Marcus on Self-Conscious Knowledge of Belief

James R. Shaw

I scrutinize an argument offered by Marcus (2021) for the claim that ordinary human beliefs are necessarily known to their subjects.[1] I sharpen a worry raised by Marcus himself that threatens to limit his conclusion to the claim that we have not knowledge but merely a *capacity* to know our beliefs in virtue of having them. I suggest that in spite of the weakening, this limited conclusion is already quite powerful and, suitably framed, can accomplish much of the work needed of self-conscious knowledge of belief in Marcus's many applications of it.

1 Assertion, Avowal, and Self-Conscious Knowledge of Belief

Marcus argues using the following premises that if a subject S believes that p, they know that they believe that p (40):

(A) S believes that p only if S is able to honestly assert that p,

(B) S is able to honestly assert that p only if S is able to avow the belief that p,

(C) S is able to avow the belief that p only if S knows that she believes that p.

Three clarifications are in order.

First, beliefs in this argument are those mental states that we generally attribute to adult humans, in ordinary contexts, using the term "belief" (9). These states, Marcus claims, are not generally shared by animals, even if human and animal cognition are each species of some common genus of thought underlying the similarities between them

Draft of June 23, 2023; please cite published version
✉: James.A.R.Shaw@gmail.com

[1] Page citations are all to this work. Thanks to Ram Neta and Japa Pallikkathayil for comments on an earlier draft.

(32). Human belief is distinguished by a number of characteristic features, including that one knows one's own beliefs better than anyone else can, and in a different manner than they do (9, 39). The above argument aims to establish a particularly strong version of that thesis.

Second, ability talk in the argument refers to *voluntary power*—the capacity to perform an act at will. The control afforded by abilities in this sense may require favorable circumstances to manifest, and may be masked by external factors (e.g. distractions) or internal ones (e.g. psychological repression) that prevent the ability's realization (42). Neither the absence of favorable conditions nor the presence of masks destroys the ability—they merely block its manifestations.

Finally, to avow a belief is to state authoritatively, though not on the basis of observation or evidence, that one has the belief (48).

(A)'s defense consists largely in developing an applying a test for the presence of abilities. The defense of (B) rests on two ideas: that to assert is to put forward a proposition as believed; and that honest assertion is an intentional action, in which one has non-observational knowledge of what one is doing *as* asserting. While these premises raise many discussion-worthy issues, I wish to set them aside for now.

It is the defense of (C) that concerns me here. The argument here has two steps. First, drawing on earlier work that harmonizes with (without fully endorsing) John Hyman's theory of knowledge,[2] Marcus claims that knowledge that p is a necessary precondition of acting in light of the fact that p. To this claim, we add that the ability to avow p comes with the ability to act in light of a belief that p. In addition to its intuitive merit, this second claim is reinforced by the observation that avowal itself is clearly a form of acting in light of one's belief that p (51). From these two claims, it follows that the ability to avow requires an ability that, unmasked, comes alongside knowledge of the belief one avows.

I've formulated the conclusion of the argument of the foregoing paragraph cautiously: the ability to avow comes with an ability that, unmasked, requires knowledge. This is what is directly supported by the premises, and it is conspicuously weaker than (C). Marcus recognizes the gap here, framing the worry as follows (and crediting it to John Phillips):

> Is knowledge a necessary condition of *the ability* to act in light of p or

[2] Marcus (2012, ch.1), Hyman (1999).

> rather only a necessary condition of the *unmasked ability* to act in light of p. If only the latter, then believing that p would not, after all, be inseparable from knowledge that one holds the belief.

Marcus's reply to the concern is brief:[3]

(i) Knowledge that p is a necessary condition of acting in light of p.

(ii) "But if [so], it is (one would think) because p's eligibility to be a fact in light of which S acts is determined by whether or not S knows that p." (52)

(iii) "If that's right, then it is difficult to see what it could mean to have the ability to act in light of p despite not knowing that p." (52)

I worry that the last of these remarks is a non-sequitur. It might be helpful to compare a parallel line of reasoning:

(i′) Having some awareness of a bike is a necessary condition of intentionally riding the bike (under that description).

(ii′) This is because: the eligibility of any given bike for use in an intentional bike-riding is determined in part by whether or not one is aware of the bike.

