**Agency and self-knowledge**

*To appear in The Routledge Handbook for the Philosophy of Agency, edited by Luca Ferrero. This is the penultimate version.*

Brie Gertler

# Self-Knowledge

The term “self-knowledge” is standardly used, by philosophers, to refer to the kind of knowledge that each of us (sometimes) has of our own mental states: what it is that we think, feel, believe, or desire. This chapter concerns self-knowledge in that sense, with a focus on knowledge of what we believe to be true, and knowledge of what we intend to do. In particular, we will examine the influential idea that beliefs and intentions are exercises of *agency*, and the implications of that idea for accounts of the individual’s knowledge of her own beliefs and intentions.

## An initial contrast: agency and knowledge

Think about some ordinary ways in which we exercise agency. I kick the can, and it bounces down the road. I call my dog, and he runs towards me. In these cases, my agency involves my affecting objects or phenomena. By contrast, ordinary cases of *knowledge* typically involve some object or phenomenon affecting *me*. I know that the can is bouncing down the road because its movement affects my visual experience—I see it bouncing. I know that my dog is running towards me because the crackling of dry leaves under his paws shapes my auditory experience—I hear him coming. So the direction of control or causation ordinarily operative in knowledge seems to be the reverse of that in which agency consists. Agency typically involves *affecting* some object or phenomenon, whereas knowledge typically involves *being affected* *by* some object or phenomenon.[[1]](#endnote-1)

This general idea, that knowledge involves being affected by the object or phenomenon that is known, clearly applies to some knowledge of one’s own mental states. When the doctor asks me whether the pain I’m feeling is a dull ache or a sharp twinge, I reflect on the pain so that my response will be shaped by the quality of the pain: its dullness or sharpness. When trying to decide what flavor of ice cream to order, my preference—the fact that pistachio appeals to me more than vanilla, say—will shape my thought that I’d prefer pistachio, and will thereby help to explain why I order pistachio.

## “Active” thoughts and attitudes

We have just seen that, on a standard model, knowledge consists in an awareness—e.g., a thought or belief—that is shaped by what is known. According to a view I call *agentialism*, this model cannot exhaustively explain self-knowledge, knowledge of our own mental states. Agentialists claim that the standard model is plausible only as regards self-knowledge of *passive* states, like pains and flavor preferences.[[2]](#endnote-2) On their view, there are also *active* mental states, and an alternative model is needed to explain self-knowledge of these active states.

Active mental states are those that involve exercises of agency. They form two broad classes. The first class is exemplified by episodes of goal-directed conscious thinking, such as imagining how Fred would look in a purple cardigan. It’s natural to say that this is something that you do, rather than something that happens to you, or an experience you merely undergo. And some philosophers will maintain that our knowledge of what we’re doing is not achieved purely through observation: after all, simply observing the image of a man in a purple cardigan does not reveal whether you’re imagining Fred or his identical twin. We will not discuss these sorts of cases here; they are discussed in [**entry xxx**].

Our focus will be the second class of mental items regarded as active: *committal attitudes* such as beliefs and intentions. According to agentialists, we don’t simply find ourselves with the belief that liberal democratic ideals are on the decline, or with the intention to visit an ailing friend. In paradigmatic cases, believing and intending are things we *do*, for *reasons*. Suppose that, upon assessing current geopolitical trends and considering them against the backdrop of history, you conclude that liberal democratic ideals are waning in influence. Weighing the evidence and drawing a conclusion is something that you *do*, not merely an experience you undergo. You are similarly active when, reflecting on the obligations of friendship, or simply thinking about your ailing friend, you decide to visit him.

These examples illustrate that beliefs and intentions often result from theoretical or practical reasoning. Moreover, existing attitudes are apt to change in the face of new evidence or practical considerations. In this sense, beliefs and intentions are often responsive to reasons. Agentialists take reasons-responsiveness to constitute a kind of agency. So they regard believing and intending, in a way that is responsive to reasons, as exercises of agency: that is, as a type of *doing*.[[3]](#endnote-3) And since we are generally responsible for what we do, we are responsible for our beliefs and intentions. We are therefore subject to normative assessment—credit or blame—according to how well our beliefs and intentions conform to our reasons. “Our rational beliefs and intentions are not mere mental attitudes, but active states of normative commitment” (Korsgaard 2009, 39).

