

RATIONALITY, CONSCIOUSNESS AND COGNITION

ANDREI MĂRĂȘOIU¹

How do consciousness and cognition relate? Smithies (2012) has offered an argument for thinking that cognitive states (beliefs) should be individuated relative to conscious states. Beliefs can be appraised for rationality, the argument goes, only if they are introspectively accessible. I think Smithies is largely correct about the relation between consciousness and cognition. However, I argue that he is wrong about thinking that we should support introspective accessibility by rational evaluation – where that evaluation is construed as amenability to reflective scrutiny. Instead, I argue that we should support the introspective availability of our beliefs to ourselves by appealing to the cognitive responsibility we hold for our believings.

1. Consciousness and believing

What is the right way to characterize cognition? To mobilize intuitions, consider zombies. Do zombies (if they are assumed to exist) have a cognitive life? Many options are on the table, but the rough trends are clearly discernible: Characterizing cognitive life in functional terms will lean towards answering ‘Yes’. Whereas characterizing cognitive life in terms of conscious experience will lean towards answering ‘No’. Let’s take a step back and think of what goes into these verdicts. In a recent paper (his 2012), Declan

¹ ICUB, Research Institute of the University of Bucharest, Str. Dimitrie Brândză nr. 1, Bucharest, Romania 060102; "Nicolae Kretzulescu" Superior Commercial School, Bd. Hristo Botev nr. 17, Bucharest, Romania 030233. Contact: aim3gd@virginia.edu .

Smithies contributes to the debate by proposing the following argument in support of a phenomenal approach to cognition:

- (1) All cognitive states are intentional states that play a rational role.
- (2) All intentional states that play a rational role are either conscious or individuated by their relations to conscious states.
- (3) So, all cognitive states are either conscious or individuated by their relations to conscious states. (Smithies 2012, p. 358)

Smithies understands cognition to cover beliefs, as opposed to sub-personal states like visual information-processing. So the first premise is not all that controversial: we should expect, if something counts as a full-fledged belief someone holds, that it should somehow feature in reasoning. It may influence what conclusions one reaches in speech or thought, or what action is outputted based on one's available motivations.

The conclusion, however, makes the controversial claim that cognition should be individuated in terms of phenomenal consciousness. Many of our beliefs may still be implicit. But what's important is that they are individuated in terms of our ability to make them explicit. Smithies contrasts this picture of cognition with a functional view of cognition, on which we individuate cognitive states in terms of their amenability to computation, be it global inference or modular processes. That view has largely been orthodoxy in early cognitive science and is even now familiar from the work of, e.g., Jerry Fodor.

Given how controversial the conclusion of this argument is, and how often-accepted the first premise seems to be, it's natural to focus the discussion on the second premise. And here Smithies has a supporting argument (numbers re-written for clearer reference):

- [4] All intentional states that play a rational role are introspectively accessible.
- [5] All introspectively accessible states are either conscious or individuated by their relations to consciousness.

[2] So, all intentional states that play a rational role are either conscious or individuated by their relations to consciousness. (Smithies 2012, p. 365)

Once again, one premise seems innocuous enough, premise (5). Our criterion for deciding whether a cognitive state counts as conscious is whether it's *possible* that it could be brought to the fore in introspection. Some beliefs are trickier than others – biased beliefs might be a good example. But one wouldn't, surely, wish to deny the bare possibility of bringing a belief to introspection if one does indeed have that belief.

And, once again, the conclusion (2) is controversial. So we should do well to direct attention to the premise on which the shift in theoretical burden seems to occur – premise (4). This states that "All intentional states that play a rational role are introspectively accessible." To that I now turn.

2. Rationality, reflection and computation

Here Smithies' argument begins to get a bit unclear. In seeming support of (4), he writes:

The concept of rationality is essentially tied to the practice of critical reflection. To a first approximation, a belief is rational if and only if it is based in such a way that it would survive an idealized process of critical reflection. On this conception, the rationality of one's beliefs depends solely upon facts that are accessible to one by means of introspection and *a priori* reasoning, since these are the methods that constitute the practice of critical reflection. (Smithies 2012, p. 364)

We can see how an argument might go: We can't subject a belief to reflective scrutiny unless we first introspect that we have it. And we can't deliberate whether we're being rational in holding

that belief unless we subject the belief to reflective scrutiny. Therefore, (4) is true.

