In order for me to become aware of another person's psychological states, I must observe her in some way. I must see what she is doing or listen to what she is saying before I can know what is on her mind. Even in cases where I learn about another person's psychological states more indirectly from testimony, somebody had to observe her behavior. Because observation is, either direct or indirectly, necessary for becoming aware of another person's psychology, I can be, and often am, mistaken about what other people are thinking or feeling. Notably, it is possible for my perceptual system to generate errors even when I am completely justified in how I form beliefs about another person's psychological states. The possibility of this kind of error in particular demonstrates that another person's psychological state is ontologically independent from my beliefs about it; the two are, in Hume's words, distinct existences. Since there is no necessary connection between another person's psychological states and my own, if they are related, it is because of some contingent causal relation.

Things are different in my own case. I do not normally have to make observations in order to know what I am thinking or feeling. Instead, it seems that I have a special kind of epistemic access to my psychological states. This first-person access is unavailable to other people. They can never come to know about my psychological states in the special way that I do. Having this mode of access, however, makes it far less clear whether my own psychological states are distinct from my beliefs about them. Indeed, many philosophers believe that they are not, thinking instead that they are in some way interdependent.¹

¹ There are various ways to spell out this interdependence. Sydney Shoemaker's view will be the primary focus of this paper. But Bilgrami (2006) and Heal (2001) describe the interdependence differently. For critical discussion of the concept of ontological dependence see Fine (1995).
This idea is endorsed explicitly, for example, by Akeel Bilgrami, who claims, "our mental states lack a certain kind of independence from our knowledge of them...possessed by the things of which we have perceptual knowledge, from that perceptual knowledge."\(^2\) Bilgrami proposes that a person's knowledge of her own psychological states is "not come by via anything at all...it comes with the states of minds themselves, for 'free', as it were."\(^3\) This suggestive metaphor of "knowing for free" is intended to capture the basic idea that a psychological state is not distinct from its owner's belief (or knowledge) of it.\(^4\) In addition to Bilgrami, several recent philosophers have defended the idea that a person's psychological states are constitutively connected to her own beliefs about them.\(^5\) These "constitutive theories" of self-knowledge aim to account for first-person access by grounding it in necessary, constitutive relations.

Despite the popularity of constitutive theories, they oppose a natural and simple thought: a psychological state is one thing and a belief about it is something else. This is obviously true with our beliefs about psychological states that belong to others. Why is it not equally true in our own case? Why should we not think that a person's psychological states and her beliefs about them are also independent from each other? I propose to capture this notion with the following thesis:

---

\(^3\) Bilgrami (2006), pg. 38.
\(^4\) Although I am not aware of anyone who holds this view, someone persuaded by Williamson’s (2000) account of knowledge as a distinct psychological type could think that a person’s psychological states are interdependent with her states of knowledge but not with her beliefs. The arguments of this paper would apply mutatis mutandis.
Distinct Existence Thesis: For any subject $a$ and psychological state $M$, $\sim (M(a) \text{ only if } a$ believes that $M(a))$.  

Plausible considerations speak in favor of the Distinct Existence Thesis. First there is Hume's well known opinion that, since all psychological states are distinguishable from each other, "they may exist separately, and have no need of any thing else to support their existence." Hume's point strikes me as prima facie true. It does seem that distinguishable psychological states could exist without each other. Second, we are both ignorant of and mistaken about many of our own psychological states. We can believe or desire something without believing that we do and we can wrongly describe our own feelings, desires or beliefs. How does this happen? Why do we make errors about our own psychological states? The combination of ignorance and error in this domain suggest that the Distinct Existence Thesis is true. Even if they turn out to be quite exceptional, we should nevertheless expect an account of why and how they occur whenever they do. Yet the most promising explanations of error will rest on the independence of psychological states from beliefs about them. They will maintain that mistakes occur because a contingent causal relation connecting a person's psychological states with her beliefs breaks down.

---

6 Dependence is also a consequence of an infallibility thesis about self-knowledge. I shall not discuss this sort of view because few people, if any, believe we are infallible about our psychological states. However, if necessary, we can reformulate the Distinct Existence Thesis in a way that accommodates infallibility. In that case, the thesis becomes a bi-conditional: For any subject $a$ and psychological state $M$, $\sim (M(a) \text{ iff } a$ believes that $M(a))$.

7 *Treatise* 1.4.5.5.

8 This point is related to Moore's Paradox. The paradox arises because the conjunction (P and I believe \~P) is not a logical contradiction. Does Moore's Paradox occur at a higher level? Consider the following: I believe that it is raining but I believe that I do not believe that it is raining. Someone who rejects the Distinct Existence Thesis could claim that this conjunction does entail a logical contradiction on the grounds that the two conjuncts were not really distinct facts. Thus, how we think about Moore's Paradox is closely related to how we think about the Distinct Existence Thesis.
But if commonsense intuitions like these support the *Distinct Existence Thesis*, why do philosophers endorse constitutive theories? One reason can be found in a well-known argument presented by Sydney Shoemaker in a number of papers. Shoemaker argues that if a person's psychological states were ontologically independent from her beliefs about them it would be possible for someone to suffer from a condition he calls "self-blindness". A person suffers from self-blindness just in case she can know about her psychological states in only a third-personal way. But, Shoemaker argues, the possibility of self-blindness is absurd so a person’s psychological states cannot be fully distinct from her beliefs about them. The impossibility of self-blindness is taken by many to show that some sort of constitutive view must be true.

