6   Counting concepts: response to Paul Boghossian

Mark Sainsbury and Michael Tye

Introduction

Paul Boghossian has mounted a sustained attack (dating back at least to 1994) on externalist semantic theories, on the grounds that they are inconsistent with transparency, understood along these lines:

> *Introspective knowledge of comparative content (IKCC)* When our faculty of introspection is working normally, we can know <u>a priori</u> via introspection with respect to any two present, occurrent thoughts whether they exercise the same or different concepts.

The originalist theory we proposed in *Seven Puzzles of Thought* accepts externalist principles, and thus rejects IKCC. It must therefore address Boghossian's arguments. We did this in the first edition of our book, and we returned to the fray, benefiting from a further version of Boghossian's criticisms[1], in the second (paperback) edition. Here the dialog continues: in the chapter to which this is a response,[2] Boghossian presents a new version of his criticisms. They are characteristically insightful and penetrating, and present a challenge to originalists: if transparency is abandoned, how can we understand rationality?

The challenge is brought out by examples, especially the cases of Paderewski–Peter and Switched–Peter, where both Peters are assumed to be entirely rational. Peter in the Paderewski case hears Paderewski at a concert, and forms the belief that he has musical talent. Later, Peter encounters Paderewski making a political speech at a rally. Peter has been told on good authority that no politicians have musical talent, and so rationally believes that he has now encountered a different Paderewski, one lacking musical talent. According to originalism, there is just one

---

[1] Presented at an Author Meets Critics session at the Pacific APA, in March 2013.
[2] Chapter 5 above. This is related to, but not identical with, his contribution to the Author Meets Critics session.

public concept *Paderewski,* which Peter exercises both when he forms the belief that Paderewski has musical talent, and when he forms the belief that Paderewski lacks musical talent. In the originalist framework, Peter has contradictory beliefs: apart from negation, the beliefs are made up of just the same concepts in the same position. The challenge is to explain how Peter can, nonetheless, be rational. If a rational thinker can believe contradictions, do we not lose all grip on what makes a thinker rational?

Switched–Peter has been moved without his knowledge from Earth to Twin Earth. For the sake of the argument, we accept that he will retain his earthly concept *water*, but will after a time acquire the Twin Earth concept *twater*.[3] His word 'water' becomes equivocal, sometimes expressing the concept *water* and sometimes the concept *twater*. Boghossian suggests that, in the simplest case, Switched–Peter will activate the concept *water* when recalling experiences involving water on Earth, but will activate the concept *twater* when thinking about the waterlike liquid in his current environment. He will remain unaware of the fact that he is activating different concepts in these different contexts. Accordingly, he will be disposed to regard the following argument as valid, even though, as we imagine it issuing from his mouth, it is not, since it equivocates with respect to 'water':

   (1) Whoever floats on water, gets wet
   (2) Pavarotti once floated on water
   (3) Therefore, Pavarotti once got wet

Peter's "regarding the argument as valid" need not amount to the explicit belief that it is valid, and so it is a state he can be in even if he lacks the concept *valid.* He was in that state if he happily made the transition from premises to conclusion and was disposed to justify belief in the conclusion by pointing to those premises.

We have assumed that Peter is rational, yet he regards an argument as valid which is not. He has failed to appreciate that the word 'water' expresses different concepts on its two occurrences (first, the concept *water*, then the concept *twater)*. If a rational thinker can make a mistake of this kind, do we not lose all grip on what makes a thinker rational?

---

[3] Both we and Boghossian regard the relevant description of the switched case as ultimately inappropriate. Boghossian thinks switching is impossible (as externalistically described), because it is inconsistent with transparency (IKCC), though the least bad version is the one in which the concepts *water* and *twater* cohabit in the mind of Switched–Peter. We think switching is possible, though we do not think that the concepts *water* and *twater* would cohabit: instead, Peter's concept *twater* would come to displace his concept *water.* These issues do not bear on the present discussion.

