological states. The analogies will be with psychological states that are sensitive to our rational judgments,
what Scanlon calls ”judgment-sensitive attitudes” (1998; see also Hieryonymi (2005)). Although he does not
offer an agential account, Byrne (2011b; section V) also believes that a proper account of how a person knows
about her own beliefs will be analogous to how she knows other psychological states. Contrary to Byrne, I am
doubtful that this is true for sensations. Sensations seem to be a distinct kind from psychological states like
belief in a number of ways that will likely matter a great deal to the question of first-person authority. Three
differences strike me as most relevant: 1) sensations are passive; 2) sensations are not, even indirectly,
determinable by a consideration of justifying reasons and 3) sensations have more of an immediate event-like
quality than psychological states like belief that tend to persist (in other words, sensations naturally tend to end
while beliefs tend to endure). I will set the topic of sensations aside for the rest of this paper. But see Boyle
(2009) for an argument against the very possibility of a completely uniform account.
In the first-personal case, my relationship to my beliefs rarely involves any reflection on or observation of them. Rather, I am usually focused on the things around me: my friends and loved ones; the best ingredients for tonight's dinner; or a difficult philosophical passage. My attention and my beliefs are both typically directed at the world outside of me. Most of the time, I am related to my beliefs simply by *believing* things, by being a believer.

This is not true when it comes to my relationship to someone else's beliefs. I must stand in a third-personal relation to your belief and take up some observational perspective or point of view in order to learn about them.

Yet even when a person does self-consciously reflect on her own beliefs (because she is asked about them, for instance), she normally continues to direct her attention out at the world, to qualities of her friend or to her dining options for the evening. When you ask me what I *believe* about living in New York, I will ordinarily answer by considering appealing features of New York: the high quality of restaurants or the accessibility of entertainment. So even in cases where you most expect a person to turn her attention inward toward her psychology, she continues to attend to things in the world outside of her.

This outward directed aspect of our ordinary way of relating to our own beliefs was famously noted by Gareth Evans:

In making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward—upon the world. If someone asks 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 somebody asked *me* whether *you* believed that there will be a third world war, it would be wrong to consider only "the same outward phenomena" that bear on the question whether

---

9 Evans (1982), pg. 225.
there will be a third world war. Reflection on your beliefs requires that I consider them as independent psychological facts from what they are about.

It is crucial to realize that Evans is not saying that when asked about her own beliefs a person must literally explore the world around her or deliberate about any facts in the world. Rather, he claims that only "occasionally" do we "literally" direct our "eye" outside in order to answer questions. For all that Evans has said, it is perfectly reasonable to expect a person to answer questions about her beliefs spontaneously. Evans's talk of directing our "eye" is a metaphor intended to show us that the kinds of considerations relevant to answering a question about one's own belief, as opposed to questions about another person's belief, are not "inner" psychological ones. Sometimes we might have to actually "get ourselves into position" to answer questions by deliberating but often we will, so to speak, already be in the right position.10

It would also be a mistake to conclude from Evans's passage that the reason we attend to "outward phenomena" is because it functions as evidence for our self-ascriptions.11 Rather, as Donald Davidson points out, "first person attributions are not based on better evidence but often on no evidence at all."12 Psychological self-ascriptions, whether in speech

---

10 This is, I think, the most charitable reading of Evans. By contrast, Shah and Velleman (2005) argue that following a procedure like Evans's is only a good way to answer questions when one has not already formed beliefs. On their view, any procedure requiring “doxastic deliberation” risks changing a person's beliefs and is therefore not a good way to answer a question about what one already believes. This argument might seem to pose problem for Evans, but only, I think, if we take his point too literally. The too literal reading is also the source of Gertler's (2011) problems with Evans. She, like Shah and Velleman, argues that active deliberation on a question can have the effect of forming a belief and, for this reason, is not a good way to find out what one already believes. Taking Evans's point metaphorically avoids these objections. Cf. Martin (1998).
12 Davidson (1984), pg. 6. This non-evidential feature of our psychological self-ascriptions is central to what Burge (1988) has in mind when he claims that "brute errors" are impossible when self-ascribing psychological states. An error is "brute" in Burge's sense when it occurs despite the person having a sound epistemic basis. The main reason that such errors are impossible for psychological self-ascriptions is that there is no epistemic basis.
or thought, are usually made independently of evidential considerations. This importantly qualifies how we should understand Evans's talk of directing our attention outward.\textsuperscript{13}

Davidson's point about independence from evidence is a purely negative point about our ordinary first-personal way of relating to our own beliefs; nothing about it entails that this way is epistemically privileged.\textsuperscript{14} As he points out, the fact that self-ascriptions of belief do not normally have an evidential basis does not imply anything about their epistemic status. In order to support a more robust epistemological thesis, one would have to explain precisely what about this particular way of self-ascribing makes it epistemically superior to those that are based on evidence. As Davidson recognized, "claims that are not based on evidence do not in general carry more authority than claims that are based on evidence."\textsuperscript{15}