(iii′) So: one cannot count as having an *ability* to ride a bike unless one is aware of the bike.

Even if one thinks (i′) is true (because without awareness of a bike, any bike-riding wouldn't be intentional under the relevant description), (iii′) is certainly false. Losing awareness of a bike does not destroy my *ability* to ride it.[4] One problem is perhaps that there are multiple readings of 'eligibility' in (ii′). On a strong reading, it essentially marks the presence of an ability in Marcus's strong sense. If so, (iii′) would follow from (ii′), but (ii′) would be false for the same reasons as (iii′). On a weaker and I suspect more natural reading, eligibility talk doesn't entail the presence of an ability. But then there is no special reason to think (iii′) follows.

---

[3] I've here restructured the text a little to get it into something like argument form.

[4] Intuitively, but especially in Marcus's sense. Cf. Marcus's example of the preservation of a surgeon's abilities in the absence of her tools (42).

The core problem is that it is all too easy to see what it could mean to have the ability to act in light of p despite not knowing that p. It could mean just what Marcus says: that the ability to act in light of p does not require knowledge—rather, what requires that knowledge is the ability *unmasked*. (Compare: we require the awareness of an available bike for the ability to ride that bike to be suitably *enabled*.)

We could of course try to repair the argument. But I want to flag that my worries go slightly beyond nitpicking about argumentative structure. When I look at the cases that would test Marcus's strongest claims about self-conscious knowledge of belief—those of belief suppression, including from self-deception—I find myself unable to accept the descriptions of them Marcus needs.

Marcus dramatizes an instance of self-deception with brothers Alfred and William who pursue careers as artists (43*ff.*). Alfred becomes an instant darling of the art world. William languishes in relative obscurity. But when Alfred becomes a recluse, galleries and critics take an unusual interest in William, ostensibly to gain insight into his brother's activities or whereabouts. Though William at some deeper level believes this is what is happening, he finds it irresistible to acquiesce in the more comforting view that the interest in his work is genuine, and merely occasioned by his brother's withdrawal from public view.

In this, or any other suitably described case of self-deception, it appears that the self-deceived agent is in a position to *learn* something about themselves (say, through therapy): that they have a latent, suppressed belief that subtly informs their actions. If Marcus is right, though, and belief necessarily comes alongside knowledge of it, this must be something the self-deceived already knows. But how can they learn what they already know?

Marcus responds to this concern by claiming that the selfsame forces that suppress belief in cases of self-deception likewise suppress, without destroying, concomitant knowledge of the belief (54). While I don't see anything incoherent about that picture, it does not strike me as a plausible description of most ordinary cases of self-deception. The description is just not intuitively apt: while we sometimes describe the self-deceived as having known all along on some level, we would not describe them as having known all along *that they knew* (even 'on some level'). The latter claim suggests they have arrived at the kind of self-awareness that would overcome their self-deception. And on a more theoretical level there is misalignment as well. For example, a suppressed belief characteristically leads to significant changes in behavior (e.g., perhaps William shows

a continued reluctance to share some of his artwork in high-stakes contexts, etc.). But in at least some cases of self-deception, there may be *no* such patterns of behavior that witness the putative bit of suppressed self-knowledge. What is the point of saying self-knowledge must be there, suppressed, if it need have no manifestations at all?

These informal concerns are hardly dispositive. But compounded by the lack of a compelling case for (C), they suggest that we step back and ask why we would want Marcus's stronger conclusion—that belief entails knowledge of it—to begin with.

2 Self-Consciousness Knowledge of Judgment

We don't have an argument for (C). But, for all I've said, we may have one for (C′).

(C′) If S is able to avow the belief that p then S has an ability which remains unmasked only if S knows that she believes p.

Together with (A) and (B), this would show that belief always comes alongside abilities which, absent interference, necessitate knowledge of one's belief. Properly framed, this result is already significant and can be set to important work.

I find it illuminating to reformulate this claim using an important bit of jargon from Marcus's book: that of a belief being *in mind*. Marcus gives several interrelated glosses and examples of this phenomenon. But underlying them is Marcus's sympathy with the idea that there is some paradigmatic form of believing p which necessarily precludes believing p's negation:

> [Beliefs in mind are] beliefs corresponding to facts (or purported facts) of which we have a standing awareness, one that excludes metaphysically—and not just normatively— awareness (or purported awareness) of incompatible facts. (25)

Marcus calls beliefs in mind in this sense *judgments* (26), a terminology I will also employ here.