My purpose in this paper is to defend the *Distinct Existence Thesis*. I shall do so by arguing in the following two sections that even if self-blindness is impossible the best explanation for this need not deny the *Distinct Existence Thesis*. Shoemaker's argument is motivated by his conviction that first-person access is epistemically unique. He thinks that if the *Distinct Existence Thesis* were true our first-person way of accessing psychological states would have to be based on contingent causal relations. But, in that case it would no longer be epistemically distinctive; it would too greatly resemble perceptual knowledge. In section three, I offer an alternative account of the epistemic entitlement to first-person access that does not base its justification on causal relations. We should not deny that first-person access is epistemically distinct but I do not believe this means we must abandon the simple thought embodied by the *Distinct Existence Thesis*. In the final section of the paper, I return to the commonsense intuitions that favor *Distinct Existence Thesis* in order to develop them more fully. I argue that the thesis is preferable to constitutive theories because it allows us to make sense of our fallibility concerning our own psychological states. Since it seems clear that we
can be mistaken about many of our own psychological states, we should avoid theories that prohibit our fully understanding these sorts of errors.

I

According to Shoemaker, “a self-blind creature would be one which has the conception of the various mental states, and can entertain the thought that it has this or that belief, desire, intention, etc., but which is unable to become aware of the truth of such a thought except in a third-person way.”\(^9\) Self-blindness is supposed to be analogous to color-blindness. Just as a person who is color-blind can learn information about color in a non-standard way, through reading a book about colors for instance, a person who is self-blind can learn about her own beliefs, desires, and intentions in a non-standard way, which is to say a completely third-person way.\(^10\) Instead of having first-person access to her own psychological states, the self-blind person will have to make observations of her behavior in order to know what she believes, wants or feels. Moreover, just as the color-blind person does not suffer from a cognitive or conceptual deficiency, self-blindness is "supposed to be perceptual or quasi-perceptual, rather than cognitive or conceptual."\(^11\) Shoemaker argues that if the Distinct Existence Thesis were true, self-blindness would be possible. It would be possible for someone to have the same kinds of psychological states any of us do without having any of the normal first-person access that we do.\(^12\)

Although Shoemaker tailors his argument to many different types of psychological states, each variation rests on a type of thought experiment that asks us to try to imagine a

---

\(^12\) (1994)
self-blind person with respect to a specified type.\textsuperscript{13} I will focus on the case of belief but the argument is analogous for other psychological kinds. When he discusses belief, Shoemaker asks us to try to imagine a self-blind man he names George.

If George were self-blind, we should be able to imagine that "the total evidence available to a man at a given time should support the proposition that it is raining, while the total 'third-person' evidence available to him should support the proposition that he does not believe that it is raining."\textsuperscript{14} In this case, it would seem reasonable for George to assert a Moore-paradoxical proposition like "It is raining, but I do not believe that it is raining," because this is what his total evidence supports.\textsuperscript{15} But, Shoemaker points out that any rational person can recognize that a Moore-paradoxical assertion is inappropriate. Since we are supposing George to be rational and not conceptually or cognitively deficient, we can see that he will avoid asserting a Moore-paradoxical proposition. Hence he would behave in the same way any of us would.

Because George has mastery of the concept of belief, Shoemaker thinks the following two things will also be true:

A) He will recognize that when asked "'Do you believe that P?' he ought to answer 'yes' just in case he would answer 'yes' to the question "Is it true that P?"

\textsuperscript{13} The different versions can all be found in Shoemaker's Royce Lectures (1994).
\textsuperscript{14} (1988) pg. 118.
\textsuperscript{15} There is another puzzle here. Shoemaker claims that the total third-person evidence might justify the assertion of a Moore-paradoxical sentence. Offhand, it may seem that this evidential situation could arise because there are two distinct facts involved; one about George's psychology and one about the weather. However, I think that when one really tries to imagine the situation it becomes quite difficult to picture the totality of George's evidence unambiguously supporting both conjuncts of a Moore-paradoxical proposition. Suppose we include George's utterance of "it is raining" as one component of the "third-person" evidence. We will no longer have an evidential situation that conclusively supports George asserting "it is raining but I don't believe that it is raining." Shoemaker oddly seems to grant this at one point, but says that "since the objection under consideration" depends on an inconceivable situation (unambiguous evidence for a Moore-paradoxical proposition) and his argument does not, the objection is defeated. But, the problem for Shoemaker is that it might be that the reason everyone avoids a Moore paradoxical assertion is because of objective third-personal evidence, because evidence always clearly supports only one conjunct. If that were true, it would be unclear how the phenomenon helps us understand a distinctively first-personal mode of access. Cf. Shoemaker (1995).
B) He will recognize the meaning of 'believe' and "preface his assertions with 'I believe' in just the circumstances in which this is pragmatically appropriate.

