

Understanding Baseless Self-Knowledge: Boyle on Reflection and Rationality

Johannes ROESSLER (Warwick)

Abstract: Matthew Boyle's *Transparency and Reflection* articulates and defends what might be called a 'transformative' view of self-awareness. The human capacity for self-awareness is part of the very nature of our minds, not added to a form of mentality we share with subjects lacking that capacity. Yet, in his recent work Boyle introduces a striking 'additive' element into the overall transformative picture. While 'non-positional' self-awareness is an essential aspect of 'human mindedness', the capacity for reflection is a further, additional accomplishment. I raise some questions about this additive element — in particular, about its compatibility with a reflectivist account of self-knowledge, and about its motivation.

Keywords: self-knowledge, reflection, rationality, psychological capacities

Abstract: In seinem Buch *Transparency and Reflection* entwickelt Matthew Boyle eine 'transformative' Theorie des Selbstbewußtseins. Selbstbewußtsein ist ein wesentlicher Aspekt unserer mentalen Zustände— nicht etwas das zu einer Psychologie hinzukäme, die wir mit nicht-selbstbewußten Subjekten teilen. In seinen jüngsten Veröffentlichungen hat Boyle dieses transformative Model allerdings um eine interessante 'additive' Dimension erweitert: während 'nicht-positionelles' Selbstbewußtsein unseren mentalen Zuständen wesentlich ist, wird die Fähigkeit zur Reflexion als zusätzliche Leistung verstanden. Der Artikel konzentriert sich auf zwei kritische Fragen: eine zur Kohärenz und eine zur Motivation des additiven Elements der Boyleschen Theorie.

Keywords: Selbstbewußtsein, Reflexion, Rationalität, psychologische Fähigkeiten

In *Transparency and Reflection*, Matthew Boyle argues that the human capacity for self-knowledge is bound to remain mysterious so long as it is treated as just another topic in epistemology. Drawing on, and seeking to revive, a tradition from Plato to Sartre, he argues that self-awareness is an essential feature of 'human mindedness'. It "informs our capacities

for knowledge and action”¹, and it “transforms the nature of our cognition in general”.² Understanding first-person self-knowledge, on Boyle’s view, requires understanding the sense in which our attitudes and experiences are inherently self-conscious. Self-knowledge is primarily a subject matter for the philosophy of mind.

There is so much I admire in, and have learned from, Boyle’s work that I wonder if I’m properly qualified to contribute to what the *Philosophische Jahrbuch* calls a *Kontroverse* (variously translated as an ‘intense dispute’ or a ‘heated discussion’). I am not sure there is anything in Boyle’s work I would intensely dispute. There is much I intensely agree with. I share Boyle’s dissatisfaction with the dominant ‘epistemologically oriented’ approach to self-knowledge, and I find his case for a ‘reflectivist’ alternative compelling and extremely illuminating. I am also struck and attracted by Boyle’s suggestion that reflectivists would do well to revive and develop the notion of ‘implicit self-awareness’. There are many things I would have liked to discuss here, for example Boyle’s diagnosis of certain challenges to the authority of our ordinary commonsensical view of ourselves as rational thinkers, or his Sartrean reflections on bodily awareness. For reasons of space I’ll focus on a single topic, a specific commitment of Boyle’s version of reflectivism. I will call this a non-conceptualist view of ‘implicit self-awareness’, or non-conceptualism for short. On that view, the ‘implicit’ awareness we are said to have of our attitudes and actions is peculiarly primitive. It provides a basis (in some sense) for reflective self-knowledge but it does not itself implicate the sorts of conceptual capacities required for such knowledge. For example, to say that S is implicitly aware of believing that p does not commit us to the idea that S has mastered the concept of belief.

I start with an initial worry about non-conceptualism: there seems to be a risk of making reflectivism collapse into a version of the ‘epistemologically oriented’ approach it is intended to replace. I go on to suggest that non-conceptualism is in any case not mandatory for reflectivists (section 2). What is the motivation for non-conceptualism? Boyle makes much of the possibility that a thinker may be rational without being reflective. But I think non-conceptualism also reflects a certain explanatory ambition – a particular view of the explanatory role implicit self-awareness needs to play in a reflectivist account. Trying to be as *kontrovers* as possible, I will raise some doubts about each of these motivations. (Sections 3 and 4)

¹ Boyle (2024), 2. The target of my contribution is of course Boyle’s article in this journal (Boyle, 2023) but I will sometimes also refer to Boyle’s more detailed treatment of the issues in his 2024 book.

² Boyle (2023), 128.

1. Speaking to the believer herself

Suppose we ask an expert on international affairs whether she believes there will be a third world war any time soon, and she replies in the affirmative. Next, suppose that we add the following, more leftfield question: ‘How do you know you believe this?’ In real life, our pundit would no doubt give us a dirty look and get on with more important matters. (I will come back to the issue of the oddity of our question.) But let’s imagine that, wishing to be helpful, she replies along the following lines: ‘My answer to the first-order question, whether there will be a third world war, is ‘yes’. Since our first-order judgements usually express what we believe about the relevant matter, my judgement provides good evidence that I believe there will be a third world war soon. That’s how I know.’

In a marvellous passage, Boyle registers the following reaction to this way of thinking about self-knowledge: “If I received this sort of report on someone’s beliefs, I should demand to speak to the believer herself, not just to her biographer, however well-informed she might be.”³ The trouble with the expert’s account is that it represents her as ‘alienated’ from her belief. By regarding the question of her belief as an empirical matter in its own right, she disables herself from using the self-ascriptions of her belief as an expression of that very belief, depriving her self-ascription of the significance we would ordinarily attach to such claims. (As Boyle’s reaction illustrates, her account is also apt to alienate her audience.) I won’t go into the details of Boyle’s argument here. I think he is right about this point, and he is also right about its implications for epistemological theorizing about self-knowledge. On Boyle’s diagnosis, to seek an explanation of self-knowledge in terms of an ‘epistemic basis’ that enables us to find out about our mental states is to falsify the explanandum. Importantly, the point is neutral on whether the explanation is thought to corroborate or to debunk traditional claims about ‘privileged access’. The distortion lies in the very idea that what we have is *access*. That idea builds a form of “alienation into the very structure of [first-person] self-knowledge”⁴.

