**Self-Knowledge**

It feels appropriate to start this entry with an autobiographical aside. I arrived at the topic of self-knowledge primarily through the philosophy of mind and it took a while to see myself as an epistemologist. Nevertheless, I plainly *was* doing epistemology, which I eventually realised after friends rightly pointed out that my topic featured ‘knowledge’ in the title.

That I am an epistemologist is one piece of self-knowledge which took time and help to acquire. And it can be harder still to discover other things about oneself, such as character traits or repressed beliefs. On the other hand, it’s common to start discussions about self-knowledge by observing that some self-knowledge, in particular, that of most attitudes and sensory experiences, seems importantly special compared to our knowledge of other people’s mental states. Suppose I believe that British food is underrated, that I want a cup of tea, or that I see a computer. It seems remarkably easy for me to learn these facts about myself, and I need not rely on any evidence to do so.

Debates in self-knowledge span various topics and disciplines, such that it’s understandable that my own self-knowledge of being an epistemologist would take time to acquire. Also, instead of marked factions with clear sides there are many lines of dis/agreement and which are motivated by different interests. Even an extended article like this can therefore only hope to capture part of the terrain. In the following, I’ll focus on self-knowledge of beliefs and our reasons for belief. For discussion of other cases of self-knowledge, see Gertler (2011, 2021).

I’ll proceed as follows. §1 introduces the starting intuition that self-knowledge is distinctive relative to other-knowledge, i.e., knowledge of others’ minds. I’ll then spend most of the article tracing certain discussions in the self-knowledge of belief in §2, and discuss the less explored topic of self-knowledge of reasons for belief in §3.

# 1 Self-knowledge as importantly distinctive

We can potentially distinguish self-knowledge from other-knowledge along four main lines, although it’s debatable which are most relevant, if indeed any. These are especially pertinent for the set of examples that count as ‘easy and special’ self-knowledge.

*1. Improved epistemic access*

First, most philosophers accept that our self-ascriptions of mental states are more reliably true than other people’s. If I say that I believe that British food is underrated, regardless of my accuracy about culinary matters, it’s surely most likely that do in fact have this belief. At the claim’s strongest, though, one might say that certain self-ascriptions will always be true if formed in the right way, i.e., they’ll be infallible.[[1]](#footnote-1) This seems less plausible regarding our beliefs as opposed to experiential states.

Second, we might say that certain features self-intimate, i.e., that there’s a necessary connection between being in the relevant mental state and knowing that one is. The strongest claim would be to say that I always know that I have the state when I in fact do.[[2]](#footnote-2) More plausibly, it might be that I will always *be in a position* to know that I have it even if I do not currently possess this knowledge.[[3]](#footnote-3) Or, weaker (and more plausibly) still, perhaps I will be in a position to know that I have a certain attitude *if I am rational/if the relevant attitude is rational*.[[4]](#footnote-4)

*2. Distinctive method and warrant*

Most philosophers think that a distinctive method and warrant underpins at least some instances of self-knowledge. Standardly, we might contrast self- and other-knowledge in saying that self-knowledge is non-inferential. I don’t seem to learn of my belief about British food by observing my behaviour (such as my tendency to hoard crumpets), listening to my speech (including my rants on the topic) and/or inferring from these facts like an external observer must. That said, it may well be that we can observe certain mental states in other people, such as seeing their goals when they reach for something. [[5]](#footnote-5) But I don’t need to observe my physical behaviour to learn what I’m thinking.

*3. First-person authority*

The standard thought is that we’re authoritative about our mental states in that others normally defer to us about them. There’s two ways of cashing out first-person authority, which relate to the two senses of ‘authority’ in English.[[6]](#footnote-6) First, a module leader is an authority concerning their class in that they know the readings and deadlines, and are more likely to be correct if you ask them for information. But they’re also an authority in the sense that they have a say over the class: they assignthe readings and are held responsible if students run amok. In the case of self-knowledge, I might be authoritative about my beliefs in that my self-ascriptions are especially reliable or also in the sense that I am in charge of my own mental life. (Although on the difference between epistemic agency and the teacher’s control over their class, see Hieronymi 2006).

*4. Rational agency*

Another difference could be that I don’t relate to my beliefs in the way that I relate to other items of knowledge – as just any old facts to be discovered. Instead, the claim is that we’re agents regarding our attitudes and are responsible for them. In this way, our attitudes differ from other mental states such as pains and perceptual experiences which merely happen to us. A notable tradition argues that this rational agency grounds self-knowledge of these attitudes.

Suggestions (1)–(4) are highly debated. The positions I will be turning to discuss now have different things to say about these suggestions, and characterise (or deflate) the asymmetry in various ways.

# 2 Self-knowledge of belief

I’ll now concentrate on the main topic of self-knowledge of belief, and tell a story focusing on two of the main turns in contemporary debates. I’ll end with some broader points.

## 2.1 Quasi-perception/observation

The obvious place to start is an inward-looking and empiricist approach to self-knowledge. The thought is that just as I perceive the world, so I can perceive my inner life.

The most minimal and plausible way of understanding this, in the context of belief, says that self-knowledge resembles perceptual knowledge in two key respects.[[7]](#footnote-7) Following the ‘broad perceptual model’, introspection parallels perception *broadly construed*, or rather, it’s to say that the fundamental characteristics of perception are very minimal to begin with (here I use Shoemaker’s (1994b) terminology). What is central to learning perceptually that there’s a dog in front of me is that the dog is independent of my knowledge of him and that he, in a roundabout way, causes my belief about him. The role of mediating visual experiences, for example, is not essential.

Armstrong (2001) develops a paradigmatic version of this view according to which self-knowledge uses a ‘self-scanning process’ (2001: 324). Armstrong also thinks that self-knowledge would be warranted along the same lines as perceptual knowledge – where the latter is reliabilist (e.g., *Ibid*.237-8). He holds that one’s belief that *p* can be non-inferentially warranted in virtue of its being sensitive to *p*. This sensitivity obtains when as a matter of ‘empirical necess[ity]’ (*Ibid.* 189) one would not have the belief that *p* had *p* not been the case. If the self-ascriptions formed by the self-scanning process are reliably accurate, as Armstrong thinks they are, then our self-ascriptions will be thusly warranted. Because the account rests on a causal mechanism, self-knowledge resembles visual perception in the potential for mistakes in that the causal mechanism might fail and could even be replicated to give us access to others’ mental states.