At places, Marcus seems to identify the state of a belief's being repressed in a way that would mask an ability to avow it with that belief's being pushed out of mind. For example:

> [S]omeone who represses a belief cannot bring it to mind—that is what it is for them to repress it. Thus such a person cannot avow or assert the

> belief. Their ability to do so is masked by whatever explains the repression …(82)

As such, using the notion of a belief being in mind—a judgment—we can contrast a stronger form of self-conscious knowledge of belief sought by Marcus with a weaker form better supported by his arguments.

> Self-Conscious Knowledge of Belief
> Necessarily, beliefs are known to be believed.
>
> Self-Conscious Knowledge of Judgment
> Necessarily, judgments are known to be believed.[5]

Because all judgments are beliefs, but not vice-versa, the first of these claims entails the second, but not vice versa. If we take the recently stated connections between the unmasked ability to avow and being in mind seriously, then the argument for (C′) (with the help of (A) and (B)) gives us an argument for Self-Conscious Knowledge of Judgment. What might we accomplish with this weakened thesis?

It may seem that this immediately gives up on a critical idea for Marcus: that the connections between belief and knowledge of belief are not merely causal (52), and are manifestations of a single capacity (53). As Marcus puts it "[K]nowing what I believe is not a matter of having checked up on my psychological state." The attractiveness of this view is not limited to the unity of the self-conscious mind that it fosters, but also lies in its ability to stave off concerns about an infinite proliferation of distinct attitudes that might otherwise seem to be brought on by self-conscious knowledge of belief (18*f.*).

But the mere in-principle separability of belief and knowledge of belief does not force a causalist picture on us. In fact, it is far from clear that a causal picture could account for even the limited necessity of Self-Conscious Knowledge of Judgment: what about being in mind could *necessarily* preclude the interruption of a causal connection between two otherwise entirely distinct mental states? A much more plausible alternative is to say that it is inherent in paradigmatic belief to be known, but that repression warps this state, and in doing so warps this otherwise essential feature of it, destroying the knowledge that would ordinarily have to come alongside it. It can

[5] Triviality threatens if 'in mind' reduces to 'known'. But being in mind is not characterized in terms of knowledge, nor is there pressure in Marcus's framework to conceptualize it in those terms.

also be that a single capacity is responsible for belief and our (self-conscious) knowledge of it, as long as we think of the capacity's operation as warped in suppression as well. Re-establishing a proper connection between belief and knowledge is not a matter of getting evidence for the presence of one's own belief, but simply comes with the restoration of belief to its paradigmatic psychological and normative role.

So nothing stands in the way of having belief assume the standard roles it plays for Marcus *provided* it is in mind. And the roles of belief in *these* contexts, I want to suggest, are primarily what matter for Marcus's applications of self-conscious knowledge of belief. To see why, let's briefly take a look at one central application of Marcus's self-consciousness thesis: Moorean Paradox. (Though from this case, we will be able to extract a general lesson.)

Statements of the form "p and I believe not p" are generally problematic.[6] Marcus frames the explananda here as follows: When the first conjunct is asserted and the second *avowed*, the total statement is just as unintelligible as the statement of a contradiction—we can't make sense of the state of mind of the speaker alongside their honesty. But when the first conjunct is asserted and the second merely asserted (and not avowed), the assertion is an intelligible expression of irrationality. To get a sense for the idea behind the second claim, consider the following elaboration of a statement with Moore-paradoxical form: "'The subway is safe; yet (it seems) I don't believe it: after all, I spare no expense to avoid taking it." (64) This individual has some kind of alienation from their beliefs that may well be irrational, but we can at least get some grip on the state of mind they are trying to convey.

Marcus's account of the central, unintelligible Moore-paradoxical statements takes the following shape, which appeals directly to his thesis about self-conscious knowledge of belief:

> Because belief and knowledge of belief are inseparable, assertion, which everyone understands as expressive of (first-order) belief, is also expressive of (an inseparable) knowledge of belief; and belief-avowal, which everyone understands as expressive of doxastic knowledge, is also expressive of (the corresponding inseparable first-order) belief. Moore-Paradoxical utterances are absurd because they purport to express the mind of someone who knowingly embraces a contradiction, which is impossible. Moore's

[6] I'll focus on the 'commissive' version of paradox.