How is self-knowledge to be made intelligible, if not by reference to some mode of access to the facts about our mental lives? Reflectivism’s core idea, as formulated by Boyle in previous work, is this: knowledge of (e.g.) one’s current beliefs is to be explained “not by appeal to some mechanism or method that allows the subject to know an otherwise unknown fact about

³ Boyle (2024) 59.

⁴ Boyle (2015), 344. The quoted passage is a diagnosis, specifically, of inferential accounts of self-knowledge, but I think it may be used without distortion to capture Boyle’s general critique of the ‘epistemically oriented’ approach.

himself, but in terms of the nature of belief itself”.⁵ Philosophical reflection on the nature of belief reveals that human beliefs are inherently self-conscious. In the light of this, it is unsurprising not just that we ordinarily know what we believe, but also that the request for an ‘epistemic basis’ or mode of access (‘How do you know?’) would be off-key. In his more recent work Boyle adds an important refinement, which turns on the notion of ‘implicit self-awareness’ (henceforth ISA). Our overall picture of self-knowledge, he suggests, should be in two parts. The first part provides an account of the inherently self-conscious nature of human beliefs and other attitudes, explicating the sense in which, in holding an attitude, we are ‘implicitly’ aware of holding it. The second part elucidates the way ISA grounds ‘reflective’ or ‘explicit’ self-knowledge – the kind of knowledge we express when we answer questions about our attitudes.

Boyle stresses the Sartrean inspiration behind his account of ISA. However, whereas Sartre’s discussion of ‘non-positional consciousness’ strikes many readers as suggestive and mystifying in equal measure, Boyle makes a sustained attempt to demystify ISA. Central to his account is the idea that ‘human mindedness’ is essentially discursive. Attitudes such as beliefs or desires, in the case of ‘rational animals’, represent answers to questions the subject herself is able to understand, consider and respond to; they are “attitudes towards questions”.⁶ In turn, Boyle argues that answering a question is a self-conscious act. In judging that *p*, for instance, a subject is implicitly aware of believing that *p*.

Once again, I am inclined to agree with all of this. The question I would like to raise concerns, specifically, Boyle’s non-conceptualist account of ISA. Attributions of ISA are supposed to have no implications for the subject’s mastery of relevant psychological concepts. The problem I see with this emerges when we ask about the relationship between ISA and propositional (or factual) knowledge. It’s useful to distinguish two steps here. First, it seems plausible that under a non-conceptualist construal, ISA needs to be distinguished from (and does not entail possession of) propositional knowledge. Second, if that is right, it can be hard to see how ISA can be anything other than an epistemic basis for acquiring propositional knowledge.

The first step is relatively straightforward. Suppose *S* judges that *p* and, in doing so, enjoys an implicit awareness of believing that *p* (very roughly, she treats the question of whether *p* as “closed”, where this means that her “own belief on the question is settled”⁷). Does

⁵ Boyle (2011), 228.

⁶ Boyle (2024), 193.

⁷ Boyle (2024), 72 f.

she therefore *know that she believes that p*? Under Boyle’s account, surely not. While S will be a ‘competent deliberator’ (ibid.), she may not have mastered the concept of belief; and it seems to be a truth more or less universally acknowledged that propositional knowledge *does* require possession of relevant conceptual capacities.

For the moment, suppose that’s correct. Now for the second step. An initial, and perhaps superficial, cause for concern is that Boyle uses terminology that is redolent of traditional ‘epistemically oriented’ explanations of self-knowledge. ISA is said to provide a “conscious basis” for explicit beliefs about our mental states.⁸ In acquiring such beliefs the subject is said to make a cognitive “transition” from one sort of state to another.⁹ Indeed, ISA is said to “justify” and “warrant” explicit self-ascriptions of the relevant states.¹⁰ A nervous reader may begin to worry whether ISA is not precisely the sort of thing that gets in the way of speaking ‘to the believer herself’ – some sort of basis for forming second-order beliefs, by exploiting which we can find out what we believe. But there is a reassuring answer. What the nervous reader overlooks is a vital qualification. ISA is not a basis for acquiring self-knowledge, merely a basis for acquiring (and retaining) “*explicit* self-knowledge” or “self-knowledge *proper*”.¹¹ Boyle makes it quite clear that when a subject exploits that sort of “basis” she is not “acquiring a new piece of information” but simply “reflecting on *what she already knows*” (ibid., my emphasis). Reflection does not amount to a ‘step from ignorance to knowledge’ or to the formation of a “new belief”.¹² In short, ISA is not an ‘epistemic basis’.

Now we come to the crux of the matter. The ‘already knows’ response would indeed be reassuring if ISA entailed possession of propositional knowledge. We could then put the nervous reader’s mind to rest by saying that in self-ascribing the belief that *p*, S is simply reflecting on something she already (implicitly) knows, viz. *that she believes that p*. However, since ISA does not entail possession of propositional knowledge, we have no right to say this. S only has propositional knowledge about her own mental states once she acquires what Boyle calls ‘explicit’ self-knowledge or self-knowledge ‘proper’. That makes it hard to see how a basis for self-knowledge ‘proper’ could be anything other than *a basis for propositional knowledge*. Far from being off-key, the question ‘How does S know that she believes that *p*?’ would have a perfectly good answer. Recall our expert on international affairs. On Boyle’s account, it seems, she could insightfully explain her self-knowledge in this way: ‘well, my

⁸ Boyle (2023), 120.

⁹ cf. Boyle (2024), 165 ff.

¹⁰ cf. Boyle (2024), 74, 194.

¹¹ Boyle (2024), 99, my emphases.

¹² Boyle (2024), 185.

conscious basis for believing that I believe there will be a third world war is the non-positional or implicit awareness of believing this which I have in judging that there will be a third world war. That awareness warrants or justifies my second-order belief and thus accounts for my self-knowledge.’