In thus describing the data offered by the examples of the two Peters, we have been careful not to ascribe to them any metaconceptual beliefs, for they might be in the relevant states even if they lacked the concept *concept*, and so could not even formulate any meta conceptual beliefs. As Boghossian stresses, we need to be equally careful in giving the originalist redescriptions of the cases in the light of which both Peters are shown to be rational.

For Paderewski–Peter, it's as if he believes he has two concepts *Paderewski* when in fact he has one. For Switched–Peter, it's as if he believes that his word 'water' expresses just one concept when in fact it expresses two. We say "it's as if" in both cases, since, for the reason given in the previous paragraph, we cannot count on Paderewski–Peter to believe he has two concepts *Paderewski*, or on Switched–Peter to believe he expresses different concepts on different occasions when using the word 'water'.

If it's as if Paderewski–Peter believes he has two concepts, his dispositions to form first–order thoughts expressible with the word 'Paderewski' coincide with the dispositions of a thinker as similar to him as possible except believing that the word is equivocal, expressing now one, now another, of two concepts. If it's as if Switched–Peter believes that his word 'water' is not equivocal, he is disposed to make invalid inferences from premises expressible with the word 'water'. More generally, his belief–forming dispositions are those of someone who explicitly believes that the word is not equivocal.

Both being contradictory and being valid essentially depend on how many concepts are involved. We agree with Boghossian that thinkers' thought–forming dispositions are normally appropriate to the number of concepts they are using.[4] This normally enables them to avoid certain simple contradictions, and to make correct assessments of the validity–value of simple arguments. We disagree with Boghossian that rational thinkers cannot have thought–forming dispositions that are out of line with the number of concepts they are in fact exercising. We regard the Peter cases as demonstrating that point: Paderewski–Peter has thought–forming dispositions appropriate to his having two concepts *Paderewski* when in fact he has just one. Switched–Peter has dispositions appropriate to there being just one concept expressible by 'water' when in fact there are two.

---

[4] We agree that Boghossian's pundit is irrational if he believes he has been on welfare but has never been helped. This is not strictly contradictory, but entails a contradiction in the presence of obvious further facts, e.g. that welfare is a form of help. The pundit has no excuse for not believing these facts or for failing to see that they ensure that he is committed to a contradiction.

This brings us to the heart of Boghossian's challenge: How in that case can we say what it is for a thinker to be rational? We will show that the Peter cases pose no problem for a theory of rationality set within an originalist framework.

Originalists allow that there are circumstances, however rare, in which thinkers behave as if they were mistaken about how many concepts their thoughts involve; or about whether a pair of thoughts is contradictory; or about whether an argument fails to be valid through equivocation. Moreover, originalists hold that sometimes it can be *rational* so to behave. When this occurs, a *rational* thinker may believe a contradiction, or make an incorrect estimate of whether or not an argument is valid. It is not that the thinker identifies the thought as a contradiction, and then rationally believes it. Rather, the thinker rationally fails to detect the contradictory character of the thought, or rationally believes that it is not contradictory.

Boghossian challenges us to provide a theory of rationality according to which this could be so. We meet that challenge as follows:

> R: Thinkers who believe contradictions, or who incorrectly assess the validity status of simple arguments, or who make fallacious inferences from simple arguments, are irrational, unless they have an excuse.

The two Peter examples on which Boghossian draws serve to illustrate kinds of excuse that may make it not merely not irrational, but positively rational, for a thinker to believe a contradiction, or to incorrectly evaluate an argument's validity status. We show that in the two examples on which Boghossian's challenge primarily turns, it is not hard to see what excuses the thinkers had. Paderewski–Peter rationally believes that there are two Paderewskis, and this blinds him to the contradictory character of a pair of his thoughts. Switched–Peter is rational *not* to believe that he expresses more than one concept by his word 'water', and this blinds him to an argument's equivocal nature.