Davidson and Evans both highlight features of the first-person point of view. Yet a person need not take this point of view on her psychological life. Suppose you were to ask me whether I believed that New York is a nice place to live. I could investigate my personal history in order to answer the question. I do participate in an above average amount of activities in New York. I also frequently tell my friends and family nice things about New York, much more than about Detroit where I grew up. It might be clear to me on the basis of objective evidence that I do indeed believe that New York is a nice place to live. In a case like this, I shift into a third-person perspective on my own beliefs, which is the same kind of perspective another person can take on them.

Thus, in order to engage with her beliefs in a first-personal way, an individual must adhere to the following two principles:

\textsuperscript{13} It also generates a \textit{prima facie} puzzle. If we do not consult "outward phenomenon" for evidence, then why do what Evans says? Why do we attend to these things in the external world if they are not evidence for what we believe? Some philosophers think that resolving this puzzle requires some kind of epistemic explanation. As I shall argue in what follows, there is another explanation.

\textsuperscript{14} cf. Moran (2001) and Byrne (2011a). Peacocke (2008) makes a similar point concerning the distinction between a person's action awareness and her perceptual awareness.

\textsuperscript{15} Davidson (1984), pg. 5
Evans Principle: One’s attention is focused on the way things are in the world, not on her own beliefs independent from the truth of what they represent.

Davidson Principle: One’s belief self-ascriptions are not based on evidence.

Someone can fail to be in accord with these principles but, when a person violates either one, she takes up a third-person perspective on her beliefs. Since it seems clear that a person cannot continue to have first-person authority when she takes a third-person point of view, an adequate account of first-person authority must explain a kind of authority that requires a person to be in accord with these principles. In the following section, I shall argue that this condition cannot be satisfied if we take first-person authority to have an epistemic sense. A person can possess epistemic authority when she violates either the Evans Principle or the Davidson Principle.

II

Let’s consider a case where a person's relation to her own beliefs is obviously third-personal, a case where she acquires knowledge of her beliefs through a process of psychotherapy.

Richard Moran calls this a "familiar therapeutic context" and describes it as follows:

The person who feels anger at the dead parent for having abandoned her, or who feels betrayed or deprived of something by another child, may only know of this attitude through the eliciting and interpreting of evidence of various kinds. She might become thoroughly convinced, both from the constructions of the analyst, as well as from her own appreciation of the evidence, that this attitude must indeed be attributed to her. And yet, at the same time, when she reflects on the world-directed question itself, whether she has indeed been betrayed by this person, she may find that the answer is no or can't be settled one way or the other...She can only learn of it in a fully theoretical manner, taking an empirical stance toward herself as a particular psychological subject.  

16 (2001), pg. 85. The passage is slightly misleading because Moran talks about attitudes like fear in the context of the analysis. But he uses the passage to bring out a truth about his "Transparency Condition", which is explicitly phrased in terms of belief. So, I'm going to assume that what is lying behind the phrasing of this passage is a view about the cognitive conditions on feelings of betrayal; i.e., in order to "feel betrayed" I must
Although I will be concentrating on an example that fits Moran's description, it would be a mistake to think that taking a more third-personal stance requires therapy. Quite often, we learn about what we believe through conversations with colleagues, family members and close personal friends. Moran's example is useful, however, because it vividly illustrates what happens when a person adopts a third-person standpoint on her own beliefs.

Let us call the person in Moran's example "Janet". For my purposes, I will focus on only one problematic belief—Janet’s belief that her sibling betrayed her. At the start of therapy, Janet is completely unaware that she believes her sibling betrayed her. She would, if asked, consistently avoid self-ascribing it and she may even claim to have no opinion on the matter at all. What Moran has in mind is that over time Janet can move to a second stage of therapy and come to learn about the existence and character of her belief. We are to imagine a scenario unfolding over the course of therapy such that Janet eventually begins to self-ascribe the belief that her sibling betrayed her. At first, she may do so tentatively, not quite trusting her therapist. But Janet's therapist, rather than simply telling Janet what she believes, coaxes her along in such a manner that she learns how to appreciate the evidence for herself. Janet begins to notice, for instance, that she is extremely hostile toward her sibling during family holidays or that she feels anxious during their telephone conversations. After months of therapy, Janet learns what she believes and this allows her to knowledgeably self-ascribe the belief that her sister betrayed her. What is important to recognize is that at this second stage, when Janet begins to self-ascribe the belief that her sibling betrayed her, she takes a third-person perspective on it.