Several possible replies to this worry would be worth considering, though I won’t be able to do so here.¹³ Let me just restate the difficulty, in the form of a dilemma. What I called the ‘already knows’ response requires that ISA is a form of knowledge. But what sort of knowledge? If it is propositional knowledge, we face the question of how lack of relevant conceptual capacities can be squared with possession of propositional knowledge. If it is some other form of knowledge (say, some kind of acquaintance¹⁴), we need to explain why ISA does not, after all, amount to an ‘epistemic basis’ for acquiring propositional knowledge (causing the two-stage reflectivist account to collapse into a version of the epistemic approach).

Perhaps there is a way around the dilemma. Still, in view of it, the question seems worthwhile whether non-conceptualism is really essential for the reflectivist project. Specifically, could reflectivists invoke the idea of implicit self-awareness without subscribing to a non-conceptualist account of it? In the next section, I sketch an affirmative answer to that question.

2. Modest reflectivism

¹³ For what it’s worth, the most promising response, to my mind, would be to reformulate non-conceptualism, by distinguishing between implicit and explicit forms of concept possession. Consider Michael Dummett’s (tentative) suggestion that a speaker who “always uses the pairs ‘I’/‘me’, ‘he’/‘him’, ‘she’/‘her’, and ‘who’/‘whom’ correctly” but has never “been taught the rudiments of formal grammar” may be said to “have an implicit grasp of the concepts” *nominative* and *accusative* (Dummett 1991, 96). *Mutatis mutandis*, a “competent deliberator” who is “intelligently responsive” (Boyle 2024, 98) to her believing that p might be said to have an implicit grasp of the concept of belief, even if she has never been taught the rudiments of commonsense psychology and is not able to answer questions about her own or others’ beliefs. We would then be able to insist that ISA entails possession of propositional knowledge without denying that propositional knowledge implicates conceptual capacities: ISA, we would maintain, is sustained by an ‘implicit grasp’ of the requisite psychological concepts. Perhaps this is not far from what Boyle has in mind. While he does not work with the distinction between an implicit and explicit grasp of concepts, he does associate ISA with an “implicit understanding of our own cognitive capacities” (Boyle 2024, 209), characterizing this as a “real species of understanding” (ibid.). Nevertheless, I am not sure the proposal would be acceptable to him. Not only is Boyle adamant that ISA “does not require the application of a psychological concept to a first-person subject” (Boyle 2024, 19), he offers an account of concept acquisition in which ISA figures as a *basis* for “framing” concepts of various cognitive activities (Boyle 2024, 210) –and so precisely not as a matter of implicitly grasping such concepts. Indeed, would it not go against the spirit of Sartrean ‘non-positional consciousness’ to suggest that this is merely a label for ‘implicit’ *positional consciousness*? In any case, we may have qualms about Dummett’s suggestion regarding grammatical knowledge; and further qualms about the analogy between grammar and psychology.

¹⁴ Note that Boyle usually characterizes ISA in ‘objectual’ terms, for example as “an awareness of our own representational states” (Boyle 2024, 20). However, some formulations encourage a ‘factive’ construal. For example, one’s implicit awareness of M is explicated in terms of one’s “intelligent responsiveness to what is in fact *one’s being in M*”. (Boyle 2024, 98, my emphasis)

A significant part of the attraction of Sartre's notion of 'non-positional consciousness' lies in its intuitive force as a 'phenomenological' description of our mental lives. Consider a variation on an example of Sartre's. Suppose you are fiddling with your cigarettes and, observing this, we ask you why. Without hesitation, you tell us 'I am counting my cigarettes.' Prior to our question, your attention may have been absorbed by the cigarettes. You may not have thought about yourself and your activity at all. Still, you can answer the question straight away, and in doing so, it is natural to think, you articulate something of which you were already, 'pre-reflectively', aware. That sort of chronicle seems compelling. But how should we understand the idea of 'pre-reflective' self-awareness? Here are two initial suggestions.

- (a) It's not obvious that the intuitive force of Sartre's example demands or supports a non-conceptualist construal of ISA. It seems natural to think that you knew what you were doing (you knew *that* you were counting your cigarettes) even before being compelled by our intervention to reflect on the matter. That knowledge would not have been available to you if you didn't have the requisite concepts. Indeed there is surely a cogent reason for attributing (pre-reflective) propositional knowledge to you. As Anscombe argued, our 'reason-seeking' question 'why?' presupposes that you know what you are doing (under relevant descriptions), in a distinctive ('non-observational') sort of way.
- (b) Nor is it clear that the phenomenology bears out the idea that ISA serves as the 'conscious basis' for explicit self-ascriptions. To answer our question, it seems, you don't need to form a belief (or 'explicit belief') on any kind of basis. All you need to do is *articulate* the content of the intention that informs your activity. An expression of your intention, in the form of first-person present-progressive self-ascription of some activity, will, under certain conditions, simultaneously express knowledge of what you are doing.¹⁵ And your intention is already conceptually structured. So, reflection will often just be a matter of putting your intention into words. Still, in the case of an intention acquired as a matter of habit or anyway without explicit reasoning, articulating its content may not be a trivial task. It may require making explicit something of which you only had a vague or dim sense.¹⁶ Importantly, there may be factors that impede or even disable the articulation of

¹⁵ Two key conditions are that the intention is informed by a realistic, knowledgeable conception of one's practical capacities and opportunities, and that its execution involves no 'mistake in performance'. See Anscombe (1957), esp. § 32.

¹⁶ Compare Charles Taylor's account of what he calls 'agent's knowledge': "[A]s agents, we will already have some sense, however dim, inarticulate or subliminal, of what we are doing; otherwise, we could not speak of directing at all. So agent's knowledge is a matter of bringing this sense to formulation, articulation or full consciousness.", Taylor (1985), 80.

an intention; for example, broadly motivational factors (say, you may be disinclined to think about the matter) or the subject's being in a state of consciousness that is not conducive to rational thought.¹⁷

Suppose we drop the idea that ISA implicates no psychological concepts. Then what would it mean to say that you were implicitly aware of counting your cigarettes? A natural suggestion, it seems to me, is simply this: while you did not think about what you were doing, you were *able* to do so – specifically, able to express knowledge of what you were doing – ‘just like that’, without relying on observation or any other mode of discovery. In a word: your activity was potentially reflective. What this means, of course, is not just that it was possible that you should reflect on your activity, but that you had the requisite capacities, and they were not incapacitated in some way. The explanatory role of ISA, on this picture, would be more modest than on Boyle's account. ISA is not a state of awareness that provides a ‘conscious basis’ for explicit self-ascriptions. Instead, it is to be understood in terms of the subject's possession of certain capacities. But note that the overall account would still be reflectivist in Boyle's sense. On an Anscombean view of ‘practical knowledge’, the capacities whose exercise accounts for our explicit self-knowledge are elements or aspects of the capacity for intentional agency itself: centrally, practical reason, or the capacity to form and execute realistic intentions (intentions informed by a proper assessment of one's practical abilities: see footnote 15).