Paderewski–Peter

Peter was told on good authority that no politicians have musical talent. It was rational for him to believe this, even though it is false, and, given the story, an immediate consequence is that there are two Paderewskis: Paderewski the pianist and Paderewski the politician. Peter believes this and it is rational for him to do so.[5] According to originalism, concepts are

---

[5] We are unclear on which side of Boghossian's division this falls: is it lacking empirical information (that there is just one Paderewski) or is it mishandling information (not appreciating how the informational elements are related)?

identical if and only if they have the same historical origin. So, *if* there are two Paderewskis, it follows that there are two concepts, one for Paderewski the pianist, and one for Paderewski the politician. So, Peter believes something (that there are two Paderewskis) that, given original- ism, entails the falsehood that there are two concepts *Paderewski,* and hence entails that there are noncontradictory beliefs to the effect that Paderewski has musical talent and that Paderewski lacks musical talent. Peter's rational beliefs entail that he can believe these things without believing a contradiction. This is his "excuse" in the sense of R. Our attribution to Peter of a single concept *Paderewski,* and the consequences of this attribution, do not entail that Peter is irrational, and are inevitable, given originalism, once it is allowed that he is rational in believing that there are two Paderewskis.

Switched–Peter

Paderewski–Peter has rational beliefs that entail he has more concepts than he actually has. By contrast, Switched–Peter rationally fails to have beliefs that would entail that he has two 'water' concepts, and so rationally behaves as if he had just one.

One option would be to iterate the structure appropriate to the Paderewski case. Then we would say that Switched–Peter has the rational belief that he has not been switched, and this entails that he has just one concept *water.* We could then appeal to the previous idea, that if a thinker's rational beliefs entail facts about how many concepts she has, she is rational to form beliefs which, *on that supposition*, would not be contradictory, even though they are *in reality* contradictory.

But, as Boghossian points out, this may be too demanding. Perhaps Switched–Peter has never envisaged or heard of Twin Earth, and so is in no position to form any belief or supposition about his relation to it. That's why we stressed, when responding to Boghossian in the second edition of *Seven Puzzles*, that a rational lack of belief can do the same kind of work as the presence of a rational belief. Switched–Peter is rational not to believe he has been switched; it would be crazy for him to believe he has been, just as crazy as it would be for you or us to believe such a thing. If he has not been switched, he has just the concept *water,* and not the concept *twater.* It would be irrational of him to believe his situation to be one in which he has two concepts expressible by 'water'. So in treating, as he does, all occurrences of 'water' in his mouth as he would if he had the explicit belief that they express the same concept, he is avoiding the irrationality of treating these occurrences as the equivocal things they really are. This is his 'excuse', in the sense of R. He is rational to lack a

belief that, if he possessed it, would unlock a correct assessment of the validity status of the Pavarotti argument. So he is rational not to believe that the Pavarotti argument is invalid through equivocation. Moreover, the argument strikes him as valid; he is rational to lack any belief that would undermine the claim that it is valid; hence he is rational to believe it is valid, and rational to use the premises as a basis for believing the conclusion. This is the position that all ordinary (unswitched) people are in when they confront the argument. The argument strikes them as valid. Knowing nothing of Twin Earth, it's rational for them not to believe any Twin–Earth–related defeater for this belief. Hence it is rational for them to hold it. Given his ignorance of the switch, Switched–Peter's rationality in regarding the argument as valid is exactly like that of ordinary people.

Conclusion

The Paderewski–Peter case relies on the idea that a false belief can be rationally held, if one is presented with misleading evidence. The Switched–Peter case relies on the idea that there can be truths that it is rational for one not to believe, for one may have no shred of evidence for them. We take it that these ideas are platitudes common to every discussion of rationality, yet are the only ones on which we have drawn in illustrating the kinds of excuses relevant to R. Hence we do not believe that our originalist defense of the rationality of both Paderewski–Peter and Switched–Peter calls for any startling novel doctrines about rationality. We agree with Boghossian that there are deep and difficult issues in this area. But we also believe that the specific attributions of rationality to the two Peters are resolved within an originalist framework by rather simple considerations.