

Incompatibility Arguments and Semantic Self Knowledge

Henry Jackman
York University

There has been much discussion recently of what has been labeled the “Brown-Boghossian-McKinsey”, “Brown-McKinsey” or sometimes just “McKinsey” arguments for the incompatibility of externalism and self-knowledge. However, while the three author’s arguments have been treated as interchangeable, they are not identical. In particular, Brown’s and Boghossian’s arguments have a fairly serious flaw that cannot so easily be attributed to McKinsey. In what follows, I’ll (1) present a version of the ‘received’ “Brown-Boghossian-McKinsey” argument, (2) outline what I take to be the most serious objection to it, (3) explain why this sort of objection does not seem, or do not seem immediately, to tell against McKinsey’s argument, and (4) suggest a number of alternative responses that might apply to McKinsey as well.

The “Brown-Boghossian-McKinsey” (BB) argument against the compatibility of Externalism and Self-Knowledge is correctly attributed to Jessica Brown, and Paul Boghossian, and it runs something like this:

BB-1: You can know *a priori* that you are thinking a thought with the content *Water puts out fires*.

BB-2: You can know *a priori* that you couldn’t have a thought with the content *Water puts out fires* unless the environment is a certain way, e.g. some samples of water have sometimes existed....

BB-3: You can’t know *a priori* that your environment is the relevant way, e.g. some samples of water have sometimes existed....

BB-1 is taken to an explication of the doctrine of self-knowledge. BB-2 is a commitment that is taken to come with externalism, since externalism, if true, seems to have been established by purely *a priori* methods such as canvassing our intuitions about thought experiments like the Twin Earth cases. According to this version of the argument for the incompatibility of externalism and self-knowledge, BB-1 and BB-2 are jointly incompatible with the overwhelmingly plausible BB-3, so one of BB-1 or BB-2 must

be given up.

However, popular as it is, the BB argument is only as plausible as BB-2's assumption that the dependence of our water-thoughts on the existence of actual water can be known *a priori*, and there are very good reasons to doubt that this is the case. Something like BB-2 might be true if we could be sure that "water" was, in fact, a successfully referring natural kind term, but while it may seem clear to us that it is such a kind term, its being one isn't something that we can know *a priori*. We might be able to tell by reflection that we *intend* the term to pick out a natural kind, but it is one of the main characteristics of externalism that such intentions are defeasible. First of all, if one endorses social externalism of Burge's sort, one must leave room for the possibility that, in spite of the fact that you take the term to pick out a natural kind, the term picks out, say, more of a phenomenal or functional kind since that is how "water" is actually used in one's community. Further, even if one's usage was typical of one's community, an externalist is committed to the possibility that an entire community could turn out to be mistaken about whether a term of theirs picked out a natural kind. Purported natural kinds such as "air", "jade", "tree" or "lily" all failed to pick out natural kinds in spite of our initial assumption that they did, and the question of whether "water" will share the same fate is an *empirical* rather than *a priori* question.

It is easy to get the impression that BB-2 is a commitment that externalists must take on given that most of their arguments are based on *a priori* 'thought experiments'. This is, however, a mistake. The externalist thought experiments typically just *presuppose* (reasonably) that our environment is a certain way, and thus can't be part of any argument to establish that it is that way. What, if anything, is determined *a priori* by the, say, Twin Earth thought experiments is the conditional claim that *if* our environment were a certain way, then our thoughts would have content *A*, while *if* our environment were another way, then our thoughts would have content *B*. Since such conditionals are incompatible with internalist accounts of meaning, these conditional claims are enough to establish externalism *a priori* as a general claim about meaning. Even if, in fact, our environment were such that what we referred to was determined completely by our beliefs and dispositions (perhaps we had been dreaming our entire life), that wouldn't change the fact that, as a general thesis about 'the meaning of "meaning"', internalism is inadequate. The equation that determines what we mean includes a place for our environment, even if the value for that term might run to zero.

Given the implausibility of BB-2, the BB argument is thus in pretty bad

Incompatibility Arguments and Semantic Self Knowledge

shape. However, while arguments of the BB form are often referred to as “McKinsey-style” or even “McKinsey’s”, McKinsey’s own argument is more like the following:

McK-1: Oscar can know *a priori* that he is thinking that water puts out fires.

McK-2: The proposition that Oscar is thinking that water puts out fires logically implies the proposition *E*, (e.g. some samples of water have sometimes existed....).

McK-3: The Proposition *E* cannot be known *a priori*, but only by empirical investigation.

According to McKinsey, since we can know *a priori* all of the logical consequences of the things that we know *a priori*, the combination of McK-1 and McK-2 is incompatible with the extremely plausible McK-3. Hence, one of McK-1 or McK-2 must be given up.

There are some differences in style between McKinsey and the BB reasoner, but the real difference of substance has to do with the relation between BB-2 and McK-2. The BB reasoner treats the relation between our water thoughts and the external conditions that are necessary conditions for them as *a priori*, while McK-2 suggests that McKinsey is inclined to treat something like the relation as representing a logical or conceptual truth.