Plausibly, A) follows from George's being a rational believer with the concept of belief.\(^1\) When asked whether or not you believe that P, you, like George, usually consider the truth of P.\(^1\) The concept of belief is that of a psychological state responsive to the truth it represents. Understanding this connection between belief and truth is a large part of having the concept.

Having the concept of belief also means George can appreciate the relevance of pragmatic considerations for self-ascribing beliefs. For example, when I lose my keys, I have to look for them. Where are they? I don't really know and there is no evidence indicating their location. But, I believe they are in my office. In this sort of case, my saying "I believe the keys are in my office" is a way of making a hesitant assertion about the location of my keys. Since he has conceptual mastery, George could behave the same way I do when I lose my keys. Like me, he could say "I believe that the keys are in my office" in the appropriate contexts.

Shoemaker therefore concludes that George will behave just like any normal person and self-ascribe beliefs in the exact same conditions any of us would. There would "be

---

\(^1\) I actually have some doubts about whether A) is true for George. Shoemaker thinks that the truth of A) is central to our concept of belief. But, it seems to me that part of why ordinary believers can answer the question "Is P true?" when asked about their beliefs is because they are entitled to presuppose that they have first-person access to their own cognitive actions. If I did not know what I judged when answering "Is P true?" it is hard to see how I would be rational in considering "Is P true?" as a way of answering a question about my beliefs. Thus, even though A) is true for us, it may not be fundamental to our concept of belief. I will grant that A) is true for George for the sake of argument but will return to these issues in section III. For a different criticism, see Brueckner (1998).

\(^1\) This is a point familiar to readers of Evans: "In making a self-ascription of belief, one's eye's 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 \(p\)." (1982, pg. 225)
nothing in his behavior, verbal or otherwise, that would give away the fact that he lacks self-acquaintance." When we try to imagine George, we can only imagine a person who behaves like us, which means that we cannot imagine a distinctively self-blind person. This, however, just means that self-blindness is not a genuine possibility. If it were, Shoemaker insists, "there should be something that would show, or at any rate provide good evidence, that someone was afflicted with such self-blindness."  

It is important to keep in mind that George is supposed to be completely unable to become aware of his psychological states except in a third-personal way. Even if we can conceive of him lacking first-person access to some psychological states, this is not self-blindness. After all, we lack first-person access to some of our psychological states some of the time, but none of us are self-blind. To truly conceive of a self-blind person, we need to imagine a person as sophisticated as we are both psychologically and conceptually without any first-person access at all. Is this possible?

I agree with Shoemaker that it is not. It is certainly difficult for me to imagine a fully rational person with a psychological life as rich as an ordinary adult human without that person having first-person access to her own psychological states. Such a person would always have to observe her own behavior to learn about her psychological life but then it can easily seem that she might have no such life to learn about. Without some prior first-personal access to her own thoughts, why would she take her own observed behavior to even be intentional behavior? Why would she not instead take it to be a series of involuntary movements? Since the person has no direct first-person access to her beliefs or desires it does not seem she has any reason to take herself to even be a cognitive agent, presuming that she can take herself to be anything at all.

19 Kind (2003) thinks it may be.
Although I personally find these considerations compelling, I don't believe they are conclusive. In fact, I think one might try to make a case for the conceivability of self-blindness by imaging someone losing first-person access late in life, after acquiring a great deal of rational maturity and conceptual skill. Just as it is difficult to imagine someone blind from birth mastering color concepts, it is hard to picture someone who has always been self-blind with the necessary resources to self-ascribe psychological states. In both cases the difficulty might be overcome by imagining the person gradually losing the specific mode of access. If we try to imagine George slowly losing his capacity for first-person access, might we have more success? I think we might and it might be instructive to try, but I still find self-blindness incoherent. So, for the remainder of this paper, I shall assume that it is impossible in order to see what this means for the Distinct Existence Thesis.

II

Does Shoemaker's argument from self-blindness refute the Distinct Existence Thesis? He himself characterizes the target of the argument as any view that holds "the existence of these states and events is independent of their being known in this way, and even of there existing the mechanisms that make such knowledge possible." Notice, however, that in describing his target this way Shoemaker conflates two separate things. First, there is the possibility that the "existence" a psychological state depends on another state, namely that of knowing the first. If this were the case, it would be a mistake "to think of the second-order belief as a distinct state that is caused by the available first-order belief."20 But Shoemaker also mentions a second possibility; that a psychological state depends on there being what he calls "mechanisms that make such knowledge possible". These "mechanisms" can be

20 (1995), pg. 92
thought of simply as whatever process realizes our capacity for first-person access. It is helpful to distinguish these two.

**State Dependence:** For any subject $a$ and psychological state $M$, necessarily, $M(a)$ only if $a$ knows that $M(a)$.

**Capacity Dependence:** For any subject $a$ and psychological state $M$, necessarily, $M(a)$ only if $a$ has a distinct capacity to know that $M(a)$.

*State Dependence* requires that constitutive, rather than causal relations connect a person’s psychological state with her knowledge of it. On the assumption that knowing that $P$ requires believing that $P$, *State Dependence* is a thesis most constitutive theorists would endorse. However, because the most promising constitutive views wish to be consistent with recent evidence from social psychology indicating that we sometimes lack awareness of our psychological states, *State Dependence* should also be understood as a generic claim. It is a stronger claim because *Capacity Dependence* follows from it trivially. Nonetheless the two are logically distinct.