Here one might worry, what \*is\* it that we are aware of in this case? After all, our beliefs aren’t objects we can become aware of, and Armstrong in particular thinks that beliefs are individuated by their causal profile. In reply, Armstrong suggests that both introspection and perception just amounts to a ‘flow of information’ which can include the fact that you have a belief. He also observes a precedent for thinking that we have perceptual awareness of causal relations, in bodily awareness. I can be aware of something pressing into me without inferring it, even though this is a causal relation and pressure is a causal notion (2001: 96-7, 326-7).[[8]](#footnote-8)

## 2.2 Neo-Ryleanism

Some argue that it isn’t just theoretically possible that self-knowledge could resemble other-knowledge. It does in fact do so right now, and we acquire both through inference. This includes the cases of self-knowledge of belief that outwardly seem special and easy. Ryle (1949) famously argued that our beliefs consist in behavioural dispositions, and we learn of them via inference from behaviour and inner speech (note how his accounts of belief and self-knowledge dovetail together). More plausible versions seek to better capture the apparent asymmetry between self- and other-knowledge by expanding the evidence base to include more mental states. Cassam (2014) and Carruthers (2013, 2010) have views about self-knowledge which can be described as *neo-Rylean* in this way (which isn’t to say that they’re behaviourists about belief).[[9]](#footnote-9) The evidence includes ‘judgement, inner speech, dreams, passing thoughts and feelings’ (Cassam 2014: 138), and sensory data in general (Carruthers 2013, 2010).

We could say that self-knowledge is inferential at the personal and/or subpersonal level, which is an epistemologically significant decision. Talking of the personal level is to say something about the subject *herself*, e.g., that she does such-and-such a thing. Or instead, we can just talk about parts of the person at a low-level; e.g., it isn’t that the subject *herself* beats her heart or fires her c-fibres.

Targeting one’s account at the personal-level, as Cassam (2014) does, allows one to say that self-knowledge is inferential in virtue of both the method and warrant involved. On this account, subjects themselves form the self-ascriptions either consciously or unconsciously by transitioning from evidence (*Ibid.* 138-9). But even if the process itself hasn’t taken place, self-knowledge can still be nevertheless warranted inferentially (e.g., *Ibid.* 139). This is to say that the self-ascription is based on supporting beliefs which are themselves justified.

Alternatively, we might just say that self-knowledge is inferential at the subpersonal level, as with Carruthers’ account (2011). He thinks that both self- and other-knowledge are acquired by a ‘mindreading’ module (e.g., 2011: 260), and therefore a distinct mechanism with its ‘own neural realisation’ (2011: 227). It transitions from sensory evidence to attribute mental states to both ourselves and other people. This mechanism is therefore just one part of the person, and so the subject herself doesn’t infer (although Carruthers says otherwise, I therefore think it’s natural to read this as a subpersonal account). Unless we add more concerning something the subject herself does or has access too, the most natural account of the epistemology would therefore be reliabilist.

These inferentialist accounts obviously challenge quasi-perception in saying that we habitually acquire self- and other-knowledge in the same way. But they all concur in two key respects, concerning the direction of attention and role of empirical warrant; these constitute important fault lines in the debate.

## 2.3 The turn to transparency

The accounts so far agree that we learn of our own mental states by attending to own minds (or at least, a mechanism does) in the sense of having an inner detection mechanism directed towards them or relying on evidence regarding the mental states themselves.

But in opposition, Evans famously writes that instead:

[I]n 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?’ (Evans 1982: 225).

The thought is that, normally, when trying to learn whether we believe that *p*, we consider whether *p* itself is true. That is, the question of whether *I believe that p* is *transparent* to the world-directed question of whether *p*; we use the so-called *transparency method* to learn of our beliefs. This also potentially applies to many other states, and proponents often extend it in various ways. E.g., perhaps whether *I intend to phi* is transparent to the question of whether *I will phi*, whether *I desire o* is transparent to the question of whether *o is desirable*, and whether *I’m experiencing pink* is transparent to whether *there is a pink object in front of me*.

Suppose the question of what I believe is transparent to the question of whether *p*, and by answering one question I can thereby – presto! – answer the other. How this ‘thereby’ works is the million-dollar question. In what way can the transparency method confer knowledge? After all, it involves transitioning from one subject matter to an unrelated one, and can result in knowledge about whether I believe that *p* even if I’m mistaken about *p* itself.

## 2.4 Empiricist transparency accounts

One tactic is to incorporate the transparency method into an empiricist epistemology.

Byrne (2018, 2011, 2005) argues that the transparency method is best understood in terms of following the inferential rule, BEL: If *p*, believe that you believe that *p* (2005: 95). So, if *British food is underrated,* then you should believe that you believe that *British food is underrated.* Using BEL is knowledge-conferring under a reliablist epistemology even though it seems like a terrible rule. Employing BEL is self-verifying since one can only do so ‘by believing that one believes that *p* when one has recognised that *p*. And recognising that *p* is (inter alia) coming to believe that *p*’ (2005: 96). And the self-ascriptions are safe – it couldn’t easily have been false if you self-ascribe the belief in this way, if judging that *p* just is coming to believe it. Even if *p* is false, it’s true that you believe that *p*. It’s interesting to note that this relies on a strong commitment regarding the nature of belief, namely, that there is a strong link between judging *p* to be true and believing that *p* is true.[[10]](#footnote-10)

## 2.5 The turn to rational agency

All these accounts so far have been *empiricist*. This is to say that self-knowledge is warranted in a way that resembles our knowledge of the natural world. More than anything, saying that self-knowledge is empiricist is to say what this kind of account is not: it’s not grounded in rationality or rational agency, in contrast to the next set of accounts we’ll consider. The thought behind these accounts is that unlike our pains and perceptual experiences which happen upon us, our beliefs are up to us in a certain sense; we can exercise *agency* concerning our beliefs, and perhaps also other attitudes.

A main motivation is what we can call the argument from responsibility. Moran writes that: ‘Without some assumption of agency, there would be no point to the interpersonal demand for justification or correction, for there would be nothing the person could be expected to do in *response* to this demand’ (Moran: 403-4). We can cash this out more precisely as follows.[[11]](#footnote-11)

Premise 1: Subjects are accountable for their beliefs.

Premise 2: Subjects can only be accountable for states of affairs they have agency over.

Conclusion: We have agency over our beliefs.

We should note two important clarifications. First, the controversy concerns whether subjects have *direct* agency over their beliefs, and not just performing intentional actions that bring about the belief like visiting a hypnotist.[[12]](#footnote-12) Second, the claim isn’t that subjects can believe at will, choose to believe, believe intentionally, or believe for pragmatic reasons.[[13]](#footnote-13) Some philosophers do endorse claims of these sorts, and an interesting project would be to explore agentialism about self-knowledge from this starting point. Nevertheless, the main discussions of self-knowledge don’t understand epistemic agency in these terms.