This might seem like a comparatively trivial difference if one thought that McK-2 would also imply that the relation between water thoughts and the existence of water was knowable *a priori* (since logical relations between propositions could be known *a priori*). However, McKinsey seems to deny precisely this implication. How he sees his own arguments as different from those of the more common BB form shows up in the following remarks he makes about Boghossian:

Boghossian defends a variant of my argument, using the following instance of modus ponens:

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

According to Boghossian, the doctrine of privileged self-knowledge implies that one can know (2) *a priori*, and externalism implies that (1) is knowable *a priori*. So if compatibilism were true, one could know (3) *a priori*, which is absurd. This *reductio* differs from mine in two important ways. First, unlike mine, Boghossian's argument does not assume that externalist theses like (1) are, if true, logical or conceptual truths. And secondly, unlike mine, Boghossian's argument *does* assume that such externalist theses are, if true, knowable *a priori*. I take this second feature to be a serious weakness, since I think it is fairly clear that externalist theses like (1) are *not* knowable *a priori*.

The criticisms of the BB argument I discussed earlier involve arguing that we are not in a position to know by reflection alone what BB-2 claims that we are. However, the passage quoted above suggests that McKinsey does not take the availability of any such reflective knowledge to be part of his argument. Indeed, he seems to be in agreement with critics of BB when he claims that "no form of incompatibilist argument is going to work if the argument itself assumes that the externalist theses are knowable *a priori*."

I find BB-2's claim that the relation between our water thoughts and actual water is knowable by reflection to be, in spite of its apparent falsity, a lot clearer than McKinsey's suggestion that it is a conceptual or logical truth. Nevertheless, I'll try to discuss briefly why McKinsey thinks that its status as a logical truth does not entail its being knowable *a priori*. McKinsey, if you remember, argues that:

(1) The proposition that I am thinking that water puts out fires logically implies the proposition *E*,

He denies, however, that this entails that we can know *a priori* that:

(2) If I am thinking that water puts out fires, then *E*.

Rather, he takes (1) to entail that that, *if* I can know *a priori* that:

(3) I am thinking that water puts out fires

Then I can also know (2) *a priori*, and thus know *a priori* that:

(4) *E*

Consequently, McKinsey's reasons for doubting that BB-2 follows from McK-2 relate back to his doubts about McK-1. If BB-1 were true, then BB-2 would follow from McK2. The conclusion that we don't have the sort of *a priori* knowledge posited by BB-2 will thus, from McKinsey's perspective, just give him more reason to deny the compatibility of McK-1 and McK-2. In other words, if externalism were true and we could know

Incompatibility Arguments and Semantic Self Knowledge

the content of our thoughts *a priori*, then we could tell *a priori* that these contents were externalist. We can't tell *a priori* that their contents are externalist, so if externalism is true, then we can't know the contents of our thoughts *a priori*.

In short, letting *P* stand for the proposition that *I am thinking that water is wet*, *E* stand for the proposition that *water exists*, and *A(x)* for *x is knowable a priori*, the BB argument can be represented as:

B1. A(P)	Assumption
B2. A(P → E)	Assumption
B3. A(E)	1, 2, AMP

McKinsey's argument, on the other hand, when fully spelt out is more like the following:

M1. A(P)	Assumption
M2. P → E	Assumption
M3. (A(P) & (P → E)) → A(P → E)	Assumption
M4. A(P → E)	1, 2, 3, conj & MP
M5. A(E)	1, 4 AMP

Brown and Boghossian's problematic assumption B2 is not assumed by McKinsey at all, and since both arguments are meant to be reductios, simply denying McKinsey's analog of B2, M4, will not get the compatibilist out of trouble, since M4 is itself taken to follow from assumptions about externalism and self knowledge.

So just what then, if anything, is wrong with McKinsey's own arguments? Ultimately, it seems that the problem with the McKinsey argument stems from the same source as Brown and Boghossian's argument. Namely, all three assume that competent speakers must have grasp of their concepts that is more 'adequate' than externalist's allow. In Brown and Boghossian's case, this shows up in their assumption that we could tell *a priori* whether or not the contents of our thoughts were 'world-involving', and in McKinsey's it will show up in the externalist's rejection of the notion of 'conceptual implication' presupposed by McKinsey's argument.

In particular, many would argue that type of dependence associated with semantic externalism is 'metaphysical' rather than 'logical'. One need not be able to tell *a priori* the *metaphysical* consequences of what one knows *a priori*, and so if the dependence in this case is metaphysical, there need be no tension between externalism and self-knowledge. That is to say, the

sort of implication in premise M2 must be logical/conceptual if M3 is to have any plausibility, but (some would argue) M2 is only plausible if the implication in question is understood as metaphysical.