Since the *Distinct Existence Thesis* implies that any of a person's psychological states can exist independently from her having a belief about it, it is inconsistent with *State Dependence*. But does the impossibility of self-blindness really show that the *Distinct Existence* is false? Does it show that there must be necessary connections between a person’s psychological states and her own beliefs about them? Shoemaker says that the thesis entails "that for each kind of mental fact to which we have introspective access, it is at least logically

---

21 It would therefore strengthen both to read the claims universally. But, I assume that the stronger version of either would be unacceptable to most philosophers; although the stronger form of *State Dependence* does seem to have been Descartes’ view.
possible that there should be creatures in which such facts obtain, and who have the ability
to conceive of them, but who are self-blind with respect to them.” However, even if it is
possible for a creature to be "self-blind" with respect to "each kind of mental fact", it may
nevertheless be impossible for that creature to be self-blind with respect to all of them. Indeed, this is precisely what would happen if *Capacity Dependence* alone were true.

In describing George, Shoemaker argues that he is be unable to become aware of *any*
psychological states except in a third personal way. If George were self-blind to all of his
psychological states, he would lack a *capacity* to know about his psychological life in a
distinctively first-personal way. Assuming that George’s self-blindness is impossible, the
argument demonstrates that a person could not have psychological states in the same way
that we do without having a capacity for first-person access to them. *Capacity Dependence* must
be true. But why is it true? Why must a subject have a distinctive capacity for first-person
access? The simplest answer to this question, which I think Shoemaker favors, is that *State
Dependence* is true; a person has a capacity for first-person access because her having
psychological states constitutively entails that she knows about them. But the downside with
this answer is that it requires we give up the *Distinct Existence Thesis*. If we wish to hold onto
it, we will need to explain *Capacity Dependence* another way.

---

22 (1994), pg. 273
23 Consider a slightly imperfect analogy. Being a student is independent from receiving an 'A' in a Philosophy
101. Thus, for any student in my class, it is possible that she does not receive an 'A'. The two states are
ontologically distinct. But it may nevertheless be impossible for no student to receive an 'A' in this course.
Suppose, plausibly, that if it began to look as if all the students in the course were going to receive less than an
'A', the grading scale would be recalibrated (this could be a part of the nature of grading for an introductory
university course). In this case, it would be impossible for every student to not receive an 'A', although it is
possible for each student.
We might consider whether Capacity Dependence could be grounded in the nature of psychological subjects rather than in the nature of psychological states. Shoemaker sometimes writes as if he has this sort of thing in mind. For instance, he endorses the idea "that it is of the essence of mind that each mind has a special access to its own contents, or more soberly expressed, that each person has a special access to his own mental states." But if "the essence of mind" includes "special access", couldn't "the essence of mind" explain Capacity Dependence? It is, as we have seen, very difficult to imagine a person without first-person access, but perhaps this is because having the capacity for this kind of access is a fundamental part of what it is to be a rational subject of psychological states.

A central aspect of being such a subject is having attitudes like beliefs, desires, intentions that are responsive to appropriate reasons. My beliefs, for example, are normally sensitive to evidence and to justifying reasons bearing on the truths they represent. If I believe that P and am confronted with what I take to be conclusive evidence that P is false, I will, if rational, immediately stop believing P. In this way, my assessment of the world, my take on evidence for or against P, makes an immediate and substantial difference to the existence and character of my psychological state. As a subject whose attitudes are responsive to reasons, my sense of good reasons for or against them directly affects them and has the potential to change them.

It seems to me that in order for a person's sense of reasons to be able to directly determine what her psychological states are she must have a unique capacity for first-person access. Tyler Burge has recently stressed that this is necessary for an individual to engage in critical reflection on her psychological states. He claims that "it is constitutive of critical

---

24 (1988), pg. 115
reasoning that if the reasons or assumptions being reviewed are justifiably found wanting by the reviewer, it _rationally follows immediately_ that there is prima facie reason for changing or supplementing them, where this reason applies within the point of view of the reviewed material (not just within the reviewing perspective).\(^{25}\) Someone reasoning critically in Burge's sense must turn her attention toward her own psychological states in order to critically assess and evaluate them. This often culminates in a judgment to the effect that a particular state is warranted or not by reasons, a judgment that Burge notes necessarily has the potential to _immediately_ change the original state. Burge thinks that if our way of knowing our own psychological states rested on some kind of epistemic intermediary, "there would never be an immediate rationally necessary connection" between, for instance, what we deliberatively judge we ought to believe and what we in fact believe. Why not?