---

The characteristics of the distinctively first-personal way of relating to one's own beliefs are clearly absent. Janet believes that she was betrayed by her sibling but her relation to the belief is exceedingly reflective; it is one that she acquires only after hours of therapy. She violates the *Evans Principle* because her attention is not directed outside at features of her sibling that ought to bear on the question of her betrayal. Even when she does consider facts about her sister, they do not seem to matter to what she believes; she must also consider that belief as something that can exist independently from those facts. This is why, in Moran's words, "when she reflects on the world-directed question, whether she has indeed been betrayed by this person, she may find that the answer is no or can't be settled."

Janet also violates the *Davidson Principle*. She bases her self-ascription on "her own appreciation" of psychological evidence. She correctly says that "I believe that my sibling betrayed me", but it is because she observes things like her own uneasiness on the phone or her hostility toward her sibling at Thanksgiving. In violating these two principles, Janet treats the belief that she was betrayed by her sibling in a third-personal manner.

But although Janet takes a third-personal stance toward her belief in the second stage of therapy, she does not seem to be in an inferior epistemic position with respect to it. Through therapy Janet learns what she believes; she therefore knows something about herself--she *knows* that she believes that her sibling betrayed her. Her knowledge, *qua* knowledge, has been based on very good evidence. Moran describes Janet's problem as one where she "cannot learn of this attitude of hers by reflection on the object of the attitude."

But, in claiming this he seems to give more too much credit to epistemic considerations. We can see that ultimately Janet will have to move from the second stage of therapy to a third stage where she relates to her belief in an engaged first-personal way. But that will not really
require Janet to “learn of” her attitude or acquire more knowledge about it.\(^{18}\) Janet already knows the relevant fact about what she believes.\(^{19}\) If this is correct, then what she must develop in the final state of her therapy is not epistemic.

One way to see this last point clearly, is to imagine Janet becoming a sort of psychoanalytic expert.\(^{20}\) In that case, she could make many knowledgeable self-ascriptions, but they would always be based on behavioral evidence, violating the *Davidson Principle*. This would not prohibit Janet’s expertise from enabling her to be in the best position to know about her psychological states. And this may well make Janet a kind of epistemic authority on her psychological life. However, any kind of authority Janet could acquire in this way would not be *first-person* authority because her way of relating to her beliefs would remain too detached and distant. Being an epistemic expert is therefore insufficient for having the special kind of authority exclusive to the first-person point of view.

I think it is clear that Janet lacks a distinctively first-person kind of authority, but the missing authority looks to be connected to her agential capacities to determine what she does believe on the basis of what she thinks are good reasons. When she relates to her belief that she was betrayed by her sibling in a third personal way, Janet lacks the crucial ability to

---

\(^{18}\) This same point was made by Freud who insisted that "If knowledge about the unconscious were as important for the patient as people inexperienced in psychoanalysis imagine, listening to lectures or reading books would be enough to cure him. Such measures, however, have as much influence on the symptoms of nervous illness as a distribution of menu-cards in a time of famine has upon hunger." (XI, 225) Freud’s point here, as elsewhere, is that what must happen in order for psychoanalysis to succeed is not an increase in the patient’s knowledge.

\(^{19}\) Someone could object that Janet actually does acquire knowledge during therapy. On a certain conception of information, the perspective Janet takes on the relation she stands in to her belief actually changes what facts are available. According to this line of thinking, shifting into a first-person relation on a psychological state is a way for Janet to learn something new because it creates a new fact rather than, as I have suggested, being a way for her to relate in a different way to a fact that she already knows. The same kind of issue arises in Frank Jackson’s famous argument about Mary (Jackson, (1982))—Jackson argues that when Mary’s perspective on color changes by leaving the room, she acquires information. Some have challenged the argument by claiming that leaving the room does not give Mary any new information (e.g., Lewis (1983)). I do not have space to adequately address this issue here. My own view is that Janet does not acquire any new information (I also think Lewis’s view is plausible and that Mary does not gain information upon leaving the black and white room).

change what she believes. If we were to ask Janet to reflect on facts about her sibling, to reflect on what Moran calls the "world-directed question", she might conclude that she \textit{ought} to believe that she has never been betrayed by her sibling or that she has no clue why she believes this horrible thing about her sibling. It may be slightly odd but not unintelligible to hear her say something like "I know I should believe that my sibling did not betray me, but I just can't". In ordinary cases, however, when a person judges that a particular belief is unreasonable she thereby extinguishes the belief. Although Janet may have a sense of what a good reason to believe is, that assessment makes no difference to her psychological condition.