## 2.6 Rational agency and transparency

Accounts that ground self-knowledge in epistemic agency can either attempt to accommodate transparency or reject it.

One way of incorporating transparency into an agentialist framework famously comes from Moran (2001). He argues that according to the quasi-perceptual account we would treat our attitudes as just a ‘passing show’ like how we observe sushi on a conveyor belt (2001: 32). This would be to approach our beliefs as some among many empirical facts in the world, i.e., to adopt the so-called ‘theoretical stance’ (2001: 65). Instead for him, the perspective we adopt when self-ascribing the belief needs to be the same as the perspective we take when holding the belief itself.

This point is reflected in the irrationality of so-called ‘Moore’s paradoxical statements’: ‘*p*, and I don’t believe that *p*’ and ‘I believe that *p*, but *p* is false’. Such statements seem problematic even though the third-person versions are fine. Yet empiricist accounts fail to capture this since whether *p* is true isn’t in itself to say anything about whether I believe that *p*. This means that under the theoretical stance, ‘the thought expressed in a Moore-type sentence would describe a perfectly coherent empirical possibility on which one could sensibly report’ (2001: 84).

Moran therefore argues that we ideally use the transparency method to learn of our beliefs, whereby the truth of *p* *does* matter from our perspective. Specifically, transparency rests in ‘the deferral of the theoretical question “What do I believe?” to the deliberative question “What am I to believe?” And in the case of the attitude of belief, answering a deliberative question is a matter of determining what is true’ (2001: 63). So, in considering whether *p*, I thereby make up my mind on the matter and am able to self-ascribe my conclusion as my belief. Understanding transparency in this way engenders a particular understanding of first-person authority. Namely, I am authoritative regarding my mind because others can’t learn about it if I haven’t made it up. This is like how the students might try to predict the course reading list, but there’s a limit to how successful this will be if the convenor still hasn’t decided.

Importantly, according to Moran, it’s not just that we can use the transparency method; we ought to do so (2001: e.g., 84, 127). Otherwise, there is ‘something wrong with [us]’ and we will be in a state of ‘alienation’ concerning the belief (*Ibid.* 68).[[14]](#footnote-14) We fall short if we either cannot do so because our belief fails to match with what we would sincerely assert, or nevertheless refuse to use it. This is roughly because otherwise we would not see ourselves as rational agents. We would thereby fail to count as such, and wouldn’t be able to deliberate at all (*Ibid.* 84).

Also, a nice feature of this account is that it captures how self-knowledge can sometimes be hard to acquire, but not because the state is alienated and we need to infer it. Suppose I ask you if you believe that it’s raining outside – you can tell me either way following a quick glance. But suppose I ask you for your views on the philosophy of self-knowledge. If you’ve only just started learning about the topic, you might ask me to wait until you’ve finished reading this article and thought about it. This is seemingly because you need to make up your mind about the topic and the world-directed question is especially thorny.

Suppose that we employ the transparency method by making up our minds in the process; the question remains as to how this provides us with self-*knowledge.* Here, Moran appeals to transcendental entitlements grounded in the nature of deliberation (2003: 406-7). Transcendental entitlements are entitlements secured by a transcendental argument. Arguments of this form show that x must be the case because x is required for some fact to obtain that would be hard to reject. Moran argues that deliberation requires several things. First, we must think that the conclusion we reach will change our beliefs. This is because deliberating is to decide what to believe, and, further, we need to recognise that this is what we’re doing. It’s less clear why this is the case. One thought is perhaps that this is just what deliberation is – we in fact always make that assumption. Also, we conclude deliberation by in effect committing and binding ourselves to a position (2001: 95). Plausibly, doing so requires recognising oneself as so bound. Deliberation is also such that we can only *assume* that our conclusion will change our belief. The fact that we are rational agents is irrelevant for deliberation since it does not help us decide what to believe (2001: 94-5). Therefore, in order to deliberate, we must tacitly assume that when we conclude that *p*, we indeed believe that *p*; this assumption entitles us to self-ascribe the belief.

Another transparency account with an agentialist slant is that of Coliva (2009). She argues that to judge that *p* is to take on the commitments associated with believing that *p* such that you make it the case that you hold the belief and are responsible for it. Given the right conceptual grasp, the subject will also express this commitment by self-ascribing the belief. And this runs both ways: self-ascribing the belief also constitutes this commitment and is a way of making it the case that you hold the belief.

## 2.7 Rational agency and self-knowledge: non-transparency accounts

Burge, on the other hand, grounds self-knowledge in reflective critical reasoning as opposed to world-directed deliberation. This is the process of reflective thought whereby you evaluate your reasoning and the bases of your beliefs. Burge writes: ‘As a critical reasoner, one not only reasons. One recognizes reasons as reasons. One evaluates, checks, weighs, criticizes, supplements one's reasons and reasoning’ (1996: 98). E.g., suppose you believe that it will rain, and but I tell you that the weather forecast is for sun. You reconsider your reasons: there are grey clouds outside, but, upon further thought, clouds dissipate quickly. Since the grey clouds are a poor reason for the belief, you update it accordingly.

For Burge, our nature as critical reasoners grounds self-knowledge of our beliefs. This has three components. First, we have an entitlement to self-ascribe our beliefs. We ought to engage in critical reasoning and to revise our beliefs in line with assessments of what is rational. Critical reasoning requires self-ascribing our attitudes and reasons. Moreover, since we can only be obligated to do things that we actually can do, i.e., ought implies can, it must be rational to self-ascribe our attitudes in this context, i.e., we must be entitled to do so (1996: 101-2). Second, the self-ascriptions in fact must also be knowledgeable, since false accidentally true beliefs wouldn’t help the process or make it rational.

Finally, the warrant needs to be non-observational and non-inferential; it can’t just rest on reliable causal relations. This is because for critical reasoning to function as it ought to do, there need to be direct rational relations between items of self-knowledge and the lower-order attitudes they concern. In general, our attitudes ought to be reasonable by our lights and responsive to our take on the reasons. But I don’t just have immediate reason to revise my belief that *p* if I discover evidence that not-*p* – I also have it if I learn that *my belief is unjustified.* But suppose I learn than my friend believes that *p* but that her belief is unjustified. There would be no direct normative force. She would only have reason to change her mind if she trusts my opinion. And the most I would have reason to do is to try and get her to change her mind and only then if I care about her believing the truth. If inner sense theory were completely correct, when I evaluate my beliefs, all I would have is reason to try and get myself to change my mind too (108-11). So, we must be entitled to self-ascribe our attitudes in a way that doesn’t rest on empirical processes of this sort.