Yet once we make this argument explicit, the problem with it is explicit as well. Why should we characterize rationality in terms of getting vetted in reflective deliberation? We might easily think of an alternative standard for rationality, namely, causal transitions between beliefs which realize some formal pattern of inference or computation which the evaluators (they need not be us) have decided upon endorsing, for whatever reason. The contrast here is between rationality as amenability to reflection and rationality as mere conformity to some types of computations.

Smithies sees the contrast. He writes:

One tempting avenue is to appeal to Dennett's (1969) distinction between personal and subpersonal levels. On this proposal, beliefs are intentional states of the person, whereas subdoxastic states are intentional states of parts of the person – namely, their computational subsystems. The problem with this proposal is that we need a more fundamental account of what makes it the case that an intentional state is properly attributed to the person as opposed to one of the person's subsystems. ...Broadly speaking, there are two options for cashing out the distinction between personal-level and subpersonal-level intentional states: one can appeal either to facts about consciousness or to facts about functional role. (Smithies 2012, p. 360)

If Smithies were right about the personal/subpersonal distinction, we would be back to square one. Functional role approaches may favor a computation-first approach to cognition, whereas phenomenal approaches may obviously lean towards individuating cognitive states in terms of conscious states. If so, then we can't find support for (4) which isn't question-begging. That would be bad news for Smithies' argument. Fortunately, I think there is a way out: to support (4) without any circles in justification.

3. Cognitive agency

We can characterize the personal/ subpersonal distinction in terms of one's *cognitive responsibility* for the beliefs one holds, *leaving it open* whether rational believing entails any reflection.

Conceiving of cognitive agency is, *a priori*, logically independent from both phenomenology and functionalism. We may form a belief implicitly, so cognitive agency doesn't imply that the belief must be conscious. And we may form a belief without abiding by some computational or inferential canon (first-order logic, classical probability theory), so cognitive agency doesn't imply that the belief must be identified with the functional role it plays in our web of beliefs, in guiding our speech or our overt behavior. Talk of cognitive responsibility codifies the fact that (other things being equal) we're to blame if we form or sustain irrational beliefs, and we deserve praise if we form or sustain beliefs in full compliance with some rational norms of choice.

We can then see why Smithies thinks reflection is needed for rationality. He seems to subscribe to what Sher (2009) calls "the searchlight view". In general,

an agent's responsibility extends only as far as his awareness of what he is doing. He is responsible only for those acts he consciously chooses to perform, only for those omissions he consciously chooses to allow, and only for those outcomes he consciously chooses to bring about. (Sher 2009, p.4)

This directly applies to cognitive agency in forming or sustaining beliefs. How else, the thought often goes, could we control our beliefs other than by reflecting on them? We deliberate about how good our beliefs are – how well supported by evidence, how well inferred, how well they cohere with each other, etc. And that reflective scrutiny is what makes us more than mere bearers of our beliefs – it makes us authors of them.

This is a "searchlight view" because reflection casts light over the beliefs it considers the way you cast light into a dark

room. Just as you're typically looking for something when you go into a dark room with a searchlight, so you're on the lookout for irrationality when you reflect on your beliefs.

As seductive and entrenched in common sense as this view of the role of reflection may seem, the view is mistaken. Regardless of the virtues and vices of excessively reflecting on the beliefs you hold, there is little by way of argument to support the idea that we're responsible for what we believe *because* we can reflect on it. And there is much that speaks against this idea (for some incisive remarks, cf. Kornblith 2010). For one thing, we rarely *form* beliefs reflectively; typically, reflective scrutiny comes in after the fact. To think we could always form beliefs reflectively even if we rarely do is to offer an advertisement for the powers of reflection instead of acknowledging a fact of our cognitive lives.

To combat the idea that we're only responsible for believing to the extent that we reflect on our beliefs, here is a different – and, I submit, more convincing – notion of cognitive agency, due to Hieronymi:

We change our minds, and so control our attitudes, not by reflecting on or thinking about our mind, but rather by thinking about the object of our thoughts. The controlling happens “behind the lens,” so to speak. The thinking subject controls its thoughts in thinking them. (Hieronymi 2008, p. 371)

Hieronymi's line meshes well with an insight from virtue epistemology: who we are as thinkers, and how well our thinking goes, depends on our actual dispositions to form and sustain beliefs, rather than the occasional, perhaps sometimes erroneous and always somewhat effortful reflective scrutiny we may subject ourselves to.