In the case of intentional activities, then, there seems to be some initial promise in a version of reflectivism that invokes the notion of ‘implicit self-awareness’ but steers clear of non-conceptualism. I'll call this sort of view ‘modest reflectivism’. Consider now the case of self-ascriptions of belief. Modest reflectivists will agree with Boyle that believing, in the case of rational thinkers, involves the capacity to judge or to answer questions, and that in judging that *p* one is implicitly aware of believing that *p*. They will suggest, though, that ‘implicitly aware’ here means something like ‘potentially reflective’. Someone who has fully mastered the capacity to judge will be able to reflect on the attitudes expressed by her judgements. One way to motivate that thought is to argue that the capacity to judge requires an understanding of the normative import of a judgement. In judging that *p* one closes the question whether *p*, as Boyle puts it. Yet, at the same time, one lays oneself open to normative questions probing one's entitlement to judge that *p*. For example: ‘How do you know that *p*?’ or ‘Why do you believe

¹⁷ See my 2013 for slightly more detailed discussion of relevant ‘disabling conditions’. I think some of the points Boyle makes in his illuminating reflections on the question of the value of explicit self-knowledge (Boyle 2024, ch. 10) are related to this.

that p?’ Someone who has fully mastered the capacity to judge will not only be able, in their judgements, to be responsive to reasons, as they see them; they will also be able to invoke their reasons in response to normative questions. In this way, the capacity to believe may be said to be internally connected with the capacity to know one’s beliefs and one’s reasons. A rational thinker’s beliefs are ‘potentially reflective’.¹⁸

There are of course a number of ways in which this sketch would need to be filled out, and there are questions as to whether it can be extended to first-person knowledge of other sorts of attitudes. I present it here merely as a rough indication of a possible way to develop the basic reflectivist idea, and as a background for examining the case for non-conceptualism about ISA. I think there are two sorts of reasons why Boyle would tend to be sceptical about a modest version of reflectivism. One has to do with the explanatory role of ISA. In Boyle’s account, ISA does explanatory and justificatory work that cannot be matched by modest reflectivism. Second, an account on which ‘implicitly aware’ means ‘potentially reflective’ is plainly in the target area of Boyle’s critique of ‘over-intellectualized’ accounts of human rationality. I consider the two challenges in turn.

3. Reflectivist explanations

One source of scepticism about modest reflectivism might come from misgivings about capacities, or more specifically about explanations of knowledge in terms of the exercise of capacities. Does a modest reflectivist have anything more illuminating to say than that S knows that she believes that p because she has the capacity to know that sort of thing? Briefly, there are three points to make in response to that general worry.

First, a modest reflectivist should make it clear that the ‘capacity for self-knowledge’ is intelligible in terms of underlying capacities; such as, in the case of intentional activities, the capacity to form, and express, realistic intentions, or in the case of belief, the capacity to express one’s beliefs by making judgements. In the light of further reflection on what it means (for example) fully to master the capacity to judge we can see that someone who exercises that capacity will be aware of what they are doing in judging that p, viz. expressing their view that p. Or so the modest reflectivist will argue. S’s self-knowledge is intelligible, not in terms of a bare capacity for self-knowledge, but in terms of an interconnected set of abilities implicated in believing that p. Second, one way in which that explanatory schema can be illuminating is

¹⁸ This echoes some formulations in McDowell’s account of perceptual belief: see McDowell (2011), esp. 17–44. For (in effect) a modest reflectivist perspective on first-person knowledge of belief and of propositional perception, see my (2011) and (in press), respectively. (The latter is influenced by Boyle’s Sartrean turn.).

in shedding light on why someone may *not* be aware of their attitudes in the way they would ordinarily be expected to be. For example, motivation or a disordered state of mind (or both) might impede S exercising her capacity to express her belief that p and disable her from knowing that she believes that p. Third, the modest-reflectivist analysis may shed light on what is wrong with the quest for an ‘epistemic’ explanation of self-knowledge. Recall Boyle’s suggestion that the idea of first-person *access* builds a form of ‘alienation into the very structure of [first-person] self-knowledge.’ Now, modest reflectivists might grant that in a particular case, one may stand back from a judgement, wondering whether it articulates one’s real view of the matter. They would insist, however, that to do so is to wonder whether one’s attempted exercise of the capacity to judge was in fact defective; whether one’s attempt at exercising that capacity was fully successful. And they will make the same point about a request for an explanation of how we find out what we believe (‘How do you know you believe there will be a third world war?’): the request only make sense if the subject’s attempt at exercising the capacity to judge is less than fully successful (a fully successful attempt would implicate *baseless* knowledge of what she believes). It’s here that we may begin to see why the request would ordinarily be considered odd and off-key (and indeed perhaps insulting: it would amount to questioning something like the integrity of one’s rational capacities). Furthermore, a completely general such request (‘What’s the epistemic basis for first-person knowledge of beliefs?’) would represent us as alienated from our judgements across the board. A modest reflectivist may ask whether that request would even be coherent; whether it would not amount to simultaneously attributing beliefs to us and denying that we have the capacities whose exercise makes it possible for us to have beliefs.