## 2.8 Rationality and self-knowledge

One can also ground self-knowledge in subjects’ rationality more broadly, without appealing to doxastic agency and responsibility. Shoemaker (199b, 2012) for example thinks that believing that *p* constitutively involves believing that you believe that *p*, provided you possess the relevant conceptual and rational capacities. He argues that self-blindness is impossible in rational agents, i.e., those ‘with normal intelligence, rationality, and conceptual capacity’ (1994b: 282). This would be a systematic unawareness of oneself resulting from brute errors, and which doesn’t reflect a failure in rationality (as with visual blindness). But rational agents know that if *p* is true, then it is instrumentally useful to act as if *p* is true in all ways, which includes asserting that they believe that *p*. So we don’t need to posit any extra mechanism to subjects in order to explain self-ascriptions – just ‘what is needed in order to give them first order beliefs plus normal intelligence, rationality, and conceptual capacity’ (1994b: 284).[[15]](#footnote-15)

## 2.9 Criticisms of rational agency accounts

I’ll now discuss some criticisms which target either agentialist accounts or rationalism more broadly.

Perhaps empiricists can agree that being a rational agent necessarily requires being able to know your beliefs, the capacity for direct agency, assuming that one is a rational agent etc… but, Gertler (2018, 2011) writes, under a *de dicto* reading of these necessities. Suppose we say that if Sally is honest, then necessarily she doesn’t lie. This states a necessary fact about Sally but only insofar as she’s honest. There’s many possible worlds in which she lies – she just wouldn’t count as honest in them. That is, the necessity claim only applies to honest beings *de dicto* (insofar as we say that they’re honest) and not *de re* (concerning whether they themselves could lack that feature). Similarly, we could say that necessarily, rational agents know their beliefs *insofar as we would count them as rational agents*. But for any such agent, they could fail to possess self-knowledge, in which case they would cease to count as a rational agent. But a merely *de dicto* necessity couldn’t ground self-knowledge in the way that agentialists and rationalists need.

We might also have doubts about transcendental entitlements. In general, they feel like a magic trick. With a transcendental argument, it’s not that the fact that we need self-awareness to deliberate is a reason for thinking that we have self-knowledge but there’s still an open question about what grounds it. Rather, this requirement itself is meant to ground the knowledge. In this sense, it feels like conjuring up the warrant out of nowhere because you need it in a way that is sure to raise suspicions.

Campbell and Greenberg (2023) argue against there being sui generis doxastic agency (i.e., one that doesn’t reduce to common-garden intentional agency) and also attempts to ground self-knowledge in this way. They claim that appealing to *agency* as understood like this can only ever be redundant. One ends up understanding agency as resting in, say, our commitments to the truth of a proposition or responding to reasons as part of critical reasoning. But then one can just as well appeal to these notions, and think that they require self-knowledge, without thinking that they amount to *agency.*

In the face of the commitments regarding sui generis doxastic agency, one could also appeal to the agency involved in intentional *physical* actions, like the original ‘agentialist’ position, that of McGeer (1996). She takes as a starting point a dispositionalist account of belief akin to Ryle’s. But instead of learning of our beliefs by inferring them from the constitutive behaviour, our self-ascriptions are often correct because we make it the case that we act in accordance with them, and thus count as holding that belief.

## 2.10 Prior awareness accounts

Apart from Cassam’s personal-level inferentialism, the accounts discussed so far don’t appeal to anything that the subjects themselves grasp per se (but instead to reliable transitions and transcendental assumptions). In contrast, other philosophers appeal to some sort of prior awareness regarding our beliefs, although there is a challenge as to how to do this without positing a spectatorial relation to our minds.

One option is to appeal to action awareness and to utilise work in the philosophy of action in a very literal way. We seem not just to have distinctive access to our mental states, but also to what we are doing. For some philosophers, this is because we have a special form of experiential awareness that represents what we are doing, which is distinct from our sensory and visual awareness. This includes mental actions such as deliberating, calculating, imagining, and making judgements. Just as I do something in saying aloud ‘p’, similarly I do something when I sincerely assert it in my own head. Both Christopher Peacocke (2009, 2007, 1998) and O’Brien (2007, 2005) think that we can have a primitive form of experiental awareness of what we are doing in judging that *p*. This isn’t just to resort to a quasi-perceptual account because, Peacocke presses, an experience can occupy our attention without being its object (1998: 65-6).

There’s various strategies for using this observation to then say something about self-knowledge of belief. Christopher Peacocke argues that the concept of belief includes a close, although fallible, connection between them and our judgements. Our awareness of judging *p* can serve as a reason for self-ascribing belief in and of itself in virtue of our conceptual grasp (1998). Antonia Peacocke (2017) holds that when we employ the transparency method and answer the world-directed question, we perform this mental action under one particular description, namely, finding out whether we believe that *p*. So, in having agents’ awareness of deliberating and forming a judgement about the world, we are aware of our conclusion as figuring out whether we believe that *p*. And Keeling (2024b) argues that we have what she calls ‘doxastic agent’s awareness’, which is a direct awareness of our beliefs in virtue of forming and sustaining them through performing the relevant mental actions.

Alternatively, in developing a transparency account, Boyle (2019) draws on Sartre’s (1956) notion of non-positional consciousness. This is a form of consciousness that doesn’t have an object, i.e., it isn’t consciousness *of* something. Boyle’s gloss is that when we represent the world via various mental states (believing something, desiring something, and so on), we do so under a ‘mode of presentation’ that already presupposes first-personal elements. In the case of belief, judging is to express the belief that *p*. It is thus to represent something in the world as being true, as opposed to just supposing that it is, and to therefore take a certain issue as settled. But when we take a question to be closed, it’s not that the world has suddenly become fixed – we have become fixed on the matter. So to treat the question as closed is already to treat it as closed *for us.* It is to have a tacit awareness that we have taken a stand on the issue, which warrants us in self-ascribing our belief (2019: 1034).

## 2.11 Implications

I’ll conclude this section on the self-knowledge of belief (and the main meat of this entry) by noting some broader epistemological upshots.

First, most obviously, our accounts of the nature of belief and self-knowledge of them intertwine. The most immediate way in which self-knowledge of belief bears on belief itself is under a constitutivist account such as Shoemaker (2009), whereby knowing that one believes that *p* is in fact part of what it is to believe that *p*.[[16]](#footnote-16)

The relation between believing and judging that *p*, and the potential for disassociations, is also relevant. Can you judge that *p* without believing that *p*, and can you believe that *p* without being disposed to judge that *p*? It strikes me that most philosophers in the self-knowledge literature think that that you can count as believing that *p* without being prepared to judge that *p*. So alienated beliefs still count as beliefs, just irrational ones.