The agentialist maintains that, because reasons-responsive beliefs and intentions are exercises of agency, self-knowledge of those attitudes contrasts with self-knowledge of passive states like sensations. In particular, self-knowledge of beliefs and intentions does not fit the picture described earlier, on which knowledge involves being affected bythe known object or phenomenon. Suppose that someone asks you whether you intend to visit your friend. If you answer this question by consulting your memory, your belief that you intend to do this is shaped by that intention itself, which you later remember. In effect, you are regarding your intention from the perspective of a passive observer, the perspective you occupy when observing an ache or tickle. You are not treating the intention to visit your friend as a commitment to be endorsed. Similarly, if you respond to the question “do you believe that liberal democratic ideals are on the decline?” by trying to remember your past statements on that topic, you’re treating this opinion as an inert state to be discovered, rather than as a conviction to uphold or disown, that is, a commitment for which you bear responsibility. In both of these cases, the method you use is available to others: someone else could check your calendar, or remember your past comments about liberal democracy.

Agentialists allow that self-knowledge of sensations and other passive mental states is a matter of observing (or otherwise tracking) those states. But, they contend, a subject cannot apprehend her own beliefs and intentions *as* “active states of normative commitment” merely by observing her cognitive life. For this reason, agentialists maintain that self-knowledge of beliefs and intentions differs significantly from self-knowledge of passive mental states.

## First-person authority

Accounts of self-knowledge generally aim to explain “first-person authority”, the special authority that each of us seems to possess as regards our own mental states. As evidence for this authority, consider how odd it is, in ordinary circumstances, to challenge someone’s report about what they believe or intend. If you say “I believe liberal democratic ideals are on the decline” or “I intend to cook dinner tomorrow”, it would be odd for someone to respond “No you don’t”.

Although it’s generally agreed that we possess first-person authority, the nature and source of this authority are disputed. On one view, the oddness just described is fully explained by the idea that we have first-person *epistemic* authority about our own attitudes: that is, each of us is in an epistemically privileged position to identify our beliefs and intentions. This means that, as compared with others’ beliefs about my attitudes, my own self-attributions are more strongly justified, or open to fewer sources of error, or produced by processes that are more reliable (etc). Claims of first-person epistemic authority are standardly expressed as claims about *privileged access* to one’s own mental states.

Agentialists argue that construing first-person authority in purely epistemic terms neglects the more significant kind of authority we possess, in believing or intending: viz., that these are acts of first-person agency. In saying “I believe liberal democratic ideals are on the decline” or “I intend to visit my friend tomorrow”, I am not simply reporting on an internal state that I’m in a special position to know about, as I might report that I have a stomachache. Instead, I am expressing my commitment to the corresponding view or plan, and thereby taking responsibility for it. Agentialists call such expressions of commitment *avowals*. My avowal conveys that you can hold me accountable for the avowed attitude: you can reasonably ask me to justify this pessimism about liberal democracy, or to explain why it is reasonable or appropriate for me to visit my friend tomorrow. The authority to avow an attitude belongs exclusively to the subject, because only the subject herself can take responsibility for her attitudes. In this sense, the agent is said to possess first-personal *agential* authority relative to her attitudes.

Agential authority contrasts sharply with epistemic authority. Others can know that I am in pain, or that I intend to cook dinner tomorrow. In this sense, privileged epistemic access admits of degrees: it may consist in the subject’s having a way of detecting her own mental states that is more reliable than the ways others can detect them, or in one’s own self-ascriptions being more strongly justified than others’ ascriptions of mental states to her (etc).[[4]](#endnote-4)

By contrast, the authority to take responsibilityfor a belief or intention is an all-or-nothing matter, as beliefs and intentions can be avowed only by the person who believes or intends. Consider the following scenario.