Suppose, then, that putting capacities centre-stage would not necessarily make reflectivism unilluminating. There are two more specific explanatory demands, however, that, so it may be argued, only a non-conceptualist account of ISA can satisfy. There is the demand for an account of the justification of first-person self-ascriptions. And there is what is sometimes called the ‘problem of transparency’: how can it make sense to answer questions about our own beliefs by making judgements about the world? Both explanatory projects play central – and I think connected – roles in Boyle’s version of reflectivism. We can bring out the connection between them by observing that Boyle frames the ‘problem of transparency’ as a problem about our basis for first-person self-ascriptions. The problem is that “nothing in my apparent *basis* [e.g. my basis for judging that I believe that p] seems to supply a *ground* for

[that judgement]”¹⁹ (since the truth of *p* usually neither entails nor provides evidence that I believe that *p*). Boyle’s account of ISA is intended to provide a solution to the ‘basis-but-no-ground’ problem. He writes: “What justifies our reflective self-ascription is a preexisting awareness of the relevant mental state [...]”²⁰ In the light of Boyle’s non-conceptualist theory of ISA, judging that *p* can after all be seen to provide me with a justifying ‘ground’ for forming the ‘explicit belief’ that I believe that *p*. In turn, that point delivers a solution to the ‘problem of transparency’: the justificatory role of ISA makes it reasonable to answer questions about our beliefs by addressing questions about the world.²¹

I think a modest reflectivist should concede that her account of ISA does not make for a neat solution to the ‘basis-but-no-ground problem’. What she should question is whether reflectivists should even try to *solve* that problem, as opposed to dissolving or deconstructing it. Why should we assume that first-person reflective self-ascriptions (and the ‘explicit beliefs’ they express) have any kind of basis? After all, reflectivists (of any stripe) hold that first-person self-knowledge is ‘baseless’: it is to be explained by reference to the ‘nature of human mindedness’, not in terms of any epistemic basis. Then why should we nevertheless think of a first-order judgement as a ‘basis’ for self-ascriptions? I am not sure I fully understand Boyle’s thinking about this question, but I can think of two possible answers. First, it might be said that we have to provide an account of what justifies *S*’s ‘explicit belief’ that she believes that *p*, and that to this end, we need to identify *S*’s ‘basis’ for forming that ‘explicit belief’. Second, it might be said that the notion of a basis is implicit in the very idea of ‘transparency’. Roughly speaking: to say that we answer questions about our beliefs by answering questions about the world just is to say that our answers to the first sort of question are in some sense ‘based’ on our answers to the latter sort of question.

A modest reflectivist will seek to counter these considerations by proposing a diagnosis: they reflect a ‘belief-first’ approach to self-knowledge that we have no reason to accept – and that reflectivists, in fact, have reason to *reject*. Re justification: why should reflectivists not reverse the order of explanation that has been treated as orthodox by proponents of an ‘epistemically oriented’ approach to self-knowledge? Why not say that what justifies *S*’s belief that she believes that *p* is that she knows she believes that *p*? (Quite generally, surely, if you

¹⁹ Boyle (2024), 38, my emphases.

²⁰ Boyle (2024), 74. Or again: the “implicit understanding of my own representational state” I have in representing the question whether *p* as “closed” “warrants me, on reflection, in explicitly ascribing a corresponding representational state to myself.” (Boyle 2024, 194)

²¹ “[I]t is our possession of nonpositional consciousness of our own mental states that warrants us in treating questions about our own mental states as transparent to corresponding questions about the non-mental world.” (Boyle 2024, 99)

know that p, it's OK to believe that p.²²) That would allow us to understand the justification of S's (second-order) belief without appeal to any 'basis', and indeed in a way that brings out why the belief has no basis (viz. the knowledge that justifies it is 'baseless'). Re transparency: the observation that we treat one question as 'transparent' to another is surely neutral on exactly how the first question gets answered. One possibility might simply be that *in answering* the second question we will be aware of – we will know – the answer to the first question. There is no mention here of a basis for the second-order belief.

The widespread tendency to pose the problem of transparency as a problem about the basis for self-ascription, I think, goes back to Moran's influential account of the 'problem of transparency'. Here is a representative passage:

“The immediate problem with the Transparency claim is that it seems to say that I can answer a question about the beliefs of a particular person [specifically, about my view on the likelihood of a third world war], which would seem to be an empirical psychological question, not by considering the evidence about that person, but by considering the evidence about some wholly distinct matter of fact, [...], evidence about the increasing militarization of world politics. This is a real problem.”²³

Although Moran is not explicit about this, it is natural to come away from this passage with the impression that we form second-order beliefs on a very peculiar basis (viz. evidence for the relevant first-order belief). Again, by characterizing the question of what I believe as an 'empirical psychological question', he encourages the expectation that the answer should be underwritten by relevant evidence or some sort of epistemic basis. But Moran offers no support for these assumptions. A modest reflectivist will argue that they reflect the tenacity of the traditional 'epistemically oriented' approach to self-knowledge.²⁴

Then how will the modest reflectivist account for the point (or reasonableness) of the practice of 'transparency'? One suggestion, as already indicated, is simply that it makes sense to address the question whether I believe that p by addressing the question whether p because in answering the latter question I will know the answer to the former. But I think there is

²² See Williamson (2007), 357 ff.

²³ Moran (2003), 405.

²⁴ Yet, it might be asked, isn't the whole point of Moran's account of the role of the 'deliberative stance' in self-knowledge precisely to dislodge that approach to self-knowledge? It is true that Moran opposes traditional introspectionist accounts of self-knowledge. Yet, his own account (as I understand it) turns on an explanation of our *entitlement* to second-order beliefs acquired by the use of the 'transparency procedure'. (See esp. Moran 2003, 405) Thus, while he certainly aims to dislodge various versions of the 'epistemically oriented' approach, his conception of (and solution to) the 'problem of transparency' nevertheless seems to be framed in terms that derive from that very approach. See my 2013 for a more detailed discussion and critique of Moran's account.

