

(This is a pre-print text)

BOGHOSSIAN'S REDUCTIO OF COMPATIBILISM*

Comment on Boghossian's "What the Externalist Can Know A Priori"

Carlos J. Moya. University of Valencia

In a number of papers,¹ Paul Boghossian has argued against the compatibility between content externalism and self-knowledge by trying to show that, if externalism is correct, then there are possible situations in which a subject would be systematically wrong about the identity and difference of his thought contents. These possible situations involve the subject's being unwittingly switched between, say, Earth and Twin Earth. In this case, he would judge that two of his thoughts are of the same type when, on an externalist individuation, they are not. In Boghossian's terms, externalism is incompatible with transparency of mental content. This line of argument has received considerable attention and been the object of a wide and subtle discussion in recent times.²

In the present paper, however, Boghossian has turned to a different line of attack to compatibility between self-knowledge and externalism. He now tries to elaborate a *reductio ad absurdum* of compatibilism by trying to show that this position is committed to the claim that a subject can come to know, purely a priori, substantial truths about the world which every reasonable person would agree can only be known a posteriori.

There are some antecedents to this line of argument, as well as some criticisms to it, and I shall eventually refer to them while discussing this paper.³

According to Boghossian, Oscar, a compatibilist, is in a position to argue, purely a priori, as follows:

* Research for this paper has been funded by the Spanish Government DGICYT as part of the projects PB93-1049-C03-02 and PB93-0683. My thanks to this institution for its generous help and encouragement. I also thank Josep Corbí and Tobies Grimaltos for comments on previous drafts.

¹ Especially in 'Content and Self-Knowledge', *Philosophical Topics* 17 (1989), pp. 5-26; 'Externalism and Inference', in E. Villanueva (ed.), *Rationality in Epistemology*, Philosophical Issues 2, Ridgeview Publishing Company, Atascadero (California), 1992, pp. 11-28; and 'The Transparency of Mental Content', *Philosophical Perspectives* 8 (1994), pp. 33-50.

² Participants in this discussion include Ted A. Warfield, 'Privileged Self-Knowledge and Externalism Are Compatible', *Analysis* 52 (1992), pp. 232-237 and 'Knowing the World and Knowing Our Minds', *Philosophy and Phenomenological Research* 55 (1995), pp. 525-545; Stephen Schiffer, 'Boghossian on Externalism and Inference', in E. Villanueva (ed.), op. cit., pp. 29-37; K. Falvey and J. Owens, 'Externalism, Self-Knowledge and Skepticism', *The Philosophical Review* 103 (1994), pp. 107-137; J. Owens, 'Pierre and the Fundamental Assumption', *Mind and Language* 10 (1995), pp. 250-273; P. Ludlow, 'Externalism, Self-Knowledge, and the Prevalence of Slow Switching', *Analysis* 55 (1995), pp. 45-49; T. Burge, 'Our Entitlement to Self-Knowledge', *Proceedings of the Aristotelian Society* 91 (1996), pp. 91-116; C. Peacocke, 'Our Entitlement to Self-Knowledge. Entitlement, Self-Knowledge and Conceptual Redeployment', *Proceedings of the Aristotelian Society* 91 (1996), pp. 117-158, among others. I myself have contributed a paper on this subject, 'Externalism and Self-Knowledge', to the VIII Inter-University Seminar on Philosophy and Cognitive Science, University of Granada, 23-25 May 1996.

³ See M. McKinsey, 'Anti-Individualism and Privileged Access', *Analysis* 51 (1991), pp. 9-16; A. Brueckner, 'What an Anti-Individualist Knows A Priori', *Analysis* 52 (1992), pp. 111-118; M. McKinsey, 'Accepting the Consequences of Anti-Individualism', *Analysis* 54 (1994), pp. 124-128; A. Brueckner, 'The Characteristic Thesis of Anti-Individualism', *Analysis* 55 (1995), pp. 146-148; J. Brown, 'The Incompatibility of Anti-Individualism and Privileged Access', *Analysis* 55 (1995), pp. 149-156. See also note 7.

1. If I have the concept *water*, then water exists.
 2. I have the concept *water*.
- Therefore,
3. Water exists.

Clearly, that water exists is something that cannot be known a priori. The form of the argument is valid. So, at least one of the premises will be either false or not knowable a priori. Boghossian, however, holds that Oscar is bound to accept that both premises are true and knowable a priori. Therefore, he is bound to accept that the existence of water can be known a priori. Since this is absurd and is implied by compatibilism, compatibilism must be false.