There seems to be more divergence on whether you can judge that *p* without thereby believing that *p*. Some authors think that you can’t – recall, for example, the role this plays in Byrne’s account and his claim that following BEL is strongly self-verifying because judging that *p* is coming to believe that *p*. Others think that it isn’t. For example, Moran doesn’t, but that failing to learn of a belief through the transparency method counts as a rational failure. This decision has consequences for the reliability of the transparency method.

Second, there’s a question, especially for agentialists, regarding the norms of reasoning. Does being an ideal believer require that you are prepared to endorse all your beliefs all the time, or does it instead permit ‘taking a step back’ following Burge’s picture of critical reasoning? One advantage to Burge’s ideal of rational reflection compared to Moran’s deliberation model concerns higher-order defeat. Sometimes we encounter defeat for our belief that *p* in the form of evidence that *p* is false. But other times, we might acquire defeat concerning our reasoning processes themselves, such as if we realise we’ve taken a cognitively-impairing drug. While these cases pose their own problems, it’s at least plausible that being an ideally rational reasoner doesn’t just involve considering whether *p* is true but also reflecting on your capacities in this way.[[17]](#footnote-17)

Third, another obvious broader choice point concerns whether we have sui generis agency regarding our beliefs (i.e., a distinctive form of agency that doesn’t just reduce to our agency in performing intentional actions). If we deny this, as many in fact do, then the standard agentialist picture is already a non-starter.

Lastly, self-knowledge raises important questions regarding epistemic warrant, especially if we reject accounts that appeal to a more common-garden epistemology, like inferentialism or quasi-perception.

# 3 Self-knowledge of reasons for belief

So far, I’ve canvassed the self-knowledge of belief. But what about *why* we hold our beliefs (and also other attitudes, and why we act as we do)? In comparison to self-knowledge of belief, this question has received less attention.

## 3.1 The ‘Orthodoxy’

The standard thought is that even if we have special access to our beliefs and other mental states, we don’t have any special access to why we hold them. Gertler quickly sets aside this possibility in her introductory book, and with seemingly good reason:

‘Even the staunchest proponents of privileged access acknowledge that we lack privileged access to these causal relations. So we should be wary of attempts to challenge the general idea of privileged access by citing cases in which subjects are ignorant of the causal sources of their attitudes or actions to challenge the general idea of privileged access. […] [T]o show that we lack privileged access to such matters is to attack a straw man’ (2011: 75).

Instead, the thought is that even if we don’t infer our beliefs, we can only learn *why* we hold them through inference (Nisbett and Wilson 1977).

Gertler presents two main reasons for thinking this (2011: 72-5). First, many experiments illustrate the phenomenon of confabulation, i.e., that we sometimes concoct incorrect explanations of our attitudes and actions. For example, in one paradigm study, Nisbett and Wilson (1977) displayed four identical pairs of pantyhose and asked participants which pair they preferred. Subjects tended to pick those placed towards the right of the table. When asked why, the subjects cited features such as the supposedly better quality, sheerness, and weave when in fact the position of the tights played the real explanatory role. This shows that subjects can be ignorant of correct explanations, and also that their self-ascriptions aren’t especially reliable. And we might think that these mistaken self-ascriptions were formed inferentially, and from this, the simplest conclusion is that we always use inference to learn of them (on this argument, see Carruthers 2013, Nisbett and Wilson 1977).

Second, explaining our beliefs is to cite a causal relation. But, so one might think, we can only learn of causes inferentially. So we cannot have distinctive, i.e. non-inferential, self-knowledge of what explains our attitudes (Gertler 2011: 74).

Let’s start by finessing the claim that subjects lack distinctive access to why they hold their beliefs. There’s several ways of cashing out ‘why’ someone believes something, and arguably different kinds of ‘reason’. Suppose I believe that the bread is fully cooked. It may well be that I myself hold this belief on the basis of reasons, e.g., for the reason that the crust looks brown. These are the reason *for which* I believe – my ‘motivating’ or ‘operative’ reasons. In the epistemic case, this is to say that I *base* my belief on these reasons. Motivating reasons may well be a kind of cause, but they’re a special kind. They differ from what we can call ‘purely causal reasons’, like how a hot oven is a reason why the bread cooks quickly. E.g., we might also explain a belief by citing c-fibres firing and biases which cause it without being *the subject’s* reasons. It’s important to note that talking of reasons in these ways isn’t to say that they’re normative reasons, i.e., good reasons. Maybe the crust isn’t in fact brown. The contentious point, then, is just whether we have distinctive access to our motivating/operative reasons. It’s obvious that we don’t have privileged access to our biases and underlying neurological processes, or to what are in fact good reasons.

## 3.2 Against the Orthodoxy

We can challenge the orthodoxy in various ways. And despite the lack of explicit discussion, various authors implicitly oppose the orthodoxy, by defending views that conflict with it.

(1) On some accounts of the basing relation, believing that *p* for the reason that *q* isn’t a causal relation after all. Instead perhaps it consists in believing that *p*, believing that *q*, and believing that p *is a good reason for believing that* q, (e.g., Setiya 2013). This avoids the worry concerning non-inferential knowledge of causal relations. Furthermore, it follows naturally from this account that we also have distinctive access to our reasons. After all, self-knowledge of our reasons in this case would just straightforwardly reduce to the self-knowledge of belief, namely knowing that you believe that p *is a good reason for believing that* q (Setiya 2013).

(2) That we have distinctive self-knowledge of our reasons follows naturally from an agentialist account of self-knowledge given the role it gives to reasoning and deliberation. For example, recall that for Burge, part of occupying the reviewing perspective is that we can evaluate our beliefs and *our reasons for them*; surely our rational entitlements would therefore extend to self-ascribing our reasons as reasons. And following a Moranean approach, doxastic deliberation involves considering the relevant reasons and forming a belief on their basis. So, if we are entitled to assume that we are deliberating, then we would presumably also be entitled to assume that we hold the belief on the basis of the reasons we considered (Boyle 2011b: 7-8).

(3) There’s also support and tools available from broad perceptual models of self-knowledge, even if we take epistemic basing to be a kind of causal relation. Recall Armstrong’s remark from earlier that we can still be said to quasi-perceive beliefs even if we individuate them by their causal relations, and that bodily awareness of pressure already serves as precedence. More controversially, some have also argued that we can visually perceive causal properties (e.g., Siegel 2009). So facts concerning subjects’ reasons could also be introspectable, especially if all that matters is the flow of information.

(4) Keeling (2021) argues that the question ‘why’ we hold an attitude plays a dual role (either when posed by others or ourselves). It at once requests explanation and also justification for the attitude. As such, we wouldn’t be taking the question seriously if we just inferred our reasons using behavioural and social evidence, since this wouldn’t take into account the quality of the reasons.