something else a modest reflectivist should say. Moran characterizes the first-person question of what I believe as ‘what would seem to be an empirical psychological question’, and it clear from the context that ‘psychological’ here means something like ‘narrowly psychological’ – concerned merely with the way someone represents things as being rather than with the way things are. Of course, it is undeniable that believing is not a factive attitude. Believing that p does not entail that p. Still, there may be a good reason why we do not ordinarily conceive of the question of my beliefs as ‘narrowly psychological’. To bring this out, recall our political pundit. Suppose we direct another leftfield question to her: ‘If we understood your thought processes well, you seem to have answered our question as to whether you believe there will be a third world war soon by thinking, not about your psychology, but about international affairs. How can that make sense, given that our question really was about psychology?’ Suppose that once again, the expert is patient and tries to be helpful. One thing she might say is this: ‘I took it you were conducting a political interview, not a psychiatric assessment. If so, then surely your question was not *merely* about my psychology. In effect, I was invited to *testify*. Ideally, I would have been able to share my knowledge regarding the likelihood of war. Even if knowledge may be to ask too much in this area, you were presumably interested in an expert opinion – a view you would have reason to accept or at least to consider seriously in your own thinking about the matter. The question I needed to ponder, therefore, was whether I *know* if there is going to be a third world war soon, and if not, whether I can nevertheless offer a well-informed view on the subject. Either way, the question is inextricably entwined with international affairs.’²⁵

We might summarize the point borrowing Boyle’s terms: to think of the question of what I believe as (narrowly) ‘psychological’ is to build a form of ‘alienation into the very structure’ of our understanding of transparency. In the light of all this, we can see that it is at least far from obvious that reflectivists should try to solve the ‘basis-but-no-ground problem’. Instead, they may seek to dissolve it, by pushing back against assumptions that create the impression that transparency presents us with (as Moran puts it) a ‘real problem’.

4. Rationality and psychological concepts

²⁵ A fuller discussion of this point would benefit from reflection on the following remark by Edward Craig (which might be heard as a novel gloss on the notion of a first-person perspective): “One great advantage of using yourself as an informant, of course, is that unlike everybody else you are always around; and in any case you are sure to be the first person you ask.”, Craig (1990), 64.

But is the modest reflectivist not guilty of overintellectualizing human psychology? Is the idea that believing/intending/perceiving is inherently open to self-knowledge not implausible, in view of the surely undeniable fact that there are believers, intenders and perceivers who lack the conceptual wherewithal to reflect on their beliefs or intentions or perception?

Boyle's discussion of these concerns is rich and subtle, and I cannot hope to engage properly with the full range of considerations he marshals. I would like to identify what I take to be two central arguments and explain why I am not convinced by them. The Argument from Non-Reflective Deliberation maintains that if we properly consider what is involved in theoretical or practical reasoning we find that there is nothing incoherent in the idea of a 'competent deliberator' who has yet to master the psychological concepts required for first-person self-knowledge. The Argument from Cognitive Development states that as a matter of fact such subjects are not particularly exceptional: in our early years we all answered that description.

Isn't a competent deliberator someone who is able to address (and so must understand) first-person questions about her reasons for belief or action? Let me quote a passage in which Boyle propounds a negative answer to that question:

“why couldn't a subject consider one of these simpler questions:

whether *p*

whether to do *A*

These questions certainly pertain to what the subject will believe or do, but their formulation does not make explicit reference to *her believing* or *her doing*. At least apparently, then, a subject might entertain one of these simpler questions without thinking of herself and her own belief or action as such: she might merely think about the *object* of her belief or action (the proposition that *p*, the deed of *A*-ing) without thinking of *herself* and her own *cognitive relation* to this object.” (2024, 17)

Boyle's point here is that there is no “quick argument” for the conclusion that a deliberator must have “the capacity to think of her own mental states as such”²⁶, and that, in the absence of a “quick argument”, the “Kantian claim that rationality requires the capacity to reflect on one's own representations looks strikingly substantive and not terribly plausible.” (ibid.). Now, one thing Boyle may have in mind in speaking of a ‘quick argument’ is the idea that reasons for action and belief just are relevant sorts of mental states (e.g. desires, beliefs, perceptions).²⁷

²⁶ Boyle (2024), 18.

²⁷ Compare Christine Korsgaard's suggestion (quoted by Boyle at p. 16 and again p. 18 n. 25) that a rational deliberator must be able to ask ‘Is this perception really a reason to believe?’ or ‘Is this desire really a reason to act?’.

In my view, Boyle is right to resist that idea. Ordinarily we think of normative reasons as (typically non-psychological) facts, not as mental states. But I think the Kantian nexus between rationality and reflection need not be motivated by a psychologistic conception of reasons. It may be motivated, I suggest, by careful consideration of something Boyle accepts: a competent deliberator is able to answer practical or theoretical questions by being responsive to relevant reasons.

First, a remark on what seems to be the most radical ingredient of Boyle's picture of non-reflective deliberation. The 'simple' question defining the goal of practical deliberation, he suggests, is *impersonal*: 'whether to do A'. It is a much-debated question whether practical deliberation is essentially *first-personal*. But even those who dispute this tend to accept that practical deliberation is personal, in the sense that it is about what some specific agent (or agents) is (or are) to do. There is more than one way to defend that assumption. One is to reflect on the central position in our practical thinking of the question 'How?' Practical deliberation, in paradigmatic cases, is a matter of determining how to achieve some objective, where the answer will need to be informed by a conception of the agent's practical abilities and opportunities for action. Suppose we find ourselves with a flat tyre in the middle of nowhere in a foreign country. 'What to do?' would be an extremely natural response to the situation.²⁸ Yet, the impersonal formulation will not get us very far. Any specific plan of action will depend on whose prospective activities we are talking about. There may be things we should do all together (e.g. wheel the car to the hard shoulder – something none of us may be able to do by themselves). As to who should, individually, do what, that will depend on the individual's talents and circumstances. You are good at changing tyres. I have a telephone. And so on. A reasoned answer to the question 'How to do A?' can only be given in the light of a conception of the relevant agent's abilities and situation.²⁹

How does this bear on the case of theoretical deliberation? I want to suggest that the deliberator's abilities are no less relevant here, though in a more indirect manner. One

²⁸ Compare Pnin's reaction to his discovery that he is on the wrong train: "'Important lecture!' cried Pnin. 'What to do? It is a catastrophe!'", Nabokov (1957), 150.