I would like to resist being forced to choose between externalism and self-knowledge. I think we can have both. I am, then, a compatibilist. My aim in this comment is going to be mainly negative: I will try to show that Boghossian's incompatibilist argument does not succeed. But I will also say something positive in defence of compatibilism, by giving some indications about how a plausible compatibilist externalism might be construed.

Premise 1 of Oscar's argument is based on the externalist thesis that there is a dependence relation between thought content and external individuation conditions. Premise 2 rests on the assumption of self-knowledge. Though Boghossian takes this premise to be obviously acceptable, I find the way it is stated rather controversial. Normally, self-knowledge is held to extend to thought contents, rather than to the concepts one possesses in thinking those contents. In fact, it is hard to accept that one enjoys privileged access to the concepts one has, at least if this means the concepts one *masters*, in having a certain thought. Normal adults, I would say, know directly, non-inferentially, what they are currently thinking, in that they can give true reports of these thoughts, but I am pretty sure that they are quite often wrong about their mastery of the concepts that make up those thoughts. Burge's social externalism might have some bearing on this claim. So, if premise 2 is to be plausible, 'having a concept' has to be understood in a weak sense. One has a concept, in this weak sense, just in case one is able to have and to express, with a sufficient degree of competence, according to normal standards, thoughts involving the concept. So understood, premise 2 is certainly acceptable, given self-knowledge.

Premise 2 and the antecedent of premise 1 in Boghossian's argument are stated under the assumption that the concept had by Oscar the compatibilist, namely the concept *water*, fulfills a number of conditions, related to presuppositions involved in Twin Earth experiments. First, water and twater (that is, any other substance macroscopically indistinguishable, but microscopically, chemically distinct from water) 'have to be thought of as distinct substances, distinct natural kinds'. Second, the word 'water' is intended by Oscar to express a natural kind concept. Third, Oscar has no views about the chemical composition of water. Fourth, Oscar's concept *water* is atomic, not compositional. Finally, the concept is to be thought of by Oscar as individuated in terms of its referent. If Oscar is to know premise 1 a priori, it is essential that these conditions are known by him a priori and that they, together with his possession of the concept, logically imply the consequent, namely that water exists.

Is a Twin Earth externalist committed to hold that premise 1 is true and knowable a priori? A natural and powerful response (which Boghossian considers and finally rejects) by an externalist would be to say that the second condition is too weak for this premise to be accepted. For imagine Peter, a compatibilist and a defender of the phlogiston theory. He intends his term 'phlogiston' to express a natural kind concept, as does Oscar with his term 'water'. Following Oscar's example, he reasons as follows:

- A. If I have the concept *phlogiston*, then phlogiston exists.
- B. I have the concept *phlogiston*.
- Therefore,
- C. Phlogiston exists.

An externalist will contend that what leads Peter to a false conclusion is that 'phlogiston' does not actually name a natural kind, so that the concept it expresses is not the concept of a natural kind. What is implied by externalism is not premise 1, but, at most, premise 1', which should replace the former:

1'. If I have the concept *water* and the concept *water* is the concept of a natural kind, then water exists.

Now, however, conclusion 3 cannot be deduced from premises 1' and 2. Besides, and more importantly, the antecedent of 1' cannot be known a priori. Whether a term actually names a natural kind and so expresses the concept of a natural kind can only be known a posteriori. If this is correct, premise 1' becomes a triviality. Nothing substantial about the natural world can be obtained from externalism together with self-knowledge a priori.

Boghossian's strategy is to block this natural externalist move by denying that an externalist can allow for the existence of atomic concepts expressed by terms with an empty extension, such as 'phlogiston' in our world or 'water' on his extraordinary Dry Earth. On the other hand, taking such terms to express compound concepts makes questions about compositionality or atomicity of concepts depend on contingent external circumstances, such as the existence or non-existence of a natural kind, and this he deems unacceptable. Compositionality and atomicity are internal, non-relational properties of a concept; these properties supervene on the molecular and functional structure of the concept's possessor and are not contingent on what there can be in the outer world. If all this is correct, an externalist must concede that if a term expresses an atomic concept, then it has a non-empty extension. Now Boghossian is in a position to establish his reductio of compatibilism against objections. He can now reject premises of the form of our premise 1'. The antecedent of the conditional need not include that the concept Oscar has is the concept of a natural kind, which could only be known a posteriori. It suffices that Oscar intends to use it as a natural kind concept and that the concept is atomic (together with the rest of conditions stated above). And all this can be known by Oscar a priori. So, Oscar, the compatibilist, knows a priori that the concept he expresses with his term 'water' is the concept *water* and that this concept is atomic. So his term 'water' has a non-empty extension. Since this term expresses the concept *water*, and he knows it does, he knows that the term's referent is the property of being water and that its extension is everything that has this property, namely all bits of water. He can conclude, then, purely a priori, that water exists or has existed at some time. But to think that this can be known a priori is absurd. Therefore, compatibilism is false.