In order for someone to engage in critical reasoning about her psychological states, she must be aware of them. Let's suppose she does so indirectly or in a more third-personal manner. In that case, she would learn about her psychological states on the basis of evidence or observation. This is the same kind of epistemic access we have to those psychological states belonging to other people and it seems that we can certainly rationally criticize what others ought to believe, want or intend. Yet when it comes to others our judgments are more like appraisals or recommendations; they do not necessarily determine the existence or character of the other person's psychological state. Depending on how convincing we are, the person may accept our judgments, but she may also ignore them. When someone takes up this kind of third-person perspective on herself, she treats her own psychological states

\(^{25}\) (1996), pg. 257. Burge's argument is directed against what he calls the observational model of self-knowledge. Such a model allows for the possibility of what Burge calls "brute errors" about one's own psychological states. Burge suggests that this possibility of "brute error" is partially responsible for the dissociation between the "reviewing" and the "reviewed" perspective. Although I agree that without first-person access there is a kind of dissociation, I do not think that the bare possibility of brute errors is the reason for it. The complete explanation for the dissociation, as I will argue, proceeds along different lines.
similarly to how she does those of other people. She treats them like facts whose existence could possibly be independent from her own rational assessments of what her attitudes should be. From a third-person point of view, her best judgments about what her psychological states should be need not settle what they in fact are. This is why critical rational judgments made from a third-person point of view lack an “immediate rationally necessary connection” to the subject’s psychological states.

In fact, because of this it seems to me that Burge overemphasizes the importance of reflection. The necessity for a special kind of first-person access is really more basic. Simply having rational attitudes requires having a capacity for first-person access whether or not we subject them to deliberative evaluation. This is because, as rational subjects, we must take the existence of these states to be exclusively determined by reasons. This is a fundamental to the nature of the first-person perspective. From that perspective, if I believe that P, it is only because some suitable reason (or some set of reasons) is a compelling indication that P is true. Hence, if my way of knowing about my belief were not first-personal, if it rested on epistemic mediation of any sort, then my belief that P might exist independently from any appropriate reasons for holding it. Basing my knowledge of what I believe on evidence therefore amounts to admitting that the existence of my belief does not depend exclusively on reasons that bear on the truth of P. This is because the evidence in favor of my believing that P is not evidence for the truth of what I believe. Thus, lacking first-person access would directly undermine the rational authority of my rational judgments, which is why, in Burge's words, "there would never be an immediate rationally necessary connection" between my sense of what I ought to believe and what I in fact believe. Burge is right to

think that such dissociation would occur without a capacity for first-person access; but it can happen prior to critically reflecting on one’s psychological states.

An example should make this point clearer. Suppose that I come to know that I believe that my neighborhood is unsafe on the basis of behavioral evidence. I recognize, for instance, that I check the locks on my windows and doors far more frequently than in any other place I have ever lived. I also notice that I walk extremely quickly through my neighborhood and regularly glancing over my shoulder. This kind of behavior is good evidence that I believe my neighborhood is unsafe, but it is not good evidence that the neighborhood is actually unsafe. It does not indicate that the proposition <my neighborhood is unsafe> is likely to be true. Indeed, the behavioral evidence for what I believe is consistent with there being no good reasons to think my neighborhood is unsafe. Even if we suppose that my self-ascription on the basis of evidence is correct, it is possible that my belief is based on no good reasons and that its existence is caused by something else (for instance, my growing paranoia). Coming to know what I believe in a way that is mediated by evidence leaves open the possibility that what I do believe is not determined exclusively by what I ought to believe. My point is that when a person relates to her beliefs in a first-personal way, this possibility must be closed off. From that perspective, I believe that my neighborhood is unsafe only because it is supported by good reasons.

Someone might object that just as considering evidence is a rational way to form beliefs about my neighborhood, considering evidence is a rational way to form beliefs about one’s own beliefs. Indeed, it might be thought that if a person is attempting to learn what she already believes, wants or intends, as opposed to what she ought to, she should consider all potential evidence indicating what psychological states she has.27 If one wishes to acquire

---

27 This objection is raised in Shah and Velleman (2005).
knowledge, it might be irresponsible to attend to anything less than the total available evidence.

I do think there are cases where it is epistemically responsible to base one’s self-ascriptions on evidence. For instance, prior to noticing my nervous habits, I was not aware of my belief that my neighborhood is unsafe. Nevertheless, the belief clearly played an important role in my life. It may be that the only way I could ever have learned about it was by noticing behavioral evidence. If so, it seems responsible to acquire knowledge of this belief by responding appropriately to that evidence. But this point does not generalize. If a person’s judgments about what she ought to believe did not normally constitute what she did believe, she would not have rational attitudes. In that case, there would be no point to her trying to engage in any critical reflection or evaluation of her psychological states. Whatever type of cognition she would be reconsidering would not be rational cognition. As I have been arguing, rational cognition requires a capacity for accessing one’s own psychological states in a direct, first-personal way. It is not that a rational agent cannot sometimes acquire knowledge of her psychological states in a third-personal way. But she can only do so by distancing herself or disengaging from the more basic first-person standpoint. If a person completely lacked the capacity for first-person access, if she were self-blind, her psychological states would remain at too far a distance from her own sense of reasons and she would cease to be a rational cognitive agent.