Nigella wants to cook eggplant parmigiana, but she loathes grocery shopping. So she tells her brother Nigel “You should cook eggplant parmigiana tonight!”, knowing that once he endorses this plan he’ll buy the necessary ingredients. Nigella stills plan to cook the dinner herself—Nigel is a terrible cook.

In this case, Nigella exercises agency in bringing it about that Nigel intends to cook tonight. Still, Nigel’s intention to cook is not an exercise of Nigella’s agency. Nigella cannot avow Nigel’s intention. Persuading Nigel to cook, and asserting “Nigel intends to cook tomorrow”, do not make Nigella responsible for justifying his intention. After all, Nigella thinks it would beunreasonable for Nigel to cook. To justify her action, Nigella need only cite possible benefits of Nigel’s *intending* to cook tonight. By contrast, Nigel can justify his intention only by specifying reasons (for him) to actually cook tonight—e.g., possible benefits of his cooking.

It is now clear why our agential authority, relative to our beliefs and intentions, strikes many philosophers as more remarkable than our privileged epistemic access to those attitudes. Agential authority is all or nothing, as only the subject can take and bear responsibility for her attitudes, whereas epistemic privilege is a matter of degree. Moreover, agential authority uniquely attaches to one’s own reasons-responsive attitudes and other exercises of agency, whereas epistemic privilege can extend to any phenomenon the subject is especially well-situated to appreciate: I have a degree of epistemic authority as regards my dog’s whereabouts so long as no one else is around to observe him.

Now it may appear that agential authority, however remarkable, doesn’t bear on the question of self-knowledge, since self-knowledge is an epistemic phenomenon. But as we will see, leading versions of agentialism take the distinctive epistemic characteristics of self-knowledge of beliefs and intentions to derive from the subject’s agential authority as regards these attitudes.

# Agential models of self-knowledge

We’ve just seen the case against assimilating self-knowledge of beliefs and intentions to self-knowledge of sensations, flavor preferences, and other “passive” states. We now face the question of how to account for self-knowledge of beliefs and intentions. Granting that beliefs and intentions are sometimes acts of agency, what does this mean for self-knowledge of these attitudes? In other words, how is first-personal epistemic authority related to first-personal agential authority?

## Rational Agency as Providing Epistemic Entitlement to Self-Knowledge

The most prominent agentialist accounts (Burge 1996, Moran 2001, Boyle 2009) take our epistemic authority, relative to our attitudes, to derive from our agential authority—specifically, our authority as rational thinkers. These accounts are motivated by the view that beliefs and intentions are *active* insofar as they are responsive to reasons.

Tyler Burge’s version of agentialism centers on the idea that, insofar as we are rational thinkers, we have certain obligations. We are obligated to (try to) avoid inconsistent beliefs, to conform our beliefs to our evidence, and to strive to satisfy other rational norms. Burge claims that we could not hope to satisfy these obligations unless we had the capacity for self-knowledge, since we must know what we believe in order to assess our beliefs for consistency with each other, and for conformity with our evidence. In this way, he argues, our obligations as rational thinkers ensure that we are capable of self-knowledge. Burge takes this to mean that our rationality gives us a kind of “epistemic entitlement” to beliefs about our attitudes; it is because of this entitlement that such beliefs (when true) qualify as self-*knowledge*.

On Burge’s account, our epistemic authority as regards our attitudes—our capacity for self-knowledge—derives from our agential authority, the responsibility we bear as rational thinkers. Burge is concerned to locate the basic source of our epistemic authority, but he does not provide an account of the specific method we use to achieve self-knowledge.

Richard Moran shares Burge’s general outlook. Like Burge, he thinks that our capacity to rationally shape our attitudes constitutes a kind of agency, and that this agential authority over our attitudes provides for epistemic authority. Moran goes beyond Burge’s earlier discussion in advancing a view about how our agency provides a route to self-knowledge.