One might object that this case does not show that first-person authority is not epistemic but only that epistemic authority must be acquired in a particular way in order to be sufficiently first-personal. However, if we interpret first-person authority in an epistemic sense, the mode of acquiring authority does not seem to matter. On the epistemic interpretation, a person has authority by being in the best position to know about a range of facts. Supposing that Janet were to become a kind of third-personal psychoanalytic expert on her own psychological states, she would be in such a position. So, if deference were merely a response to the epistemic status of a person’s assertions, Janet should be entitled to deference even when she abdicates the first-person point of view. But it does not seem like she is. When we learn that Janet violates the \textit{Evans Principle} and the \textit{Davidson Principle} in making self-ascriptions, we are naturally hesitant to trust what she says. Her third-personal way of acquiring knowledge of her \textit{own} psychological states actually seems to undermine our justification for immediately believing what she says about them. But if her expertise were sufficient to justify trusting what she says, this would not be a response to a distinctive kind of authority exclusive to the first-person perspective.
Other philosophers have raised concerns about epistemic explanations of first-person authority, but these criticisms often focus on explanatory shortcomings of epistemic approaches.\textsuperscript{21} The problem that I am raising is different. I do not think the principal difficulty with epistemic accounts is one of explanatory sufficiency. It is rather that the entire epistemic approach misidentifies the proper explanandum. It explains something that does not require any of the fundamental features of the first-person point of view.

In order to make this criticism more perspicuous, I shall briefly consider one example. Jane Heal’s account of first-person authority has one central thesis: "the existence of a second-level belief about a first-level psychological state is itself what makes it true that the first-level state exists."\textsuperscript{22} Heal wishes to explain how we cannot "just be mistaken" about our own minds, which shows that she takes the epistemic sense of authority for granted. Her purpose, as she states it, is to provide an account allowing her to combine the following three claims:

(1) People’s first-person present tense self-ascriptions of their psychological states are authoritative, while second and third person ascriptions are not authoritative.
(2) Psychological predicates have the same meaning and role in first, second and third person uses.
(3) Persons and their psychological states are among the public and effect-producing occupants of the universe. (pg. 3)

\textsuperscript{21} An important exception is contemporary expressivist theories. Bar-On (2004) and Finkelstein (2003) both criticize traditional epistemic accounts of first-person authority and argue for sophisticated forms of expressivism that deserve a more careful and detailed discussion that is possible in this paper. Briefly, I believe that expressivism is right to think that in deferring to a person’s psychological self-ascriptions we are not responding to the epistemic properties of what she says. Therefore, I think expressivists do identify the correct explanandum. But, the problem for expressivism is one of explanatory sufficiency. The fact that a particular utterance expresses an underlying psychological state does not justify our practice of deference. Just as it may be inappropriate for me to question your outburst of “ouch”, it may also be inappropriate to question other types of expressive speech acts. But, although this may excuse the practice, it does not justify presuming what someone says is true.

\textsuperscript{22} Heal (2001), pg. 4. Heal’s explanation of first-person authority is metaphysical; she accounts for the phenomenon on the basis of the nature of psychological states. It is nevertheless an example of what I call an “epistemic view” because it takes the phenomenon to have an epistemic sense.
Heal thinks that these claims are paradoxical because she thinks it is puzzling how a person could be an epistemic authority for things that are publicly observable. If psychological facts are public, how could one person just be in a better position for knowing about them? The only way to reconcile (1)-(3), according to Heal, is to adopt her constitutive thesis. But this overlooks a prior question. What does (1) mean? In what sense are psychological self-ascriptions authoritative? Heal’s trouble stems from assuming that "authoritative" as it occurs in (1) has an epistemic sense, in thinking she must explain how we cannot "just be mistaken" about our own minds.

Let us suppose Heal is right that my second-level beliefs constitute my first-order psychological states. By believing that I believe that it is raining, I make it true that I believe it is raining. Is this kind of constitutive relation sufficient for explaining first-person authority? Unfortunately, it does not explain a kind of authority that depends on the first-personal way we relate to our own beliefs. We can see this because the account allows for violations of the Davidson Principle. It is completely consistent with Heal’s view that I base self-ascriptions entirely on behavioral evidence. For all Heal has said, I could form beliefs about my own psychological states because of what you tell me or because of the pronouncements of an oracle and that would not prohibit second-level states from constituting first-order states. But there is no first-person authority in these cases. Heal gives us no reason to think her constitutive thesis would only be true if we did form second-level beliefs in accord with the Davidson Principle. She could always say that they have this constitutive power only when they are so formed, but that seems ad hoc. It also gets us no
further in understanding why the first-personal way of relating to our beliefs matters for securing a special kind of authority.\textsuperscript{23}

\textit{III}

If what I have argued so far is true, we should not think of first-person authority as an epistemic phenomenon. This motivates understanding it as a type of agential authority. In this section, I shall argue that a person has agential authority for her own beliefs because only she has the capacity to change or maintain them directly on the basis of reasons. Moreover, only when a person relates to her beliefs in an engaged first-personal way is she able to affect their existence and character in this manner. In addition to explaining a kind of authority exclusive to the first-person point of view, the agential account also explains the \textit{Evans Principle} and the \textit{Davidson Principle}.