(5) Perhaps the patterning of confabulation cases actually speaks in favour of there being a distinct method for self-ascribing our reasons. Note that in confabulation cases (like the pantyhose experiment), subjects erroneously self-ascribe motivating reasons as opposed to purely causal explanations. This suggests that subjects habitually try to answer the question why they think the way they do by considering the good reasons for this position (Cox 2018), and are perhaps also motivated by certain desires to satisfy rational requirements (Keeling 2018).

## 3.3 Developing an account of how we know our reasons

The best account of this self-knowledge, and the ways in which it’s distinctive, has been less explored compared to self-knowledge of beliefs themselves. But there’s many corresponding options. Perhaps subjects quasi-perceive the fact that they believe that *p* for the reason that *q* using a self-scanning mechanism. Or maybe they use a transparency method, and treat the question ‘why do I believe that *q*?’ as transparent to the question ‘what are the normative reasons for believing that *q*?’, or to avoid jargon, ‘what are the good reasons?’ or ‘what is the evidence?’ (Keeling 2021, 2019, Neta 2019b). Maybe we can be neutral on the method but insist that the self-knowledge is transcendentally warranted in virtue of the nature of critical reasoning. There’ll also be many more options corresponding to other accounts of self-knowledge.

Some of the available responses to Gertler’s doubts should now be clear. When assessing the threat of confabulation, it’s important to recall the scope of the distinctive access claim. Illustrating that subjects can fail to know about the position effect and other heuristics and biases doesn’t show that we’re ignorant of our *motivating* reasons. Confabulation does persuasively illustrate a tendency to mistakenly ascribe motivating reasons we don’t have. But self-knowledge can still be acquired and warranted in a distinctive way even if it isn’t especially reliable. And we’ve already seen some responses to the worry that we can’t learn of causal relations without inference: perhaps motivating reasons aren’t causes, and even if they are, there’s other possible examples of non-inferential knowledge of causes. Also, even if we don’t know many causes non-inferentially, this isn’t to say that we can’t know *any*. Reasons are a special sort of cause which play a special role in rational thought, and so we could make a principled exception for them (Keeling 2021: 338, Cox 2018: 181).

## 3.4 Implications

As with belief, our accounts of what it is to have a motivating reason and subjects’ self-knowledge of them interact in various ways. First, it might be that part of what it is to believe for a reason is to be aware (or be in a position to become aware) of doing so. This seems a promising line given the notable tradition of this sort of idea in the philosophy of action (e.g., following Anscombe 2000). Neta (2019a) argues that basing consists, very roughly, in representing the relation as the sort of relation that would confer justification, i.e., as being an instance of epistemic basing. Keeling (2024a) writes that motivating reasons are such that necessarily, we are in a position to self-ascribe them. And relatedly, we might think that the act of consciously inferring from a premise to a conclusion is self-conscious (Jenkins 2018, Koziolek 2021, Marcus 2021). Second, different moves regarding self-knowledge will be available depending on if we adopt a causalist or non-causalist account of basing. And third, the relation between having a motivating reason and believing that it is a good reason is also significant, especially for transparency accounts that say that we learn what our reason is by considering what is in fact a good reason.

\*\*\*

In conclusion, self-knowledge cuts across a number of philosophical fields especially in the philosophy of mind and moral psychology. I’ve therefore hopefully shown why it’s understandable why it took so long to realise that I also work in epistemology. And it should also be clear why this was a mistake, given the numerous implications for epistemological issues such as forms of warrant, reasoning, the nature of belief, and doxastic agency.

**Works cited**

Anscombe, G.E.M. (1957) *Intention*. 2nd edition. Cambridge, MA; London: Harvard University Press, 2000.

Armstrong, D.M. (2001) *A Materialist Theory of the Mind*. London; New York: Routledge.

Bilgrami, A. (2006) *Self-Knowledge and Resentmet*. Cambridge, MA; London: Harvard University Press.

BonJour, L. (2003) ‘Back to Foundationalism’, in BonJour, L. and Sosa, E., *Epistemic Justification: Internalism Vs. Externalism, Foundations Vs. Virtues*. Oxford: Blackwell Publishing, pp.60–76.

Borgoni, C. (2015) ‘Epistemic Akrasia and Mental Agency’, *Review of Philosophy and Psychology*, 6(4), pp.827–842.

Boyle, M. (2011a) ‘“Making up Your Mind” and the Activity of Reason’, *Philosophers’ Imprint*, 11(17), pp.1–24.

Boyle, M. (2011b) ‘Transparent Self-Knowledge’, *Proceedings of the Aristotelian Society Supplementary Volume*, 85(1), pp.223–241.

Burge, T. (1996) ‘Our Entitlement to Self-Knowledge’, *Proceedings of the Aristotelian Society*, 96(1), pp. 91–116.

Burge, T. (1999) ‘A century of deflation and a moment about self-knowledge’, *Proceedings and Addresses of the American Philosophical Association*, 73(2), pp.25–46.

Byrne, A. (2005) ‘Introspection’, *Philosophical Topics*, 33(1), pp.79–104.

Byrne, A. (2011) ‘Transparency, Belief, Intention’, *Proceedings of the Aristotelian Society Supplementary Volume*, 85(1), pp.201–221.

Byrne, A. (2018) *Transparency and Self-Knowledge*. Oxford: Oxford University Press.

Campbell, L. and Greenberg, A. (2023) ‘Mental agency and rational subjectivity’, *European Journal of Philosophy* [preprint].

Carruthers, P. (2010) ‘Introspection: Divided and Partly Eliminated’, *Philosophical and Phenomenological Research*, 80(1), pp.76–111.

Carruthers, P. (2013) *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford: Oxford University Press.

Cassam, Q. (2014) *Self-Knowledge for Humans*. Oxford: Oxford University Press.

Cassam, Q. (2017) ‘What asymmetry? Knowledge of self, knowledge of others, and the inferentialist challenge’, *Synthese*, 194, pp.723–741.

Chalmers, D.J. (2003) ‘The Content and Epistemology of Phenomenal Belief’, in Q. Smith and A. Jokic (eds) *Consciousness: New Philosophical Perspectives*. Oxford: Clarendon Press, pp.220–272.

Chisholm, R. (1982) *The Foundations of Knowing*. Brighton: The Harvester Press.

Coliva, A. (2009) ‘Self-knowledge and commitments’, *Synthese*, 171, pp.365–375.

Cox, R. (2018) ‘Knowing Why’, *Mind & Language*, 33, pp.177–197.