4. Reflection, introspection, and thinking well

We found a way to account for the cognitive agency we manifest in forming and sustaining beliefs while leaving it open whether

we need to reflect on those beliefs or not. And the question of the rationality of our beliefs (the end-products of our thinking) turns, on this view, on whether the very activity of thinking is rational or not. I've remained deliberately agnostic about the chances of reflection – how often we do it, how good we are at it, and how much of a significance we attach to it.

Yet this picture can provide better reason to think premise (4) is true, without any detour through reflection, and so improves on Smithies' own considerations in favor of it. Recall (4) said: "All intentional states that play a rational role are introspectively accessible."

How can we do that? I think the basic idea is to give a fillip to the searchlight view without buying into the whole thing. Once we've *done* something – formed a belief, changed how strongly we hold it, or come to question it – we should be able to introspect having that belief, and to ascertain in introspection whether we really believe it, or only weakly lend it some credence. All this we can do by "looking within", as it were. This much is needed in support of (4), and I agree with it. Introspection can come after the fact, and it is no precondition for rationality. What matters is introspective *accessibility*, not the introspective experience itself.

But distinguishing introspection from reflection allows us to put to one side all the intricacies concerning reflection, and stay with introspection and its role in retrospectively testifying to our cognitive responsibility. This is support enough for (4), since all (4) requires is that intentional states be accessible in introspection, not in reflection.

We now have a new boundary to separate the personal from the subpersonal. It's not which states are conscious, since some conscious states of mind may be mere happenings to us, as cognitive agents: we wonder about things, find ourselves daydreaming, etc. And it's not what functional role our cognitive states play, because we may wish to dissociate ourselves from some automatic connections our semantic memory routinely achieves, e.g. in cognitive and affective biases. Rather, the new boundary to separate the personal from the subpersonal is this: those cognitive states are personal that we are cognitively responsible for.

This criterion of what is personal affords a new criterion of rationality: those *beliefs* are rational which *we* are rational in forming or sustaining. This leaves open which norms rationality imposes, and how we conform to them. But it does pivot in the way needed: avoiding reflective requirements while giving its due to introspective accessibility.

We can now raise the question of what rationality is again. Being true to oneself – thinking according to one's extant thought habits – may differ from thinking in a way that others find acceptable. Which norms are appropriate for rational thought is a difficult further question. Yet we can support the introspective accessibility of our beliefs by seeing believing as something we do, and are responsible for, without settling that question.

5. Argument for thinking **Conclusion**

How do consciousness and cognition relate? Smithies (2012) has offered an that cognitive states (beliefs) should be individuated relative to conscious states. That argument trades on thinking that beliefs are rationally assessable, and only introspectively accessible states can be assessed for rationality. I think Smithies is largely correct about the relation between consciousness and cognition. However, I argue that he is wrong about thinking we should support introspective accessibility by rational evaluation – where that evaluation is construed as amenability to reflective scrutiny. Instead, I argue that we should support the introspective availability of our beliefs to ourselves by appealing to the cognitive responsibility we hold for our believings.

Acknowledgements

The paper benefited from questions and suggestions from Declan Smithies (Ohio State University) and an anonymous reviewer for

the Southern Society for Philosophy and Psychology. I hope to be able to more fully address prof. Smithies' suggestions in future work. Financial support was provided by the Research Institute at the University of Bucharest through an ICUB Fellowship for Young Researchers.

REFERENCES

- Hieronymi, P. (2008) Responsibility for believing. *Synthese* 161, pp. 357-373.
- Kornblith, H. (2010) What Reflective Endorsement Cannot Do. *Philosophy and Phenomenological Research* 80, pp. 1-19.
- Sher, G. (2009) The Searchlight. In: *Who Knew? Responsibility without Awareness*. OUP.
- Smithies, D. (2012) The Mental Lives of Zombies. *Philosophical Perspectives* 26, pp. 343-372.

