

DISCUSSION NOTE

TRANSPARENCY AND INFERENCE

KIERAN SETIYA

Contrary to an argument by Matthew Boyle (2011), doubts about the inference from *p* to *I believe that p* do not support reflective theories of self-knowledge over an inferential or rule-following view.

In his enlightening response to Alex Byrne's 'Transparency, Belief, Intention' (2011), Matthew Boyle contrasts two theories of self-knowledge that emphasize 'transparency': the idea that, in self-ascribing mental states, one's attention is directed outwards, to the world and one's assessment of it, not to one's obscurely introspectible inner life.¹ In the case of belief, on which I will focus here, one asks oneself whether *p* is true; if one's answer is yes, one self-ascribes the belief that *p*. As Boyle explains, however, this description is ambiguous between two very different views. On what he calls the *inferential* model, one makes an epistemic transition, or inference, from the premiss that *p* to a conclusion about oneself. This is Byrne's proposal in the paper to which Boyle responds. According to the alternative, *reflective* view, the belief that one believes that *p* is not formed on the basis of a prior belief that *p*; instead, for rational creatures, to believe that *p* is already to have tacit knowledge of that belief.² As Boyle insists:

This still leaves room for the reflectivist to recognize an important truth underlying the phenomenon of transparency. On his view, the important truth is this: the very same actualization of my cognitive powers that is my believing *P* is, under another aspect, my tacitly knowing that I believe *P*. Hence, to pass from believing *P* to judging I believe *P*, all I need to do is reflect—i.e. attend to and articulate what I already know. (Boyle 2011, p. 229)

¹ This note is a reply to Boyle (2011). Although I disagree with him, I found his treatment of these issues very helpful.

² Boyle (2011, §§II, IV); the restriction to 'rational creatures' appears on pp. 235–6.

Both views purport to explain the impossibility of ‘self-blindness’: of subjects who have only third-person access to their beliefs, even though they have the concept of belief and their capacity for rational inference is intact. The inferential view explains this by accounting for self-knowledge of belief as a product of that capacity.³ The reflective view explains it by treating beliefs as tacitly self-conscious, at least in rational beings.

Although they differ significantly, these views have much in common. Boyle argues, however, that we should prefer the reflective view. Indeed, his reaction to the inferential model is, as he puts it, ‘incredulous’ (Boyle 2011, p. 226).

The basic reason to reject the idea that I infer a fact about my own psychology from a fact about the world is this: the inference is mad. ... To believe that I believe *P* is to hold it *true* that I believe *P*. Being a reflective person, I can ask myself what grounds I have for holding this true. The answer ‘*P*’ is obviously irrelevant. (Boyle 2011, p. 230)

The purpose of this note is to show that this argument misfires. Concerns about the madness of the ‘doxastic schema’, which takes me from *p* to the belief that I believe that *p*, do not favour the reflective over the inferential view.⁴ In making my case, I will not argue that the inference here is anything but mad.⁵ Rather, I will argue that if this inference is mad, the madness is equally present in reflective views.

Let us grant, then, that it is irrational to form the belief that I believe that *p* on the basis of my belief that *p* itself, by way of the doxastic schema. This transition is irrational because, as I know quite well, the putative fact that *p* is not good evidence that I so believe. I am not omniscient, nor do I think I am.⁶ On the reflective view, there is no transition of this kind. In believing that *p*, I already have tacit knowledge of my belief, which self-conscious judgement makes explicit. This is not mad inference because it is not inference at all.

So far, so good. But now consider how rational inference looks on the reflective view. From *p* and *if p, q*, I infer *q*. At least on occasion,

³ See Byrne (2011, p. 213), Setiya (2011, §2).

⁴ The terminology here is due to André Gallois (1996), who also notes the puzzling character of the inference.

⁵ In defence of the inference, see Byrne (2005, pp. 93–8; 2011, pp. 206–7), Setiya (2011, pp. 186–7).

⁶ See Boyle (2011, p. 231 n.8) for development of this point.

this inference must be rational. But if I make explicit the tacit knowledge involved in my beliefs, according to the reflectivist, here is how my inference looks:

$$\frac{p \text{ and I believe that } p \\ \text{if } p, q \text{ and I believe that if } p, q}{q \text{ and I believe that } q}$$

Although it differs from the doxastic schema, this pattern is equally mad. Its premisses, that p and I believe that p , that *if* p, q and I believe that *if* p, q , may be good evidence for q . They are not, at least not typically, good evidence that I believe that q ; and I am quite aware of this. Belief is not closed under logical consequence: there are many implications of my beliefs I fail to believe. Why should q not be one of them? The point is even clearer with non-deductive inference. That I know the premiss of a good inductive argument is not by itself much evidence that I believe the conclusion. I may never have considered the conclusion; I may not have made the inference. On the reflective view, when I make a rational inference, I form the belief that q on the basis of premisses that support its truth; but I also form the belief that I believe that q , which I know my premisses do not support. The madness has not been cured: it has become much worse. For the reflectivist, it is not just inference in accord with the doxastic schema that is irrational, but inference as such!