When a person is engaged with her belief, she has an attitude fundamentally directed at the truth.\textsuperscript{24} Beliefs are and should be responsive to good reasons for believing, which are, because of the nature of belief, evidential reasons. If a person has a belief that P and is presented with what she takes to be conclusive evidence that P is false, she will normally immediately stop believing that P. Similarly, a person typically need only judge that there is good reason to believe that P and she will immediately believe that P. This is not always the case. A person may, like Janet, have beliefs that do not respond to her judgments about what she has reason to believe. Nonetheless, when a person engages with her own beliefs in a first-personal way, the existence and character of those beliefs depends exclusively on what she takes to be reasons for believing. We all take beliefs to be states responsive to

\textsuperscript{23} As I hope will be clear, once we understand more clearly what makes (1) true, its coherence with (2) and (3) will be obvious and not at all paradoxical.

\textsuperscript{24} Cf. Shah and Velleman (2005).
truth in this way and when we attribute beliefs to others we presume that they respond appropriately to relevant reasons for believing.\textsuperscript{25} 

The fact that the beliefs a person relates to in a first-personal way are responsive to her sense of reasons explains the \emph{Evans Principle}. When we relate to our beliefs in this way, we do not treat them as psychological facts independent from the truths that they represent. When asked whether we believe that P, we attend to what Evans calls "outward phenomena" because we think relevant facts in the world ought to determine what we believe. We look to considerations bearing on the truth of P because these are good reasons for believing that P. Although someone like Janet can also look to these "outward phenomena" when asked about her belief, her belief that her sibling betrayed her is unresponsive to what she thinks she ought to believe. She is therefore not warranted in following the \emph{Evans Principle}. If Janet were to try to answer a question about her belief in accord with the principle, she would give the wrong answer. However, when a person relates to her belief in a first-personal way, her judgments about what she ought to believe \emph{constitute} what she believes. This is why such a person is justified in answering questions in the way Evans suggests.

In order for my beliefs to be responsive to truth I need not form them through deliberation. Many beliefs are causally generated without conscious deliberation or judgment. Nevertheless, a person is committed to the truth of her beliefs in virtue of being \textit{disposed} to adjust them accordingly in the face of evidential or justificatory reasons. When questions about my beliefs arise, I may exercise my capacity to adjust or abandon them. But, if they do not, I remain committed to the truth of what I believe in the sense that my beliefs

\textsuperscript{25} Cf. Stroud (2011), Chapter 4. As Stroud rightly points out "we make sense of such a person's believing the things he does by finding that he comes to believe them on the basis of what he regards as reasons to believe it." (pg. 104) In that chapter, Stroud also draws on an analogy between believing and desiring and intending. This is the sort of analogy that suggests that the account of first-person authority I develop here can be extended to those types of psychological states.
would be responsive to what I take to be good reasons—I would immediately change my belief that P in the face of compelling evidence that not-P.

In everyday life, regardless of how my beliefs are actually formed, only I have a capacity to take responsibility for their truth in this way. Another person does not have to assess the reasons for or against what I believe in order to attribute them to me; she can accurately attribute beliefs to me on the basis of my behavior. But, in the first-personal case, I cannot be indifferent to the truth of what I believe. Since my beliefs express my commitments to the truth, I cannot attribute them to myself on the basis of evidence without undermining or suspending that commitment. If, in order to self-ascribe the belief that P, I first considered something that did not bear directly on the truth of P, I would thereby acknowledge that the existence of my belief may depend on more than whether or not P is true. This helps to explain the Davidson Principle. Since I must treat my beliefs as states whose existence depends only on the truths they represent, I cannot self-ascribe them on the basis of evidence for their existence. Doing so would be a way of disengaging from them and treating them as facts whose existence did not depend only on what I take to be true.

The first-personal way of relating to our own beliefs is therefore intimately connected to our cognitive agency. I am the only person capable of engaging with my beliefs in such a way that my sense of reasons for believing directly determines what I do believe. Having this capacity requires that I adhere to both the Davidson Principle and the Evans Principle; it requires that I stand in the first-personal relation to my beliefs. Although Janet shows that a person can sometimes have beliefs she relates to in a third-personal way, this cannot be the case for most of a person's beliefs. If the majority of Janet's beliefs were insensitive to her assessment of what she ought to believe, attributing beliefs to her would
begin to lack intelligibility. Normally a person is engaged in a first-personal way with what she believes such that her judgments about what she should believe constitute what she does believe. I believe that first-person authority is derived from this aspect of our cognitive agency:

*First-Person Agential Authority:* A person X has first-person authority for a belief B iff X is the only person with a capacity to directly determine B on the basis of justifying reasons.