But none of this entails that our psychological states are necessarily connected; it does not entail that *State Dependence* is true.28 According to the view I have briefly sketched, it is not the nature of psychological states that explains why self-blindness is impossible. Rather, this impossibility is a consequence of a special kind of rational relation a person can

---

28 For an alternative explanation see Peacocke (2009). Peacocke also recognizes that *Capacity Dependence* can be explained without denying the *Distinct Existence Thesis.*
stand in to her own psychological states. One is a rational subject only on the condition that she relates to those states in a first-personal way, in a way that enables her judgments to directly constitute her attitudes. The nature of this relation, rather than the nature of psychological states, is what requires a subject to have immediate access to both her cognitive actions and to her psychological states. However, these states can be distinct existences without necessary connections between them. What would necessary is that if a subject relates to a psychological state in a first-personal way she will have a special kind of access to it upon which she can base her beliefs about it.  

IV

*Capacity Dependence* is a modal claim stating that rational subjects necessarily have a capacity for first-person access. One might think that we are hardly justified in saying we have a capacity like this unless we did in fact know about most of our own psychological states. That is, it might seem that if one has a capacity to *phi*, one actually does *phi*. So it seems that any subject of psychological states will be aware of a majority of them and, moreover, this will be necessarily true.

Thus, from *Capacity Dependence*, one can reasonably infer:

FPA: Necessarily, for any subject *a* and psychological state *M*, it is highly probable that if *M*(a), *a* knows that *M*(a).

---

29 For a different view that rejects *State Dependence* see Chapter 5 of Peacocke (1999). Although I agree with much of what Peacocke says, he attempts to explain first-person access by emphasizes the nature of conscious states over the nature of rational subjects.

30 Even this, however, is questionable. See Squires (1969).
Because *Capacity Dependence* implies this, we are guaranteed to know about most, even if not all, of our own psychological states. We need only to actualize our capacity for first-person access, a capacity we naturally have as rational subjects of psychological states.

The inference to FPA may need to be spelled out a bit more carefully. ³¹ First, we might need a suitable *ceteris paribus* clause to rule out inhibitors of the capacity. ³² The elevator in my building would have the capacity to hold 2,500 pounds even if it someone cut the cables every time more than 2,000 pounds was placed inside. ³³ Secondly, when it comes to an agent's capacities (as opposed to an elevator's), the agent in question must try to exercise the capacity in question. But with these qualifications in mind, it does seem that if a person in the right circumstances does not *phi* when they are trying to, then they lack the capacity to *phi*. So if a person who is relates to her psychological states in the ordinary-first personal way has the capacity for first-person access, she will regularly actualize this and therefore know about those states. I think this form of reasoning and FPA are both plausible so I will assume they are true. ³⁴

From either a pragmatic or epistemic perspective there is not much of a difference between constitutive theories and FPA. Either way, it is necessary that most people know about the majority of their psychological states. The only significant difference is whether a person’s psychological states depend on her having beliefs about them. FPA is compatible with the *Distinct Existence Thesis* whereas constitutive theories are not. But does anything philosophically important hinge on this difference? Why insist that the *Distinct Existence Thesis* is true?

---

³¹ Ideally, one would like to make the degree of probability more precise.
³² There is a risk that a *ceteris paribus* clause will make the conditional trivial. This is a familiar point from Goodman. However, since I do not intend FPA to be an analysis of any concept, I will set the concern aside.
³³ Discussions of “masking” can be found in Johnston (1992) and Lewis (1997).
³⁴ I actually think that FPA could reasonably be considered a datum in discussions of self-knowledge.
Early in this essay, I mentioned that our commonsense intuitions favor the Distinct Existence Thesis. I think that this is enough reason to believe it. We should not posit necessary connections between things that seem to be causally discreet. But even if we agree with Hume that psychological states “may exist separately, and have no need of any thing else to support their existence," this is not a conclusive argument against constitutive views. Moreover, there may be countervailing theoretical reasons to accept a constitutive theory. For example, some functionalist theories of mind individuate psychological states in terms of their causal relations to each other. Thus, for something to be a desire it would need to be related to other states, including, plausibly, beliefs about it. To the extent that one is drawn to a functional analysis, one may also be drawn to a constitutive view rather than the Distinct Existence Thesis.

However, constitutive theories face some difficulties. One of the more serious is that we are often ignorant or mistaken about our own psychological states. The existence of errors in this domain strongly suggests that we should be fallibilists about knowledge of our own psychological states. We should accept that any justification we have for beliefs about them is defeasible. The trouble for constitutive theories is that necessary constitutive connections between psychological states do not allow for errors. If the existence of a person’s psychological state constituted her knowledge of it in some way, it is hard to see how she could ever be mistaken. A universal constitutive theory is therefore incompatible with fallibilism. Since nearly everyone rejects infallibilism, an advocate of a constitutive theory must allow for exceptions to the necessary relations she thinks normally hold. She will typically do so by citing additional conditions on those relations.

---

35 A different objection than the one discussed here has received a fair amount of attention. It is that constitutive theories make self-knowledge neither a cognitive nor an epistemic “achievement”. For discussion see Boghossian (1989), Fricker (1998) and Peacocke (1999). At least some constitutive theorists seem happy to accept this consequence (cf. Bilgrami (2006)).
For example, Shoemaker believes that if one has a psychological state then one will “automatically” have a “corresponding second-order belief”. However, he also thinks this requires "a certain degree of rationality, intelligence, and conceptual capacity (here including having the concept of belief and the concept of oneself)." It is therefore possible that a person’s psychological states to not be constitutively connected and for her to lack a corresponding second-order belief, but only if she also lacks "a certain degree of rationality, intelligence, and conceptual capacity". These features must be in place in order for her first- and second-order states to be constitutively related.