Moran’s proposal is inspired by what he takes to be a distinctive characteristic of the agent’s relation to her own attitudes. What rationalizes your belief that *p*,or your intention to *w*, are reasons directly bearing on the attitude’s content. For example, your evidence bearing on *democracy is on the decline* rationalizes your belief to that effect; your reasons *to make eggplant parmigiana*, rationalize your intention to do so. In contrast, others’reasons for you to have a belief or intention are at best indirectly related to the attitude’s content. Nigella’s reason for persuading Nigel to (intend to) make eggplant parmigiana is that this intention will lead him to buy groceries. This rationale is entirely compatible with Nigella firm belief that there is no good reason for Nigel *to make eggplant parmigiana*. But Nigel’s intention is rationally justified, according to these philosophers, only if *Nigel should make eggplant parmigiana* issupported by reasons, at least from his perspective. Similarly: If I lead you to believe that democracies are on the decline, I can justify this act of persuasion without invoking any reason to think that *democracy is on the decline*: perhaps I think your pessimism would be somehow beneficial. But you can’t rationally justify *believing* that liberal democracy is on the decline except by providing evidence that liberal democracy is indeed on the decline.

To summarize: reasons bearing directly on a belief or intention, rather than on *the* *having of* the attitude, concern the attitudes’ content (*democracy is on the decline* or *to cook eggplant parmigiana*)directly, and directly justify the attitude only for the subject herself, the person who believes or intends.[[5]](#endnote-5)

These considerations lead Moran to argue that the signal characteristic of agential attitudes is that such attitudes are “transparent” to the subject (see also Boyle 2011). I can determine what I believe about *p* by considering evidence bearing on *p* itself, rather than evidence about *whether I believe* that *p*. I can determine whether I intend to *w* by considering reasons for or against *w* itself, rather than evidence about *whether I intend* to *w*. This is known as the transparency method: I grasp my attitudes by “looking through” them, as it were, directly to the facts or actions they concern. Moreover, on this view the transparency of my attitudes (to me) reflects what makes those attitudes truly my own: viz., my rational authority over them. Others may possess the power to control my attitudes through non-rational means (such as brainwashing). My attitudes genuinely belong to me because I exercise rational agency over them.

Moran argues that our agential authority, rooted in the capacity to conform our attitudes to our reasons, explains our epistemic authority relative to our attitudes. On his account, we have the right to presume that we are rational thinkers: that is, that our beliefs and intentions are responsive to reasons. This “epistemic right”, a kind of epistemic entitlement, licenses the inference from “the evidence favors *p*” to “I believe that *p*” (Moran 2003). Your epistemic authority consists in the fact that you—and only you—can grasp your attitudeby directly considering your reasons. This epistemic authority is constituted by your special *agential* position, vis-à-vis your attitudes: your capacity to conform your attitudes to your reasons.

## Self-knowledge as assured by our normative commitments

Akeel Bilgrami also maintains that our responsibility for our attitudes ensures that we possess self-knowledge, but offers an alternative view about how self-knowledge is linked with responsibility. On his view, our rational agency guarantees self-knowledge in two ways. The first focuses on the fact that beliefs and intentions motivate our behavior. Suppose that Nigel buys tomatoes because he *intends* to make eggplant parmigiana and he *believes* that tomatoes are a necessary ingredient for that dish. Assuming that purchasing the tomatoes is a free action, Nigel is responsible for it; and, Bilgrami says, we cannot be responsible for an action unless we understand the attitudes that motivated it. This means that we possess self-knowledge of beliefs and intentions that rationally motivate our behavior.

The second way that rational agency guarantees self-knowledge, on Bilgrami’s view, derives from the fact that beliefs and intentions are *normative commitments*. Having a belief commits me to behaving in ways that are rationalized by this belief: e.g., my belief that it will rain today obligates me to carry an umbrella (assuming that I want to remain dry). Bilgrami maintains that, insofar as we have a commitment, we are disposed to try to live up to it, and to upbraid ourselves if we fail. This link to dispositions requires that we are aware of our commitments. He concludes “We cannot therefore have commitments without believing that we have them.” (Bilgrami 2006, 287) In this way, the normative dimension of beliefs and intentions ensures that we have self-knowledge of those attitudes.