Only *my* sense of what counts as a good reason to believe something can affect my beliefs in this way, can cause me to believe or cease to believe that something is so. You might have a better grasp on what is, in fact, good evidence or be far better at assessing reasons but this, by itself, has no effect on what I believe. This capacity to directly determine beliefs on the basis of reasons is unique to the first-person and it is why we each have a special kind of first-person authority.

Someone might object, however, that Janet does have this capacity and merely fails to exercise it. If so and if Janet lacks authority for her belief that her sister betrayed her, my account is wrong. According to this objection, Janet would fulfill the second half of the bi-conditional, she would be the only one with the capacity to determine her beliefs for reasons, but she would not have any first-person authority. This objection is not trivial because if anyone is going to have the capacity, it is going to be Janet.

---

26 We would have a difficult time attributing a type of psychological state that we take to be fundamentally responsive to reasons, while knowing it is unresponsive to reasons in Janet’s psychological life. Cf. Stroud (2011).

27 Unfortunately, a full explanation of why this justifies others in deferring to our psychological self-ascriptions is beyond the scope of this paper. For the details of this explanation, see XXXX.
Part of my reason for going into the case of Janet was to bring out how the capacity to determine beliefs for reasons depends on the way a person relates to them. It does not depend on mere ownership of beliefs. I think it is clear that Janet lacks this capacity with respect to the belief that her sister betrayed her. When she tries to follow the *Evans Principle* by considering characteristics of her sister she can only judge that they support the opposite of what she actually believes. Her judgments about reasons for believing do not affect what Janet actually believes. Janet thinks that she ought to believe that her sister did not betray her but she cannot get herself to believe it. I do not think there is a point to insisting Janet has some hidden capacity when she is trying as hard as she can to exercise that very capacity and failing. Janet is simply incapacitated.

Another reason to think that someone like Janet lacks this capacity is that this explains why a person lacks authority whenever she takes up a third-person perspective. Recall that on an epistemic interpretation, Janet need not lack authority when she takes a third-person stance toward her beliefs. As we saw, she could acquire a kind of third-personal expertise and thereby come to have epistemic authority. On the epistemic view, the connection between Janet's authority and her standing in a first-personal relation to her belief is obscure or contingent. It is not clear why taking up a third-person point of view undermines Janet's authority. But, by adopting the agential account, we can see fundamental connections between our authority and the agential capacities intrinsic to the first-person perspective.
Recently, a number of philosophers have appealed to agency to try to explain first-person authority. For example, Richard Moran, in *Authority and Estrangement*, proclaims that a proper understanding of the first-person “will require bringing the agent more explicitly into the picture.” Although I believe there is a great deal to be learned from Moran's book, I think his view faces some difficulties. This is primarily because, despite his emphasis on the importance of agency, Moran seems to believe that first-person authority includes an epistemic element. He claims that a central task of his book is to explain why psychological self-ascriptions made in the first-person way "enjoy a particular epistemic privilege not accorded to corresponding third-person judgments." The primary problem for Moran’s view is that agency alone cannot explain this. It cannot account for an epistemically privileged way of knowing about our own psychological states.

On Moran’s view, we are entitled to answer questions about our own beliefs without having to appeal to evidence. Instead, a person can answer questions about "whether he believes that P in the same way he would address himself to the question whether P itself. From the first-person point of view, the one question is treated as 'transparent' to the

---

28 In addition to Moran, Bilgrami (2006) and O’Brien (2007) offer the most developed accounts. But Boyle (2009) also defends an agential account similar to Moran’s.
29 (2001), pg. 33.
30 He does not take it to be purely an epistemic phenomenon. Rather, Moran thinks that first-person authority has both what he calls an epistemic dimension and an agential dimension. He seems to think, mistakenly in my opinion, that the two are interdependent. See Moran (2001), pg. 92.
31 Moran (2001), pg. 10.
32 The exact same problem confronts Lucy O’Brien’s (2007) account of self-knowledge in terms of agency. O’Brien is a bit more explicit than Moran that she intends to explain an epistemically authoritative kind of self-knowledge. She defines first-person authority as follows: “An agent seems to be in a better position to know her own actions than others are.” (pg. 158) As I will argue in this section, agency does not necessarily involve this. Unlike O’Brien and Moran, Bilgrami (2006; 2010) explains first-person authority on the basis of constitutive connections between beliefs and a person’s knowledge of them. For this reason, his view does not face this problem. It does, however, have a difficult time accounting for the possibility of error. Space prohibits a fuller discussion of Bilgrami’s account in this paper but see XXXX.
other.”