It is plausible that one must have a degree of conceptual skill and intelligence in order to self-ascribe a psychological state. Let us therefore assume that without these there is no constitutive connection between a subject’s first-order states and her beliefs about them. However, it seems rather implausible that every time a person makes an error about one her psychological states it is due to a deficiency in her intellect or her conceptual abilities. What about rationality? Could it be that a person commits errors only when she is not being rational?

There are two directions in which one could try to explain how rationality is necessary for constitutive connections between psychological states. First, according to a popular version of a Freudian idea, we lack awareness of our unconscious psychological states, which are “irrational”. It is because they are irrational that we push them outside of

36 (1994), pg. 243
37 Shoemaker offers two proposals for understanding the dependence between a person’s psychological state and her belief about it. The second is that the two states have different “core realizations” but overlapping total realizations. One might prefer this if she wished to maintain that some causal relation connected a person’s psychological states with her beliefs about them. It might be thought that if we allow causal relations at the level of core realizations, then we can account for fallibility. The cases in which we make mistakes are just the ones where that causal relation breaks down. The problem, however, is that since the causal relation does break down, the states are not always connected. If we insist that their “total realizations” are nevertheless necessarily connected in some way, it can only be under certain conditions. To account for our fallibility, as I shall argue, one must explain why those conditions matter. Cf. Campbell (2011).
our awareness. Someone could argue that unconscious states are of a fundamentally different kind and, therefore, not constitutively connected to any second-order beliefs. Moreover, it is natural to think that bringing unconscious states to awareness requires one to first make them less irrational. This is what Bilgrami thinks happens at a “second stage” of therapy. He suggests that gaining awareness of unconscious states means that a person can “consider first whether the discovered dispositional or motivational mental state is in conflict with her commitments, then, if it is, to see whether she wishes to bring the conflicted disposition in alignment with her commitments by changing the commitment or by changing the disposition.”

This process sounds quite similar to the sort of reflective criticism and evaluation we considered earlier. If this were the result of “cognitive inquiry” into our unconscious states, then it seems one could, in this way, transform irrational states into fully rational ones, or at least abandon them. Yet even if this were to happen on occasion can it explain errors outside of a psychoanalytic context? While focused on writing this paper, I may be unaware of my desire to listen to country music. But must I be so ashamed of it that I repress it? The constitutive theorist’s main suggestion is that irrationality explains ignorance; it is why my desire is not necessarily tied to any beliefs about it. But not all of the psychological states we misjudge are irrational in some quasi-pathological or neurotic sense.

A better way to support a constitutive view is to claim not that my desire to listen to country is irrational in some problematic manner but that it is not even within the realm of reason—that my desire to listen to country music is a non-rational state, something more like an urge or an impulse to listen to country, something that could never be based on reasons. In that case, my lack of knowledge would be explained if primitive states of this kind were

---

38 (2006), pg. 318.
ontologically independent from beliefs about them. It seems that one could think that constitutive relations hold only for more rational psychological states.

For the sake of argument, let us stipulate that non-rational states are not constitutively related to beliefs about them, which would allow us to explain simple ignorance. But could we explain errors the same way? It seems to me that we all make mistakes about, for instance, what we want for dinner tonight or what we believe about a difficult philosophical passage. It might seem that we can sufficiently understand these by explaining both a person’s ignorance and her incorrect self-ascription. It could then be argued that if the constitutive view can explain the first it need not address the second.

There is, however, a problem with this line of thinking. Conscious self-ascription occupies a person’s attention and it is often the result of her self-conscious reflecting on her psychological states. For example, you might ask me what sort of music I would like to hear with dinner. This invites me to reflect upon relevant desires. After doing so, I might say that I would like to hear country but suppose I am wrong and that I really desire to listen to the blues. Given the level of conscious engagement with your question it seems unlikely that there is nothing in the scope of my conscious attention. Even if reflection did not reveal my actual desire to listen to the blues, it seems that it must have uncovered some kind of motivational state. Otherwise, the sensible thing to report would be my lack of preference.

---

39 Consider, however, how we would make sense of losing awareness of a particular psychological state. Suppose that I am fully aware of my desire to listen to country music, which in my mind is based on very good reasons, but suddenly become distracted by something in my vicinity. Should my desire persist, it does not seem that it would be less rational simply because I was not presently aware of it.


41 It seems even less plausible in cases where I am asked to reflect on what I believe. When someone asks me to think about what I believe about a difficult philosophical passage, I may give her the wrong answer. It does not seem that in this case the belief I actually have is of a more primitive non-rational kind.