Descartes, R. (1644) ‘Principles of Philosophy’, in J. Cottingham, R. Stoothoff, and D. Murdoch (eds) *The Philosophical Writings of Descartes Volume I*. Cambridge, UK: Cambridge University Press, 1985.

Evans, G. (1982) *The Varieties of Reference*. Edited by J. McDowell. Oxford; New York: Clarendon Press.

Fernández, J. (2013) *Transparent Minds: A Study of Self-Knowledge*. Oxford: Oxford University Press.

Gallois, A. (1996) *The World Without, the Mind Within: An Essay on First-Person Authority*. Cambridge, UK: Cambridge University Press.

Gertler, B. (2001) ‘Introspecting Phenomenal States’, *Philosophy and Phenomenological Research*, 63(2), pp.305–328.

Gertler, B. (2011) *Self-Knowledge*. London; New York: Routledge.

Gertler, B. (2012) ‘Renewed Acquaintance’, in D. Smithies and D. Stoljar (eds) *Introspection and Consciousness*. New York: Oxford University Press, pp.89–123.

Gertler, B. (2018) ‘Self-Knowledge and Rational Agency: A Defense of Empiricism’, *Philosophical and Phenomenological Research*, 96(1), pp.91–109.

Gertler, B. (2021) *Self-Knowledge*, *The Stanford Encylopedia of Philosophy (Winter 2021 Edition)*. Available at: https://plato.stanford.edu/archives/win2021/entries/self-knowledge/.

Ginet, C. (2001) ‘Deciding to Believe’, in M. Steup (ed.) *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford: Oxford University Press, pp.63–76.

Hieronymi, P. (2006) ‘Controlling Attitudes’, *Pacific Philosophical Quarterly*, 87(1), pp.45–74.

Horgan, T. (2012) ‘Introspection about Phenomenal Consciousness: Running the Gamut from Infallibility to Impotence’, in D. Smithies and D. Stoljar (eds) *Introspection and Consciousness*. New York: Oxford University Press.

Hurlburt, R.T. and Schwitzgebel, E. (2007) *Describing Inner Experience? Proponent Meets Skeptic*. MIT Press.

Jenkins, D. (2018) *How inference isn’t blind: Self-conscious inference and its role in doxastic agency*. Doctoral thesis. King’s College London.

Keeling, S. (2018) ‘Confabulation and Rational Obligations for Self-Knowledge’, *Philosophical Psychology*, 31(8), pp.1215–1238.

Keeling, S. (2019) ‘The Transparency Method and Knowing Our Reasons’, *Analysis*, 79(4), pp.613–621.

Keeling, S. (2021) ‘Knowing our Reasons: Distinctive Self‐Knowledge of Why We Hold Our Attitudes and Perform Actions’, *Philosophy and Phenomenological Research*, 102(2), pp.318–341.

Keeling, S. (2021) 'Knowing Our Reasons: Distinctive Self-Knowledge of Why We Hold Our Attitudes and Perform Actions', *Philosophy and Phenomenological Research,* 102(2), pp. 318-341.

Keeling, S. (2024a) ‘Believing for a Reason is (at Least) Nearly Self-Intimating’, *Erkenntnis*, 89, pp. 241-260.

Keeling, S. (2024b) ‘Doxastic Agent’s Awareness’, *European Journal of Philosophy* (online

advance) [ttps://doi.org/10.1111/ejop.13006](https://doi.org/10.1111/ejop.13006)

Koziolek, N. (2021) ‘Inferring as a way of knowing’, *Synthese*, 198(Suppl 7), pp.S1563–S1582.

Leary, S. (2017) ‘In Defense of Practical Reasons for Belief’, *Australasian Journal of Philosophy*, 95(3), pp.529–542.

Leite, A., 2018. Changing One’s Mind: Self-Conscious Belief and Rational Endorsement. *Philosophy and Phenomenological Research,* 97(1), pp.150–171.

Leite, A. (2008) ‘Believing one’s reasons are good’, *Synthese*, 161, pp.419–441.

Locke, J. (1689) *An Essay Concerning Human Understanding*. Edited by P.H. Nidditch. Oxford: Oxford University Press, 1975.

Marcus, E. (2021) *Belief, Inference, and the Self-Conscious Mind*. New York: Oxford University Press.

McCormick, M.S. (2015) *Believing Against the Evidence: Agency and the Ethics of Belief*. Abingdon; New York: Routledge.

McCormick, M.S. (2018) ‘Responding to Skepticism About Doxastic Agency’, *Erkenntnis*, 83, pp.627–645.

McGeer, V. (1996) ‘Is “Self-Knowledge” An Empirical Problem? Renegotiating the Space of Philosophical Explanation’, *The Journal of Philosophy*, 93(10), pp.483–515.

McHugh, C. (2013) ‘Epistemic Responsibility and Doxastic Agency’, *Philosophical Issues*, 23, pp.132–157.

Moran, R. (2001) *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton; Oxford: Princeton University Press.

Moran, R. (2003) ‘Responses to O’Brien and Shoemaker’, *European Journal of Philosophy*, 11(3), pp. 402–419.

Neta, R. (2019a) ‘The Basing Relation’, *The Philosophical Review*, 128(2), pp. 179–217.

Neta, R. (2019b) ‘The Transparency of Inference’, in A. Nes and T. Chan (eds) *Inference and Consciousness*. Routledge, pp. 275–89.

Nisbett, R.E. and Wilson, T.D. (1977) ‘Telling More Than We Can Know: Verbal Reports on Mental Processes’, *Psychological Review*, 84(3), pp. 231–259.

O’Brien, L. (2005) ‘Self-Knowledge, Agency and Force’, *Philosophy and Phenomenological Research*, 71(3), pp.580–601.

O’Brien, L. (2007) *Self-Knowing Agents*. Oxford; New York: Oxford University Press.

Parrott, M. (2015) ‘Expressing first-person authority’, *Philosophical Studies*, 172, pp.2215–2237.

Peacocke, A. (2017) ‘Embedded mental action in self-attribution of belief’, *Philosophical Studies*, 174, pp.353–377.

Peacocke, C. (1998) ‘Conscious Attitudes, Attention, and Self-Knowledge’, in C. Wright, B.C. Smith, and C. Macdonald (eds) *Knowing Our Own Minds*. Oxford: Clarendon Press, pp.64–98.

Peacocke, C. (2007) ‘Mental Action and Self-Awareness (I)’, in B.P. McLaughlin and J.D. Cohen (eds) *Contemporary Debates in Philosophy of Mind*. Malden, MA: Blackwell, pp.358–376.