Whereas Bilgrami claims that commitments partly consist in dispositions to think or behave in certain ways, a related view focuses on *events* associated with commitments, namely, avowals. It’s worth noting that, on all of these agentialist views,avowals like “I intend to cook dinner” are themselves exercises of agency. To avow a belief or intention is to take responsibility for it. And to take responsibility is not simply to describe an independent aspect of reality: instead, it is to *make it the case* that I can reasonably be held accountable for the attitude. If I avow an intention, I can be criticized for failing to carry through on it, or even for *having*  it—if the intention is not aligned with my reasons, or if it is malevolent or foolish, say. Because avowals play this role, they are “commissive, not descriptive” (McGeer 1996, 508).

The idea that avowing a belief or intention is commissive forms the basis for Annalisa Coliva’s version of agentialism (Coliva 2012). She argues that, so long as a thinker is rational and possesses the relevant concepts, self-ascriptions of beliefs or intentions will createthe attitudes ascribed, by committingthe thinker to those attitudes. She extends this “constructivist” view to rationally held conative attitudes, such as desires.

When understood in the way proposed, a judgement (or a sincere assertion) such as “I believe/desire/intend/wish/hope that P” is like a *performative*, namely like “I promise to buy you an ice-cream”… : it makes a certain thing *happen*, for it does create the first-order propositional attitude as a commitment. … [This] is possible precisely because judging “I believe/desire/intend/wish/hope that P” becomes just an alternative way of undertaking the same commitments one would make by judging that *P* (is worth pursuing or having)… (Coliva 2012: 235–6)

For Coliva, judgments about one’s attitudes *constitute* forming (or committing to) the attitude. The resulting commitment partly consists in dispositions, as in Bilgrami’s view. But the act of agency just described is what makes it a genuine commitment, for which the thinker is responsible.

## Doubts about Rational Agency

Agentialists take our capacity for self-knowledge of beliefs and intentions to derive, in one way or another, from our rational capacities. Some commentators have objected that agentialist accounts rest on an overly idealized picture of human cognition. Psychological studies indicate that our attitudes are often shaped by non-rational factors such as emotions, moods, and arbitrary features of the context or situation in which we form the attitude. Worse yet, we often manufacture or selectively attend to evidence that confirms our beliefs. The capacity to conform our attitudes to our reasons is of questionable value if our “reasons” are shaped by emotions or biases. And the fact that our attitudes correspond to what we take to be our reasons—and are thus transparent to our reasons—is not a triumph of rationality if our sense of our reasons is shaped by non-rational factors.

Suppose that Roy’s daughter asks him for a loan so that she can realize her dream of opening a restaurant. Roy carefully deliberates about this decision. He is aware that most restaurants close within the first year. But because of his affection for his daughter, and his desire that she succeed, that statistic does not weigh as heavily in his deliberations as it would if the request were coming from someone else. His hopes for his daughter incline him towards an optimistic view of the restaurant’s prospects even before he deliberates. Moreover, this optimistic outlook directs his attention away from counterevidence, such as his daughter’s lack of business experience. And it leads him to overvalue favorable evidence, such as her impressive cooking skills. In the end, Roy concludes that the restaurant would likely succeed. This belief is transparent: Roy can report on it simply by considering (what he regards as) the evidence. But the transparency of his belief cannot be attributed to his rational agency, since his belief shapes his sense of his reasons, rather than vice versa.