42 In a case different than the one I am imagining, a person can self-ascribe an urge or impulse. For instance, I might say ‘I have a hankering for opera’ or even ‘I have an urge to listen to opera’. Do urges self-ascribed in this manner remain outside the domain of rational self-awareness because they are “non-rational”? 
Because self-conscious reflection on one’s psychological states can result in errors, it seems that at least some mistakes involve more than a lack of awareness.43

Consider the following example. During dinner, my spouse may gently nudge me and say she can tell that I’m not enjoying the country music. She knows me pretty well; so she says “Wouldn’t you prefer the blues?” Alternatively, after you begin playing the country music, I may notice that I am not enjoying it as much I thought I would. A natural thought would be that I was wrong about what I desired, not wrong that I desired something. In this case, it seems that I know of the existence of a particular state but get its valence wrong. I personally think this sort of thing happens often, but my argument against the constitutive theory requires only that it happen often enough to be worthy of explanation.44

An advocate of the constitutive view may wish to deny that a person could make a mistake in the way I am suggesting, by misidentify her psychological states.45 She might object that misidentification only seems conceivable on the assumption that a person’s beliefs about her psychological states rest on causal relations, which is precisely what the constitutive view denies. But there is no independent reason to think that a person can be aware of the existence of her psychological states but wrong about its valence. Even if this objection is correct, even if a psychological state cannot be misidentified, a successful constitutive view must still offer a more substantial explanation of error. This is because,

44 One might be tempted to say the error consists simply in the fact that I assert or believe something false. Since I do not possess a desire to listen to country music, my saying that I do is a kind of mistake. Notice, however, that this changes the nature of the error. It may be true that violating a norm of assertion or of belief is, in some sense, a mistake, but it is not a mistake a person makes about her psychological state. In fact, the explanation for why I wrongly assert or believe something false usually rests on my misjudging an actual psychological state.
45 Bar-On (2004) and (2009) denies that it is impossible for a person to misjudge or misidentify a psychological state. In her estimation, first-personal judgments are immune to these sorts of errors. I think this is wrong. It seems to me possible for a person to wonder what it is they want to listen to during dinner while acknowledging that they want to listen to something. Presumably, Bar-On would argue that in these cases one’s desire is indeterminate but I do not see why that must always be the case.
according to the constitutive theory, second-order beliefs exist primarily because they are necessarily related to the first-order states they are about. However, in cases of error, there is no necessary connection; so they must exist for other reasons.

The most plausible explanation will appeal to a transition from one state to another, a minimally causal transition that could on occasion fail.\textsuperscript{46} A causal failure might derail a reliable tracking mechanism or, due to a lack of attention, we may fail to grasp the proper justification or evidence for what we should think about our psychological states. If there were normally some transition between psychological states and our beliefs about them, mistakes could be explained on the basis of failures in the causal processes that supported this transition. I think a causal explanation along these lines is very difficult to provide if we accept a constitutive view. If there is no causal relation between my psychological states and my beliefs about them, it does not seem that a contingent transition could fail. But then how do errors occur?\textsuperscript{47} It is possible for a proponent of a constitutive view to propose that, although second-order beliefs are usually constituted by first-order states, false ones are causally generated by non-rational states. Whether or not this proposal can succeed will, I think, ultimately be determined empirically. But \textit{prima facie} it does not seem promising. One issue is that, since non-rational states do not always cause false second-order beliefs, since they often exist without causing any self-ascriptions, one must clarify the circumstances in

\textsuperscript{46} cf. Boghossian (1989): "I know of no convincing alternative to the following style of explanation: the difference between getting it right and failing to do so (either through ignorance or through error) is the difference between being in an epistemically favorable position with respect to the subject matter in question--being in a position to garner the relevant evidence--and not." (pg. 76) Boghossian's discussion seems to indicate that a problem for views like Shoemaker's is that they have no explanation of error, but he mistakenly thinks the only alternative explanations are either evidential or inferential. The view I have proposed in this paper is neither.

\textsuperscript{47} Zimmerman (2006) argues that a constitutive theorist must embrace disjunctivism about first-person access. The proposal that our ignorance or error involves a more primitive non-rational state would be a sort of disjunctivism. However, Zimmerman also claims that disjunctivism about psychological states is far less plausible than a similar view about perception on the grounds that no qualitatively indiscernible states that could reasonably be mistaken for rational psychological states.
which they do causally generate errors. More qualifications will therefore be needed to
support the constitutive approach. The problem is not that constitutive views make errors
impossible but they do needlessly complicate their intelligibility.

At first glance, the constitutive approach can seem attractive because we are not
usually wrong about what we think or feel. Most of the time we get the existence and the
character of our psychological states just right and we usually do so in a special way. But an
equally important truth is that a person could be wrong about the existence or character of
any one of her psychological states. I have tried to illustrate how we can accept both of these
important truths about what it is like to for us to be psychological subjects. An important
lesson to take from this essay is that the Distinct Existence Thesis is consistent with an account
of first-person access that implies FPA. In that case, because first-person access would
grounded in the nature of rational cognitive agents, one could accept the Distinct Existence
Thesis while nevertheless acknowledging that first-person access is completely unique from
an epistemic perspective. This would not be a constitutive theory but I think it would have
all that makes those sorts of views appealing.

References


Bilgrami, Akeel (2010).


Boghossian, Paul (1989). “Content and Self-Knowledge”. In Peter Ludlow and Norah

Martin (eds.), Externalism and Self-Knowledge.


-(2009). “Two Kinds of Agency”