Boyle might complain that this argument misunderstands the reflective view. If belief is tacitly self-conscious, the premisses from which I infer that q are not grounds on which I self-ascribe the relevant belief. The belief that I believe that q is groundless; or since it is not distinct from my belief that q , it justifies itself. If you ask me how I know that I believe that q , I will say ‘Because I believe that q ’, or question the need for grounds; I will not cite q or the evidence for its truth. What would be mad is to base a belief on grounds that I know to be irrelevant. And that is not what I have done.

The problem with this response is that it can be made by the inferentialist, too. In other work, Byrne calls the transition from p to the belief that I believe that p an instance of ‘epistemic rule-following’, not inference (Byrne 2005, §7). In a related discussion, I took a similar view:

Unlike inference, in the ordinary sense, [the doxastic schema] does not draw on evidential support: [its premise], that *p*, typically is not good evidence for its conclusion, that I believe that *p*. Also unlike inference, the justification of the conclusion does not depend on the justification of the premise. When I infer from *p* to *q*, my belief that *q* will not be justified if my belief that *p* is not. By contrast, when I follow [the doxastic schema], I come to know that I believe that *p* even when that belief is irrational or unjustified. (These facts are related: the rationality of inference depends on the provision of evidence, which unjustified beliefs cannot supply.) (Setiya 2011, p. 184)

The conclusion I drew is that there are instances of epistemic rule-following that are not properly described as ‘inferential’. Since the doxastic schema presents an instance of this kind, the inferential view is poorly named. But this is not a substantive objection. In the absence of an argument against non-inferential rules, we can adapt and rename the inferential view.⁷ This leads to the crucial point. In describing the doxastic schema, the rule-following theorist—as we may now call him—should deny that I take the ‘premiss’ of the schema, that *p*, as evidence for its ‘conclusion’, or that it is the ground of my self-ascription in any sense other than being the content of the belief from which the self-ascription derives. The reflectivist cannot doubt that this is possible, or that it is sometimes rational. As we saw above, it is part of what happens in rational inference, on the reflective view: I form a belief, that I believe that *q*, by following a rule whose premisses are known to be no evidence for its truth.

In effect, there is a dilemma for Boyle’s objection. If we understand inference narrowly, the rule-following theorist can agree that it is irrational to make an inference of this kind. But he will deny that one is making an inference, in this narrow sense, when one follows the doxastic schema. If we understand inference broadly, as ‘a non-accidental transition between belief contents, where the reasonableness of the transition is open to assessment’ (Boyle 2011, p. 227), the rule-following theorist will admit that he takes an inferential view. But the reflectivist cannot hold that such transitions are insane when the contents with which they start are known to be no evidence for the contents with which they end. For his own view violates that constraint.

⁷ I defend the possibility of non-inferential rule-following in Setiya (2011, pp. 184–6).

The reflectivist could make a final move. Instead of pressing the negative claim that such transitions are invariably irrational, he can ask the rule-following theorist to explain, in positive terms, how the subject understands the rationality of her self-ascriptive belief.⁸ She cannot justify this belief by citing what she takes to be good evidence. But she must have something to say. What could it be? The premiss behind this challenge is a form of ‘internalism’ on which it is irrational to form or hold a belief unless one is in a position to recognize its epistemic rationality. The demand is controversial, and it is not clear to me how the reflective theory meets it. But on a plausibly weak conception of ‘being in a position to know’, the rule-following theorist can do so. So long as you have the concepts with which to articulate the doxastic schema, you are in a position to explain why it is non-accidentally reliable: why it is safe to self-ascribe the belief that *p* on the basis of (what you take to be) the fact that *p*. You may not give this explanation to yourself, but at least in principle, you could.

None of this supports the inferential or rule-following view, as against reflectivism. The point is that the complaint of madness does not tilt the other way. In permitting one to form the belief that one believes that *p* on the basis of beliefs one knows do not give evidence that one believes that *p*, these views are on a par. If such transitions are mad, neither view can be correct. If they can be rational, it is no objection to the doxastic schema that it involves a movement of this kind. Either way, Boyle is wrong to find in the madness of ‘inferring’ from *p* to *I believe that p* any reason to prefer the reflective view.⁹

Department of Philosophy
University of Pittsburgh
Pittsburgh, PA 15260
USA
kis23@pitt.edu

⁸ This is how Boyle replied to an earlier version of this note; see also Boyle (2011, p. 231).

⁹ Thanks to Matt Boyle, Alex Byrne and Casey Doyle for generous discussion of these topics.

REFERENCES

- Boyle, Matthew 2011: 'Transparent Self-Knowledge'. *Proceedings of the Aristotelian Society Supplementary Volume* 85, pp. 223–41.
- Byrne, Alex 2005: 'Introspection'. *Philosophical Topics*, 33, pp. 79–104.
- 2011: 'Transparency, Belief, Intention'. *Proceedings of the Aristotelian Society Supplementary Volume* 85, pp. 201–21.
- Gallois, André 1996: *The World Without, The Mind Within*. Cambridge: Cambridge University Press.
- Setiya, Kieran 2011: 'Knowledge of Intention'. In Anton Ford, Jennifer Hornsby and Frederick Stoutland (eds.), *Essays on Anscombe's Intention*, pp. 170–97. Cambridge, MA: Harvard University Press.