The well-established fact that attitudes are often shaped by non-rational factors challenges the epistemic basis of rational agency accounts of self-knowledge. If non-rational factors lead us to overvalue some reasons and blind us to others, then our capacity for rational thought is seriously flawed and limited, and it is hard to see how this capacity could ground an epistemic right to assume that our attitudes are transparent. Even if our reasons result from confirmation bias, post-hoc rationalization, or other dubious sources, our attitudes may well be transparent to *what appears to us to be* our reasons. But that kind of transparency may derive from our psychological need to fit our reasons to our attitudes—or to selectively attend to reasons that justify our attitudes—rather than our rational capacity to conform our attitudes to our reasons. The influence of non-rational factors challenges the idea that our rational agency, and the attendant obligation to reason critically about our attitudes, could provide an epistemic entitlement that undergirds self-knowledge. Put another way: since in some cases the transparency of a belief or intention is due to the fact that the attitude shapes the thinker’s sense of her reasons, the epistemic authority associated with transparency is not (always) grounded in agential authority.[[6]](#endnote-6)

Victoria McGeer presses this kind of objection (McGeer 1996, 2007). She argues that a more accurate depiction of our cognitive lives will rely on a more expansive understanding of first-person agency, beyond the agency involved in responsiveness to reasons. Given that we are not ideally rational cognizers, she says, fulfilling our responsibility for our own attitudes sometimes requires adopting a detached, third-person perspective on our cognitive lives. (Moran regards this as a kind of self-alienation.) From this perspective, we can better identify biases, impulses, and other non-rational factors affecting our sense of our reasons. Recognizing these non-rational factors is required if we are to curb their influence (McGeer 2007, 102).

McGeer’s argument is one strand in a larger skepticism about agentialist accounts of self-knowledge. Given our cognitive flaws and limitations, it’s hard to see how our capacity for rational thought could ground the kind of epistemic rights or entitlements envisioned by agentialists like Burge and Moran.

Another strand of skepticism concerns the claim, shared by Bilgrami and Coliva, that there is a conceptual connection between having a commitment and believing or judging that one does. We are often ignorant about the motivations for our decisions and actions, and we sometimes invent—or “confabulate”—motivations that rationalize our choices or otherwise support our self-image (Nisbett and Wilson 1977). Bilgrami does not deny this, but claims only that we could not be *responsible* for our actions unless we knew the attitudes motivating them. But that claim may imply that we are responsible for far fewer of our actions than we ordinarily assume. Similarly, we may be disposed to criticize ourselves for failing to live up to our commitments, if by “commitments” we mean beliefs and intentions that we’re aware of, or at least would readily endorse. But some of our beliefs and intentions operate outside of conscious awareness. And some of these implicit attitudes—including some racist beliefs and retributive intentions—are at odds with what we would avow. This perhaps means that they are not *commitments* in Bilgrami’s sense. The upshot is that Bilgrami’s account applies most naturally to conscious attitudes that we endorse. The thesis that we possess self-knowledge of conscious attitudes that we would readily endorse is comparatively uncontroversial.

Similar questions can be raised about Coliva’s view. We can grant Coliva’s claim that a judgment or sincere assertion creates a commitment, and that in those cases the commitment is known to the subject. But this doesn’t explain self-knowledge of attitudes that are not created in this way. (This indicates a limitation of Coliva’s view, but not a decisive objection to it, as the claim in question is not intended as a comprehensive account of self-knowledge.)

**3. Conclusion**

We typically act as if individuals exert rational agency over their beliefs and intentions: we hold ourselves responsible for our attitudes, and we think it appropriate to demand justifying reasons for beliefs and intentions. Some philosophers take these practices to reflect a deeper truth, that our beliefs and intentions express our rational agency, and maintain that our rational agency is key to self-knowledge of our beliefs and intentions. Others are skeptical of the extent and significance of our alleged rational agency, relative to our attitudes.

*RELATED TOPICS*

*intentional agency and rational agency; agency and cognition; agency and the first person; agency, reasons, and rationality*

*FURTHER READING*

Cassam, Quassim. 2014. *Self-Knowledge for Humans*. Oxford: Oxford University Press.

*This book contains an extended examination and critique of agentialist accounts of self-knowledge.*

Coliva, Annalisa. 2016. *The Varieties of Self-Knowledge*. London: Palgrave Macmillan.

*This book evaluates a wide range of accounts of self-knowledge. It makes the case for a pluralist view, according to which self-knowledge takes a variety of forms.*

*Works cited*

Bilgrami, Akeel. 2006. *Self-Knowledge and Resentment*. Cambridge, MA: Harvard University Press.