Peacocke, C. (2009) ‘Mental Action and Self-Awareness (II)’, in L. O’Brien and M. Soteriou (eds) *Mental Actions*. Oxford; New York: Oxford University Press, pp.192–214.

Pillar, C. (2015) ‘Normative Practical Reasoning’, *Aristotelian Society Supplementary Volume*, 75(1), pp.195–216.

Pitt, D. (2004) ‘The Phenomenology of Cognition or “What Is It like to Think That P?”’, *Philosophy and Phenomenological Research*, 69(1), pp.1–36.

Reisner, A.E. (2009) ‘The Possibility of Pragmatic Reasons for Belief and the Wrong Kind of Reasons Problem’, *Philosophical Studies*, 145(2), pp.257–272.

Reisner, A.E. (2014) ‘A Short Refutation of Strict Normative Evidentialism’, *Inquiry*, 58(5), pp.477–485.

Rinard, S. (2019) ‘Believing for Practical Reasons’, *Noûs*, 53(4), pp.763–784.

Russell, B. (1917) ‘Knowledge by Acquaintance and Knowledge by Description’, in *Mysticism and Logic*. London: George Allen and Unwin.

Ryle, G. (1949) *The Concept of Mind*. London: Routledge, 2009.

Sartre, J.-P. (1956) *Being and Nothingness*. Translated by H.E. Barnes. New York: Philosophical Library.

Schwitzgebel, E. (2008) ‘The Unreliability of Naive Introspection’, *Philosophical Review*, 117(2), pp.245–273.

Schwitzgebel, E. (2009) ‘Knowing Your Own Beliefs’, *Canadian Journal of Philosophy*, 35(supplement), pp.41–62.

Schwitzgebel, E. (2012a) ‘Introspection, What?’, in D. Smithies and D. Stoljar (eds) *Introspection and Consciousness*. Oxford University Press.

Schwitzgebel, E. (2012b) ‘Self-Ignorance’, in J. Liu and J. Perry (eds) *Consciousness and the Self: New Essays*. New York: Cambridge University Press, pp.184–197.

Setiya, K. (2013) ‘Epistemic Agency: Some Doubts’, *Philosophical Issues*, 23(1), pp.179–198.

Shoemaker, S. (1994a) ‘Self-Knowledge and “Inner Sense”: Lecture I: The Object Perception Model’, *Philosophy and Phenomenological Research*, 54(2), pp.249–269.

Shoemaker, S. (1994b) ‘Self-Knowledge and “Inner Sense”: Lecture II: The Broad Perceptual Model’’, *Philosophy and Phenomenological Research*, 54(2), pp.271–290.

Shoemaker, S. (2012) ‘Self-Intiimation and Secong-Order Belief’, in D. Smithies and D. Stoljar (eds) *Introspection and Consciousness*. Oxford; New York: Oxford University Press, pp.237–257.

Siegel, S. (2009) ‘The Visual Experience of Causation’, *The Philosophical Quarterly*, 59(236), pp.519–540.

Siewert, C.P. (1998) *The Significance of Consciousness*. Princeton: Princeton University Press.

Smithies, D. (2012) ‘A Simple Theory of Introspection’, in D. Smithies and D. Stoljar (eds) *Introspection and Consciousness*. New York: Oxford University Press, pp.259–293.

Wilson, T.D. (2002) *Strangers to Ourselves*. Cambridge, MA; London: Belknap.

1. E.g., Descartes in the *Meditations* advocates something like this concerning our thoughts, provided one employs the distinctive method with ‘great care’ Descartes (1644: I.66). And Chalmers (2003: 241-246) and Horgan (2012) argue that a certain sort of belief about our experiential states is infallible. [↑](#footnote-ref-1)
2. E.g., Locke claims that ‘[w]hen we see, hear, smell, taste, feel, meditate, or will any thing, we know that we do so’ (1689: II.27.ix). [↑](#footnote-ref-2)
3. E.g., Chisholm claims this about thoughts, feelings, and beliefs (1982: 9-11, 25), Siewert about conscious experience (1998: 197-8), and Smithies about mental states including belief and pain (2012). [↑](#footnote-ref-3)
4. E.g., Shoemaker (1994a, 1994b, 2012), Burge (1996, 1999), Moran (2001), Bilgrami (2006), and Boyle (2011b). [↑](#footnote-ref-4)
5. Against this, see Cassam (2017: §4). [↑](#footnote-ref-5)
6. E.g., see the contrast in Moran (2001) and Parrott (2015). [↑](#footnote-ref-6)
7. Due to space, I must leave aside ‘acquaintance’ accounts, whereby the lower-order state partly constitutes the self-ascriptive belief. E.g., Descartes (1644), Russell (1917), Gertler (2012, 2001), Chalmers (2003), Pitt (2004), and BonJour (2003) think that the lower-order state partly constitutes the self-ascriptive belief. Another related alternative is an ‘object perception model’, which takes the state and self-ascription to be metaphysically distinct, but relies on a rich notion of perception (for description without endorsement, see Shoemaker 1994b). [↑](#footnote-ref-7)
8. A related inner sense account is Lycan (1995). Or we could even have an account that’s more minimal still. MacDonald (2014) denies that we have a self-scanning mechanism in her account of self-knowledge of conscious thought. [↑](#footnote-ref-8)
9. Accounts in the vicinity but which aren’t as extreme include Wilson (2002), Hurlburt and Schwitzgebel (2007), and Schwitzgebel (2008, 2009, 2012a, 2012b). [↑](#footnote-ref-9)
10. Related positions include Fernández (2013). [↑](#footnote-ref-10)
11. See also McHugh (2013), McCormick (2018: ch.6; 2015: 630) and also Hieronymi (2006) more generally on responsibility and (non-voluntary) control over belief. [↑](#footnote-ref-11)
12. We might also distinguish between managerial/evaluative control (Hieronymi 2006), or believing for attitude-related/content-related reasons (Pillar 2015). [↑](#footnote-ref-12)
13. E.g., Ginet (2001), Leary (2017), McCormick (2018, 2015) Reisner (2014, 2009), and Rinard (2019). [↑](#footnote-ref-13)
14. Borgoni (2015) argues that many such beliefs aren’t in fact alienated even if they aren’t under one’s rational control, because they are nevertheless one’s own. The subject still sees them as *their own* resistant beliefs, which would leave open the possibility of an alternative form of first-personal self-knowledge of them. [↑](#footnote-ref-14)
15. See also Gallois’ (1996) rationalist transparency account. [↑](#footnote-ref-15)
16. See also Schwitzgebel (2009). [↑](#footnote-ref-16)
17. See also Leite (2018) on double-checking our beliefs. [↑](#footnote-ref-17)