When a person answers questions by following this procedure, they utter something Moran calls an "avowal". Answering by "avowing" is a way of adhering to both the Evans Principle and the Davidson Principle. Moran thinks we are entitled to do this because we must assume that we can "make up our minds" about what we believe. I agree that we are justified in answering questions just as Moran indicates and that having this agential capacity grounds first-person authority, which I believe is the most significant insight of his book. But Moran also thinks this explains why our judgments enjoy “a particular epistemic privilege” or why we have epistemic authority. This may be because Moran believes that avowing explains our special epistemic access to our own beliefs. He sometimes even describes avowing as a "form of awareness" or as a way of having first-person non-evidential access.

I think it is a mistake to think that a first-person "form of awareness" can be explained by a procedure used to answer questions about our beliefs. It rather seems that we

---

33 Moran (2004), pg. 457. See also Moran (2003), pg. 410. Because a question about one's beliefs is treated as "transparent" to a question about the world, Moran calls this the "Transparency Condition". I prefer to avoid the term "transparency" because it conjures images of psychological states being self-intimating (this is, for instance, the way that Bilgrami (2006) uses the term).

34 This is obviously a special sense of "avowal". Moran defines it as a statement of one's beliefs that is made in accord with his procedure. Moran (2001) pg. 101.

35 Many times Moran writes that our rational agency entitles us to "assume" that we are able to answer questions about our own psychological states without an appeal to evidence. For instance, he claims that "a person is entitled to this assumption insofar as his answering the question proceeds from the understanding that his sense of the reasons in favor of P itself does determine what his belief about P is." (2004, pg. 466, my emphasis) The idea that we must "assume" something like this before we adhere to the Davidson Principle strikes me as incorrect. It makes it seem as if sophisticated aspects of our cognitive agency must be apprehended in order to legitimize our non-evidential self-ascriptions. I think it is better to say that we, as agents, are entitled to adhere to the Davidson Principle without having to assume anything at all. But perhaps this is all Moran intends by his talk of "assuming".

36 I read Moran differently from Boyle (2011). On Boyle’s reading, Moran’s central idea is that certain psychological states are not distinct from our knowledge of them. In Boyle’s words, “believing P and knowing oneself to belief P are not two cognitive states; they are two aspects of one cognitive state.” (pg. 228) Interestingly, this reading would align Moran with other recent philosophers—Bilgrami (2006) and Shoemaker (1994) in particular—who think our psychological states and our beliefs about them are not distinct existences. However, as I read him, the central philosophical importance of Moran’s work is in its attempt to understand our distinctively first-personal way of relating to our own psychological states in a way that does not deny that they are all distinct existences but also does not collapse into understanding this relation as a kind of inner perception.

37 Moran (2001), pg. 134.
must presuppose some distinctive first-personal form of awareness in order to answer questions by avowing. There would be no point to trying to answer questions about whether I believe that P by considering facts that bear on the truth of P if I did not have some kind of first-person access to my cognitive actions. For it is not passive contemplation of P that determines what I believe but my judging whether or not evidence indicates that P is likely to be true. That is, I must consider relevant facts to be reasons for believing P. My judgment that they are good reasons need not be explicit; but it is nevertheless my judgment that determines whether or not I believe that P, not P itself and not facts bearing on the truth of P that I do not understand as such.  

This seems to require some first-personal form of awareness or access to my own cognitive activity. I must be aware, for example, that I take Q to be a reason in favor of believing P. This is especially true if my consideration of Q is supposed to be relevant to my answering questions about whether or not I believe that P. If I were not aware of what I judged, no amount of attention paid to “outer” facts could help me learn what I believe. If, as Moran claims, part of our entitlement to avow rests on our capacity to "make up our minds", part also rests on our being aware of what we are "making up".

For the moment, let's set this worry aside and suppose that Moran can explain first-person access fully on the basis of our having the ability to avow. Does this mean we have epistemically privileged or authoritative access? The way we know our own beliefs is clearly distinct but is there any reason to think it is epistemically superior? Moran never attempts to

---

38 These judgments should not be thought of as a response to a question. Nor should they be thought of as overly intellectual or deliberative. In a very ordinary way, I can look outside and see that it is raining and also judge that it is raining. But I do not have to see it before I judge it. Afterward, I may also declare to you that it is raining, but my declaration should not be confused with my judgment.

39 Peacocke (2008) makes a similar claim. He argues that having a capacity to engage in what he calls "mental actions" requires having a special kind of access to those actions.
explain why it would be superior to other ways of knowing. He rather seems to take this for granted.