Boyle, Matthew. 2011. Transparent Self-Knowledge. *Proceedings of the Aristotelian Society*, Sup Vol. 85: 223-241

Burge, Tyler. 1996. Our Entitlement to Self-Knowledge. *Proceedings of the Aristotelian Society* 96: 91-116.

Byrne, Alex. 2018. *Transparency and Self-Knowledge*. Oxford: Oxford University Press.

Coliva, Annalisa. 2012. One Variety of Self-Knowledge: Constitutivism as Constructivism. In *The Self and Self-Knowledge*, A. Coliva (ed.), Oxford: Oxford University Press, pp. 212-242.

Fernandez, Jorge. 2013. *Transparent Minds: A Study of Self-Knowledge*. Oxford: Oxford University Press.

Gertler, Brie. 2018. Self-Knowledge and Rational Agency: A Defense of Empiricism. *Philosophy and Phenomenological Research* 96: 91-109.

Korsgaard, C.M. 2009. The Activity of Reason. *Proceedings and Addresses of the American Philosophical Association* 83: 23-43.

McGeer, Victoria. 1996. Is “Self-Knowledge” an Empirical Problem? Renegotiating the Space of Philosophical Explanation. *Journal of Philosophy* 93: 483-515.

Moran, Richard. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, NJ: Princeton University Press.

Moran, Richard. 2003. Responses to O’Brien and Shoemaker. *European Journal of Philosophy* 11: 402–19.

Nisbett, Richard and Timothy Wilson. 1977. Telling More Than We Can Know: verbal reports on mental processes. *Psychological Review* 84: 231-259.

1. The idea that knowledge involves one’s beliefs being shaped by the facts is most plausible for *a posteriori* knowledge. Many philosophers deny that necessary truths knowable *a priori*, such as analytic and mathematical truths, are part of the causal order. On that view, such truths cannot affect thinkers or their beliefs. We will not be concerned with *a priori* knowledge, since the objects of self-knowledge—e.g., that one is undergoing a certain experience, or has certain beliefs or intentions—are knowable only *a posteriori*. [↑](#endnote-ref-1)
2. Pains and flavor preferences are passivestates, even if they arose because of past exercises of agency. If I deliberately kicked myself, or cultivated an appreciation of pistachio ice cream, then I exercised agency in bringing about my pain or my flavor preference. But neither phenomenon—the experience of pain, or the craving for pistachio ice cream—is itself an exercise of agency. [↑](#endnote-ref-2)
3. States like pains and flavor preferences are correspondingly passive is that they are not sensitive to reasons. Reasoned argument to the effect that feeling a pain is unwise or unwarranted does not alleviate the pain. And although you could invite me to sample some ice cream in an effort to persuade me that vanilla tastes better than pistachio, you can’t convince me of this purely by enumerating reasons that I *should* prefer vanilla, or providing objective evidence that vanilla is gustatorily superior. [↑](#endnote-ref-3)
4. Some philosophers argue that there is one kind of first-person epistemic privilege that is an all-or-nothing matter, namely, *acquaintance*. No one else can be acquainted with your mental states. Loosely speaking, a subject is *acquainted* with a mental state when her awareness of that state is directly linked to the state itself – where the notion of directness here is metaphysical, and contrasts with the much weaker causal relations that are implicit in the standard model discussed at the outset of this entry. For more on the notion of acquaintance, see Hasan 2014. [↑](#endnote-ref-4)
5. Doesn’t my evidence that liberal democracy is on the decline justify my persuading you of this? Not unless there is some reason for me to maximize your true beliefs. [↑](#endnote-ref-5)
6. The claim that attitudes are transparent to reasons is in itself purely epistemic; it is not committed to the idea that attitudes are exercises of agency. Accounts of self-knowledge that center on transparency, and do not rely on claims about rational agency, have recently been advanced by Byrne (2018) and Fernandez (2013). [↑](#endnote-ref-6)