Let me try to assess Boghossian's incompatibilist argument. A crucial question is the role that empirical information, e.g. that water is H₂O, that water is a natural kind, or that the term 'water' names a natural kind, plays in externalism as a philosophical theory of meaning and content. This question is controversial, but is essential in discussing incompatibilist arguments. Once the role played by empirical information is rightly understood, it can be seen that no substantial truths about the external world can be obtained from externalism together with self-knowledge a priori. Externalism is the view that our thought contents, including our

concepts, are individuated, in part, by external conditions. But there are several possible ways of developing this basic insight. Twin Earth experiments are supposed to show that tokens of 'water' do not mean the same on Earth as on Twin Earth. But why is this so? The most common answer is: because Earthians and Twin Earthians causally interact with different substances, water and *twater* respectively. This answer corresponds to a widely held version of externalism. 'Causal externalism' might this version be called. This simple version faces, in my opinion, serious, maybe insurmountable difficulties.⁴ But it is not the only possible version. On the version I would favour, the answer to the above question would be: because Earthians and Twin Earthians learn and teach the meaning of 'water' in connection with samples of what are in fact different substances. On this version, our words' meaning depends on external conditions because certain bits of the external world are used to *define* those words, to give those words its meaning. If, in order to define 'water', I say to someone 'water is *this*', then the ostended bit itself gets into the meaning of the word and into the concept it expresses. This external bit becomes a norm for a correct use of the word. It is part of the structure of the concept this word expresses. This construal of externalism, inspired in some Wittgensteinian reflections,⁵ might be called 'normative externalism'. Normative externalism includes causal externalism, but not conversely. Normative externalism is not really far from Putnam's conception. Remarks about paradigmatic samples, with a clear Wittgensteinian flavour, can also be found in Putnam's "The meaning of 'meaning'". But these remarks are often forgotten in favour of cruder causal versions. On the normative construal, when I point to a sample of water to define the word 'water', I need not know what its chemical composition is, nor do I need to be able to distinguish it from, say, *twater* if I were confronted with it. But the actual essence of the sample, whether we know it or not, helps giving the word its meaning, and this meaning is different from that it would be given to the word on Twin Earth. If this is on the right lines, it might reconcile Boghossian's view that compositionality and atomicity are internal to concepts with dependence of these properties on external conditions, as far as these conditions are an aspect of the concept itself. Let us see how.

On this construal of externalism, the real essence of a certain substance, which is only discoverable a posteriori, is given a constitutive role in individuating content and concepts. It is plausible to think that our present conviction that *water* is an atomic concept is not independent of the empirical discovery that water is a natural kind, namely H₂O. Before this discovery, human beings might well be agnostic about atomicity or compositionality of their concept *water*. We would not say that the liquid on Twin Earth is not water if we did not know that water is H₂O, and this is only knowable a posteriori. However, the constitutive relevance of this a posteriori information in determining our concept *water* and its difference from Twin Earthians' concept *twater* depends on our prior commitment to define terms like 'water' by means of external samples.⁶ It might be objected that our acceptance of the difference between the concepts *water* and *twater* shows that our concept *water* was atomic in the first place, prior

⁴ I do not know whether this causal version of externalism can meet Boghossian's objection of reductio we are dealing with. But I am pretty sure that it cannot meet objections based on 'switching cases', which I referred to in the first paragraph of this paper. It also faces the 'disjunction problem' that Fodor discovered in causal theories of meaning and content.

⁵ See especially Wittgenstein, L., *Philosophical Remarks*, Blackwell, Oxford, 1975, I, 6, and *Philosophical Investigations*, Blackwell, Oxford, 1963, I, 50. Remarks about paradigmatic samples, with a clear Wittgensteinian flavour, can also be found in Putnam's "The meaning of "meaning"".