Suppose, however, that our way of knowing our own beliefs is not epistemically privileged. I do not think this would undermine any of the central aspects of Moran's account. It would still be true that we are entitled to adhere to the Evans Principle and the Davidson Principle; in Moran's terms, we are justified in answering questions about our beliefs by avowing. Adhering to these principles requires having some kind of first-person access. Yet, by hypothesis, this is not an epistemically privileged mode of access. That does not prohibit our being justified in answering whether we believe that P "in the same way" we would answer the question whether P. Contrary to what Moran claims, these features of cognitive agency do not require us to have any "particular epistemic privilege".

If I have a special way of accessing facts that is unavailable to others, don't I thereby enjoy an epistemic privilege? One might think that having a distinctive from of awareness is equivalent to having a privileged way of knowing.\footnote{One place where this is explicit is in Boyle (2009), who writes: "self-ascriptions of these kinds of mental states are not normally liable to the same kinds of error that afflict ascriptions of such states to other people," which Boyle writes is "a manifestation of their authority." But even if my self-ascriptions are not liable to the "same kinds of error" they may nevertheless be subject to different kinds of error. One should not infer that a way of knowing is epistemically privileged just because it is unique. Cf. Peacocke (2008) and Byrne (2011).} This is mistaken. Having a distinct mode of awareness doesn't entail that it is epistemically privileged. There is a strong temptation among philosophers to think otherwise. I believe this is because the entailment would hold if we made further assumptions. Two in particular seem to be common. First, it is sometimes thought that the existence of a person's psychological state is not distinct from her knowledge of its existence.\footnote{Shoemaker (1994), Bilgrami (2006).} Second, it is often assumed that our first-personal way of knowing is more reliable than ways of knowing requiring perception. Both of these are
substantial assumptions. Either one, if true, could help explain a person's epistemic privilege over others but, without one of these assumptions, our cognitive agency cannot.\footnote{I believe that both assumptions are false. I criticize the first in XXXX. As I mentioned earlier, recent psychological research seems to indicate that the second is false.}

Some philosophers have voiced a different objection to Moran's view. They argue that many beliefs are not up to us. Shoemaker puts this objection as follows:

I know that I am wearing pants, so I believe that I am wearing pants. I know and believe that I believe that I am wearing pants…But it is hard to think of circumstances, other than those of a dream, in which it could be a question for me whether I believe this. I would also have a hard time saying what reasons I have for believing it. And I cannot think of any good sense in which it is 'up to me' whether I believe it.\footnote{Shoemaker (2003), pg. 396.}

It seems to me that there are two separate issues in what Shoemaker says here. Shoemaker highlights the familiar fact that we cannot adopt or discard beliefs voluntarily. If that were required for first-person authority, an agential account would be hopeless. But we can easily avoid the worry because, as we have seen, beliefs are fundamentally constrained by evidence, reasons, and justification.\footnote{See Moran's reply to Shoemaker (2003) and also Moran (2004).} Yet Shoemaker also objects that much of the time we cannot even make up our minds in this sense. Many beliefs, like the one that I am wearing pants, are ones I simply find myself with.\footnote{See Heal (2004) for the same criticism.} I believe that I am wearing pants, that the earth exists, and many other things besides that I never deliberated about and probably never will.

Shoemaker's criticism poses a challenge to thinking of first-person authority in terms of agency. When it comes to a belief like the one that the earth exists, my mind is simply made up and it will stay that way no matter how much attention I pay to it. It is difficult to imagine what a rational reconsideration of this belief could even look like. Nevertheless, these sorts of beliefs are ones that I relate to in a first-personal way, I am not detached or
disengaged from them like Janet is from her belief that her sister betrayed her. But how could I have agential authority for beliefs that I never reflect upon?

This objection brings out why we should set aside any language suggesting a person must actually “make up” her mind in order to possess first-person authority. Intuitively, first-person authority does not seem to require much deliberation. However, as I have repeatedly emphasized, we have a capacity for cognitive agency whether or not we activate it in deliberation.46 Because of this, my belief that I am wearing pants would be extinguished were I to see that I was not clothed. My cognitive agency involves being disposed to respond to the appropriate reasons. This is how my beliefs are supported by reasons even when they are not objects of critical reflection. If the reasons do not change, I would not display this sensitivity to reasons but I have the capacity nonetheless provided that my beliefs would respond appropriately in the relevant possible circumstances, no matter how unlikely.

Sensitivity to reasons for believing is likely to come in degrees, to be something that has its own characteristic failures, good days and bad days. But we lose sight of it altogether when a person abrogates her cognitive agency by adopting a third-person point of view on her own beliefs. For any given belief that a person has, so long as her own sense of reasons for believing determines whether she maintains, reforms, or abandons it, it is one for which she alone has authority.

46 See Raz (1997) for a discussion of agency without direct action.