So why should we believe that *P* logically or conceptually implies *E*? After all, as McKinsey himself puts it,

Let us say that a proposition *p* *conceptually implies* a proposition *q* if and only if there is a correct deduction of *q* from *p*, a deduction whose only premises other than *p* are necessary or conceptual truths that are knowable a priori, and each of whose steps follow from previous lines by a self evident inference rule of some adequate system of natural deduction. I intend the relation of conceptual implication to be an appropriately *logical*, as opposed to metaphysical, relation.

Given this explication, for the proposition that I am thinking that water is wet to logically/conceptually imply the proposition that water exists, there would have to be a deduction from the proposition that I am thinking that water is wet to the conclusion that water exists that adds to *P* only “necessary or conceptual truths that are knowable a priori,” and it is hard to see how there could be such a deduction in the absence of an antecedent commitment to M4. It may be true that all the worlds picked out by *P* are also picked out by *E*, but in the absence of this being knowable *a priori*, the necessity involved seems like more metaphysical than logical.

The worry that the relation in question may be metaphysical rather than logical is reinforced by the fact that McKinsey’s argument against it’s being a metaphysical relation seems to rely on a somewhat tendentious conception of what externalism is. In particular, McKinsey characterizes externalism as something like:

(EXT) Some neutral cognitive states that are ascribed by the *de dicto* attitude sentences (e.g. “Oscar is thinking that water is wet”) necessarily depend upon or presuppose the existence of objects external to the person to whom the state is ascribed.

McKinsey then argues that the notion of “necessarily depends” can’t be understood as mere metaphysical necessity, since:

given certain materialistic assumptions that are pretty widely held, it would follow that probably *all* psychological states of *any* kind would be wide... For instance, it is plausible to suppose that no human could (metaphysically) have existed without biological parents... If this is so, then Oscar’s thinking that water is wet metaphysically entails that Oscar’s mother exists.

Incompatibility Arguments and Semantic Self Knowledge

This criticism is fair enough, but the conclusion we should draw from it may be that a formulation of externalism requires us to be more specific about the sorts of ‘external objects’ that the thesis requires our thoughts to be necessarily dependent upon rather than changing our understanding of necessary dependence to a logical one. Or perhaps better still, one could rely on one of the many formulations of externalism that don’t make mention of any sort of necessary dependence. After all, externalism can be characterized as the more minimal thesis that, say, “some mental properties with content may not preserve across physical, internal replicas”, and on such a version of the doctrine, there is less reason to think that the externalist is committed to *a priori* conceptual dependencies between thoughts and things.

There are, of course, Quinean reasons for thinking that the distinction between ‘conceptual’ and ‘empirical’ entailments is itself somewhat tenuous. Further, the distinction between ‘conceptual’ and ‘metaphysical’ necessity is precisely one of those that becomes blurred when one moves from an internalist to an externalist framework. (Or perhaps better put, for the externalist, there is no clear distinction between conceptual and empirical implications, so ‘conceptual’ implications need not be available *a priori*.)

For the externalist, while we know what we are thinking, our grasp of those thoughts is often ‘partial’ or ‘incomplete.’ If our grasp of a thought is incomplete, then we cannot be expected to have access to all of that thought’s logical implications. While such a response would concede to McKinsey that externalism is not compatible with the most robust forms of self-knowledge possible, this sort of concession seems to be entailed by many versions of semantic externalism *anyway*.

This is most explicit in Tyler Burge’s work, and Burge argues that “What I have called ‘partial understanding’ is common or even normal in the case of a large number of expressions in our vocabularies.” One may be able to know *a priori* all of the logical consequences of those propositions that we completely grasp *a priori*, but such a ‘complete’ grasp of our concepts is precisely what the externalist often denies that we have. (A novice chess player could thus know that he was thinking about a king without knowing that he was thinking about a piece that was capable of castling, even if the latter was, in some sense, conceptually implied by the former.) If ‘partial’ understanding is the norm, then the inference from McK-1 and McK-2 to McK-3 is blocked.

Of course this ‘partial’ understanding of our concepts may itself amount to a type of failure of self-knowledge, but it is not a failure that is philosophically problematic for the externalist. People freely admit that they can be mistaken about the essential properties of the objects they talk and

think about, so there is nothing problematic about denying self-knowledge of this form.

In conclusion, then, while his reliance on the notion of logical or conceptual implication make his argument more elusive than the transparently problematic arguments of Brown and Boghossian, the source of the response to both is the same. For the externalist, our grasp of our concepts is rarely complete or 'adequate', and there is no reason to think that we need know the properties that are *essential* to a concept's correct application. This is enough to simply falsify the second premise of the BB argument, but it tells against McKinsey's as well. This more modest understanding of how we grasp our concepts entails an understanding of what it takes to know that you believe, say, that water is wet, is not robust enough to generate any problems when combined with claims about what such a proposition does, or does not, conceptually imply. Just as a small dose of an illness can prevent one from getting a serious case, the comparatively modest lack of conceptual knowledge that has always been associated with externalism prevents it from entailing a full-blown lack of self-knowledge of the sort familiar from the arguments of Brown, Boghossian, and McKinsey.