⁶ This commitment might not be optional. Meaning might need, as a condition of its existence, such external, objective samples, at least is something like Wittgenstein's private language argument is accepted as cogent. The indispensability of external samples is a plausible way of reading this argument.

to the empirical discovery that water is a natural kind, so that atomicity does not really depend on a posteriori information. My answer would run as follows. For some concepts, such as *water*, *gold*, or *tiger*, we may hold an assumption that there is a hidden unitary essence lying behind the items that fall under the concept. But this assumption, which we may know a priori, is not enough to make those concepts atomic, for, on the normative version of externalism, the assumption is defeasible, in so far as we leave the last word about atomicity or compositionality of those concepts to the real nature of paradigmatic samples. Therefore, we do not have, a priori, a sufficient amount of data to decide the issue. In saying 'water is *this*', the ostended bit becomes an aspect of the concept, so that, if we discover that no common unitary essence underlies the different samples, we cancel the atomicity assumption, and compositional aspects become decisive. If atomicity of such concepts as *water* was settled in advance, then, were we to discover that there is no common essence to water, we would say 'so, there is no water' or 'water does not exist'; but we would not say this; what we would say, instead, is that water was, in fact, a collection of different substances. Counterfactuals such as 'if there were no H₂O in our environment, our term "water" would express a different concept from the one it now expresses' rest, I contend, on our decision to understand by 'water' *this*, where the ostended bit is a sample. In this frame, the empirical discovery that paradigmatic samples of water are H₂O confirm the (defeasible) assumption that *water* is an atomic concept. If this is correct, we cannot know a priori that the concept *water* is atomic. Therefore, an externalist cannot conclude, purely a priori, that water exists. He *might* conclude that, since he has the concept *water*, there have to be external possibility conditions for that concept, but this transcendental reasoning to the effect that there are some external conditions or other, controversial as it may be, is not absurd.⁷ If all this is correct, an externalist may plausibly hold that tokens of the same word-type can express an atomic concept in one world and a compositional concept in another, even if their inhabitants are molecular and functional duplicates. Thereby, he can vindicate premise 1' and avoid reductio.

Suppose, however, that Boghossian is right and that atomicity or compositionality of concepts supervenes on molecular and functional structure of their possessors, independently of what there is in the outer world. If so, the concept expressed by 'water' on Dry Earth would be atomic. Could an externalist accept, against Boghossian's contention, that a term with an empty extension expresses nonetheless an atomic concept? I think he could.⁸ What are the truth conditions of 'water is wet' on Dry Earth? We can start with the following statement:

'Water is wet' is true on Dry Earth if, and only if, Dry Earth's water is wet.

But 'Dry Earth's water' has no referent. Before this, an externalist could opt, couldn't he, for saying that the right part of the biconditional is false or for saying that it has no truth value. Only in this latter case might he be forced to say that 'water is wet' has no meaning on Dry Earth. If he opts for saying that lack of reference of a component makes the sentence false, he can say that 'water is wet' is false on Dry Earth, but has meaning. But then, in what conditions would it be true? Well, it would be true if there were in fact a substance, a natural kind, in the extension of 'water' and this substance were wet. Why an externalist may not say this? He could then say what is the concept expressed by tokens of 'water' on Dry Earth: it is the concept Dry Earthians would have if there were in fact a natural kind, with the right external appearance,

⁷ This is something that Brueckner points out against McKinsey's incompatibilism. See note 3.

⁸ I have been led to think of this by Peter Carruthers' treatment of empty demonstrative thoughts in his article 'Russellian Thoughts', *Mind* 96 (1987), pp. 18-35.

denoted there by tokens of 'water'. It must be conceded, however, that in order to be more precise about the concept an externalist should know *what* natural kind would be denoted. But the above description of the concept seems precise enough to counter Boghossian's claim that an externalist cannot allow for atomic concepts in empty cases. If all this works, it could again support premise 1' in Oscar's argument, which, to recall, included in the antecedent the condition that the concept had by Oscar has to *be* the concept of a natural kind, instead of just being intended by him to be so.

These possible replies to Boghossian's incompatibilist argument may not be compossible. But if at least one of them is successful, this would suffice for compatibilists to elude *reductio*.

Department of Metaphysics and Theory of Knowledge
University of Valencia (Spain)
E-mail: Carlos.Moya@uv.es