²⁹ It might be said that a deliberator may employ 'monadic' rather than relational concepts of their practical abilities. They might, for instance, think of an object as 'within reach' or 'liftable', rather than as 'something I am able to reach/lift'. The trouble with that way of thinking is that it would make it hard to reason practically *with others* – say, to work out that since we need to inform the car hire company I need to switch on my phone. The point weighs particularly heavily if, as seems plausible, the capacity for practical reasoning is acquired by exercising it together with others. For example, an infant participating in a turn-taking game learns to reason from 'we are playing together' to 'it's now my [or your] turn to do A'. See Moll 2007 for relevant developmental evidence. Briefly: from the age of 14 children develop an understanding of the relation between shared goals and individual contributions. They may, for example, 'remind' a partner who fails to do their bit. See also Kern & Moll (2017) for theoretical discussion of the role of shared activities in children's cognitive development.

difference between theoretical and practical deliberation is that the sort of question defining the goal of the former is not personal, let alone first-personal. The question, as Boyle points out, is simply whether *p*, not whether to believe that *p*.³⁰ And Boyle is surely right that a competent deliberator can sometimes answer that sort of question without reflecting on her beliefs, simply “on the basis of consideration of reasons for believing”.³¹ She might manifest her ability to believe that *p* for the reason that *q*, for example, by judging ‘clearly *p*, given that *q*’. Yet a competent deliberator also needs to be able to deal with less favourable situations. Suppose she has no evidence as to whether *p*. Or suppose her reasons for and against thinking that *p* are finely balanced. Surely the first thing a competent deliberator needs to think in such cases is: ‘I don’t know whether *p* – I can’t tell’. No ‘impersonal’ alternative, it seems, could capture the reason for withholding judgement to which she needs to be responsive in that sort of situation. ‘The question whether *p* has no answer’ or ‘there is no relevant evidence’ or ‘yes and no’ would not be quite right, and would not demonstrate much competence in theoretical deliberation. Being able *not* to judge, for good reason, is as much part of rationality as making a judgement for a reason. Arguably, the point has further implications. A competent deliberator needs to be able not only to reflect ‘I don’t know’ but also ‘I do know’ (how could she do the former without being able to do the latter?). So she needs to be responsive to reasons for claims to knowledge; in particular, she needs to be able to corroborate a claim to knowledge by reflecting on how she knows what she knows. Perhaps the basic case is one in which she can just see or otherwise perceive that *p*, so no deliberation is indicated. But in cases where she knows by working things out on the basis of relevant reasons, substantiating a claim to knowledge will be a matter of reflecting on her reasons for belief. It is here that we find a rationale for the view that a fully competent deliberator needs to be (potentially) reflective, primarily about what, and how, she knows but also, connectedly (in some cases), about what, and why, she believes. Just as practical rationality requires being able to form intentions in the light of consideration of one’s practical abilities and opportunities for action, so theoretical rationality requires being able to judge in the light of consideration of one’s ability to know.³²

³⁰ Moran’s account of what he calls the ‘deliberative stance’ appears to be committed to denying this. At least, he frequently characterizes theoretical deliberation as a matter of determining ‘what I am to believe’. (See for example Moran 2001, 63)

³¹ Boyle (2024), 18.

³² The line of thinking sketched here would provide support for an assumption on which I relied in my exposition of modest reflectivism: that the capacity to judge requires appreciating that judgements are open to requests for their credentials (e.g. ‘How do you know?’). Note that there is no commitment here to the idea that as a matter of complete generality, knowledge is the norm for judgements. For my purposes here it would be enough to suppose that *in some contexts* a judgement would legitimately be expected to express knowledge; and that someone who has fully mastered the capacity to judge will be alive to that sort of expectation.

This completes my lightning sketch of a (non-psychologicistic) case for the Kantian nexus. But is the nexus not undermined by the developmental evidence? Or, setting experimental findings to one side, is it not intuitively clear that young children’s cognitive abilities illustrate “the patent conceptual distinction [and possible one-way dissociation] between the capacity for first-order thought and the capacity for meta-representation”?³³

It will be useful to put Boyle’s argumentation here formally:

- (1) Young children have “the capacity to think intelligently about whether *p*”.
- (2) Young children lack “the capacity to think explicitly about *one’s own perspective on whether p*”.
- (3) Therefore, the former capacity does not presuppose the latter. (See 2024, 18 f., n. 26)

Boyle seems to think that all of this is more or less commonsense. No “special information from developmental psychology”³⁴ is required to see the force of the argument. I will make a final stab at being properly *kontrovers* by suggesting that on the contrary, the argument commits a fallacy of equivocation.

Mastery of a capacity is not an all-or-nothing matter.³⁵ Are typically developing 14-month-olds able to walk? ‘Yes’, we might say, ‘somewhat short distances, with a little help from others.’ On the other hand: ‘not for miles, by themselves, across difficult terrain.’ In a word, they are toddlers: someone who, as my dictionary explains, ‘moves with short unsteady steps while learning to walk’. Modest reflectivists, I suggest, should make analogous points about the capacities at stake in (1) and (2). Re (2): neither commonsense nor experimental findings suggest that young children are completely unreflective thinkers. From the middle of the second year, they begin to experience rudimentary self-conscious emotions, akin to pride or shame. Around the same time children begin to be able to think about their own and others’ perspectives; specifically, 2- and 3-year-olds have an increasingly rich understanding of perception and knowledge. To mention just one finding: many 2-year-olds are adept at using ‘I don’t know’ appropriately in conversations.³⁶ Whether this reflects a fully developed mastery of the concept of knowledge is a further question, but if we read (2) as the claim that young

³³ Boyle (2024), 18, n. 25.

³⁴ Boyle (2024), 18, n. 25.

³⁵ Here I follow Timothy Williamson: “On a thicker conception of abilities [which Williamson appears to endorse], they are not all or nothing. Diachronically, one may retain the ability to do something even through changes in one’s ability to do it. For instance, many young children have the ability to speak English, but their ability to speak English will improve as they grow up.”, Williamson (2018), 246.