> Paradox is resolved. (82)

This description disguises that it is not the *mere* inseparability of belief and knowledge which does key work here, but their inseparability alongside the fact that the belief expressed by a first conjunct, and that avowed by the second, are both *in mind*. This comes out when Marcus accounts for Moore-paradoxical statements by those who do not avow the second conjunct and so form higher-order beliefs merely on the basis of self-observation or other evidence. Won't the account of the foregoing quotation predict these statements are also unintelligible, violating Marcus's own characterization of the Moorean explananda? Marcus replies:

> [A]vowing and asserting a belief entails bringing it to mind …someone who represses a belief cannot bring it to mind … Thus such a person cannot avow or assert the belief. Their ability to do so is masked … But there is no bar to such a person ascribing the belief to themselves on the basis of evidence … They could express doxastic self-knowledge in one conjunct (the assertion, which pleases them), but not the other (the evidential self-ascription, the painfulness of which is buffered by alienation). Such a statement would reveal irrationality but it would be intelligible. (82–83)

Because all beliefs are self-consciously known for Marcus, any belief truly self-ascribed in the second conjunct (and we can suppose for now it is truly self-ascribed) is one that is known to the subject just as much as any other belief. But because the belief is not in mind, it cannot be expressed in the relevant sense. Marcus takes this to block the previous line of argumentation that led to unintelligibility. Of course, nothing about the problem here requires that the belief be known to the subject—on the contrary, the suppression of the belief which takes it 'out of mind' negates any interest that knowledge would have.

Putting all this together, we see that self-conscious knowledge of belief plays a role in the two foregoing explanations when, and only when, it is paired with that belief's being in mind. Accordingly, if Marcus's account using Self-Conscious Knowledge of Belief contributes to a satisfying an explanation of Moorean Paradox (of both forms), so could an account using Self-Conscious Knowledge of Judgment, and for exactly the same reasons.

This feature of Self-Conscious Knowledge of Judgment does not seem accidental, or particular to this case. Here is why. Self-Conscious Knowledge of

Judgment differs from the thesis Marcus endorses only for suppressed belief. In such cases, Marcus will admit that even if knowledge of belief is present, it is out of mind in just the way that the suppressed belief is. Indeed, intuitively it is even 'further' out of mind than the belief, for the reasons I gave in expressing my concerns for the existence of this hidden form of knowledge: any such knowledge often seems to play no role whatsoever in explaining the activities of the agent who possesses it. With its efficacy so diminished, self-knowledge in these contexts cannot be drawn on to do any substantial explanatory work. Marcus is rightfully mindful of this in his many applications (including: to impossible contradictory beliefs, Moorean Paradox, a variant of Moorean Paradox for inference, etc.), which do not trade on giving a significant role to self-knowledge in the case of repressed belief—and indeed are persistently qualified to avoid giving it such a role. But this of course means that such self-knowledge is itself in-principle dispensable in those applications.

The role of self-consciousness knowledge in contexts of repression appears largely ornamental, giving a pleasing uniformity to the characterization of belief. I am not suggesting that that there is anything inherently wrong with thinking in these terms, of course, should we have independent grounds for embracing them. The concern of §1, however, was precisely that Marcus's arguments are unable to establish that beliefs are necessarily known to be believed so as to give belief the desired uniformity.

What I've tried to do in this section is to suggest that the the presence of self-knowledge for belief generally not only is more than Marcus establishes, and is more than is intuitively desirable, but is also more than Marcus really needs. It is enough for the self-consciousness knowledge of belief to belong exclusively to *paradigmatic* belief—to judgment. This should safeguard any intuitively plausible application of the idea that beliefs are marked by the presence of self-conscious knowledge. Moreover, if the shift to the weakened, but still powerful thesis of Self-Conscious Knowledge of Judgment is accepted, Marcus will still have made several steps forward in helping us see how we could, and why we should, accept it.

References

Hyman, John. 1999. "How Knowledge Works." *Philosophical Quarterly*, vol. 49 (197): 433–451.
Marcus, Eric. 2012. *Rational Causation*. Harvard University Press.

—. 2021. *Belief, Inference, and the Self-Conscious Mind*. Oxford, United Kingdom: Oxford University Press.