³⁶ See Harris (2017). There has been much debate recently about children’s understanding of knowledge. Moll provides a helpful overview of the development of children’s understanding of perspectives.

children lack *any* understanding of their own perspective on whether p, (2) enjoys little empirical support. Re (1): do young children have the capacity to think intelligently about whether p? Are they competent deliberators? There are surely various ways in which an affirmative answer would need to be qualified. Yes, 2- and 3-year-olds can answer (certain) questions fairly intelligently, but their exercise of intelligence is often ‘scaffolded’ by caregivers, at least initially. Competence at shared deliberation may precede competence at individual deliberation. And even when aided by others, young children’s mastery of deliberation is substantially limited. They may be able to answer questions when they can see what the answer is or when the evidence is clear-cut. But they may find it hard to confront, and reason about, conflicting evidence. When asked questions about an object that presents conflicting appearances, for example, they tend to plump for one appearance, ignoring or denying the other.³⁷

Without being able to argue this in detail, I suggest (1) would surely be implausible if construed as the claim that young children have *fully mastered* the capacity for theoretical deliberation. Equally, as we saw, (2) is implausible if read as the denial of *any* mastery of the capacity to think about one’s perspective. The trouble is that if we adopt weaker, more sensible readings of (1) and (2), they lend no support to (3). They leave open the possibility that young children have the ability to deliberate, and to have beliefs, in the sense in which they have the ability to walk. They are beginning to master the capacities sustaining human beliefs, but full mastery of these capacities only emerges over the later preschool years, when children become able, among other things, to reflect on their own and others’ beliefs. It is not as if they first learn to negotiate the space of reasons and then, in a separate development, become reflective thinkers. The two things form part of a single, extended developmental process. Young children, as we might put it, toddle in (or into) the space of reasons, moving with short unsteady steps while learning to think.

5. Conclusion

The main thrust of Boyle’s *Transparency and Reflection* is to articulate and defend what might be called a ‘transformative’ view of self-awareness.³⁸ The human capacity for self-awareness ‘informs’ and ‘transforms’ our first-order attitudes and abilities. It is part of the very nature of our minds, not added to a form of mentality we share with subjects lacking that capacity. Yet,

³⁷ The classical study is Flavell, Flavell and Green (1983). For a nuanced perspective on more recent work in this area, see Moll and Tomasello (2011).

³⁸ Adapting Boyle’s distinction between transformative and additive views of rationality: see Boyle (2016).

in his recent work Boyle introduces a striking ‘additive’ element into the overall transformative picture. While ‘non-positional’ self-awareness is an essential aspect of ‘human mindedness’, the capacity for reflection is a further, additional accomplishment.³⁹ I have raised some questions about this additive element – in particular, about its compatibility with a reflectivist account of self-knowledge, and about its motivation. Looking back, I am struck by the fact that the notion of propositional knowledge plays a key role in the various questions I have raised (starting with the question of how to understand the place of propositional knowledge in Boyle’s two-stage version of reflectivism). I’d like to emphasize that this does not (I think) reflect a dogmatic commitment to a ‘knowledge-first’ style of epistemological theorizing, but merely the conviction that the notion of propositional knowledge is a crucial (and perhaps in some ways basic) ingredient of the way we ordinarily understand ourselves and others as rational thinkers. One thing we can learn from Boyle’s work is how much is to be gained from taking that understanding seriously.

REFERENCES

- Anscombe, E. (1957), *Intention*, Oxford.
- Boyle, M. (2011), “Transparent Self-Knowledge”, *Aristotelian Society Supplementary Volume*, Vol. 85 (1), 223–241.
- (2015), “Critical study: Cassam on Self-Knowledge for Humans”, *European Journal of Philosophy* 23, 337–348.
- (2016), “Additive Theories of Rationality: A Critique”, *European Journal of Philosophy* 24(3), 527–555.
- (2023), “Self-Consciousness, Transparency, and Reflection”, *Philosophisches Jahrbuch* 130(2), 110–129.
- (2024), *Transparency and Reflection*, Oxford.
- Craig, E. (1990), *Knowledge and the State of Nature*, Oxford.
- Dummett, M. (1991), *The Logical Basis of Metaphysics*, London.
- Flavell, J., Flavell, E. and Green, F. (1983), “Development of the appearance-reality distinction”, *Cognitive Psychology* 15.1, 95–120.

³⁹ An interesting complication: there is a sense in which reflection transforms the character of the reflected-upon-attitude (leading us to “inhabit it reflectively”, Boyle 2024, 254). But the complication is compatible with the additive element I mention in the text.

- Harris, P., Yang, B. and Cui, Y. (2017), “‘I Don’t Know’: Children’s Early Talk About Knowledge”, *Mind & Language* 32(3), 283–307.
- Kern, A. and Moll, H. (2017), “On the transformative character of collective intentionality and the uniqueness of the human”, *Philosophical Psychology* 30.3, 319–337.
- McDowell, J. (2011), *Perception as a Capacity for Knowledge*, Milwaukee.
- Moll, H. (2007), „Person und Perspektivität – Kooperation und soziale Kognition beim Menschen“, in F. Kannetzky & H. Tegtmeier (eds.), *Personalität – Studien zu einem Schlüsselbegriff der Philosophie*, Leipzig, 37–56.
- Moll, H. and Tomasello, M. (2012), “Three-year-olds understand appearance and reality—just not about the same object at the same time”, *Developmental Psychology* 48.4, 1124–1132.
- Moran, R. (2001), *Authority and Estrangement*, Princeton: Princeton University Press.
- (2003). “Responses to Shoemaker and O’Brien”, *European Journal of Philosophy* 11 (3), 402–419.
- Nabokov, V. (1957), *Pnin*, London.
- Roessler, J. (2013), “The Silence of Self-Knowledge”, *Philosophical Explorations* 16 (1), 1–17.
- (in press). “Perception, Knowledge, and Reflection”, in A. Giananti, J. Roessler & G. Soldati (eds.), *Perceptual Knowledge and Self-Awareness*, Oxford.
- Taylor, C. (1985), “Hegel’s Philosophy of Mind”, in his *Human Agency and Language*, Cambridge, 77–96
- Williamson, T. (2007), “Replies to Critic”, in P. Greenough & D. Pritchard (eds.), *Williamson on Knowledge*, Oxford.
- (2018), “Hyman on Knowledge and Ability”, *Philosophy and Phenomenological Research* XCVII(1), 243–248.

Johannes Roessler
Department of Philosophy
University of Warwick
Coventry CV4 7AL
United Kingdom
J.Roessler@warwick.ac.uk