What's Wrong with McKinsey-style Reasoning?

James Pryor
NYU
<jim.pryor@nyu.edu>
12/5/2006
forthcoming in *Internalism and Externalism in Semantics and Epistemology*, ed. by Sanford Goldberg

I

It's widely accepted nowadays that some of our thoughts are externalist. What does that mean? I understand a **thought's content** to be those of its representational aspects that one *experiences as* representational and are essential to its being *that thought*. And we can understand a thought to be **externalist** when subjects who are internally the same can differ with respect to whether they're thinking the thought's content.[1]

Our thoughts about water are widely thought to be externalist. So too are our demonstrative thoughts.

---

[1] I leave it open whether a thought's content suffices to determine the thought's truth-conditions. On a view like Lewis 1979's, when I think to myself "My pants are on fire," and you think to yourself "My pants are on fire," we're thinking *the same* content though our thoughts have different truth-conditions. Segal 1989 proposes a similar view about demonstrative thoughts. I don't count extra-contentual semantic differences of these sorts as externalist. Neither do I count as externalist views that say that Earthling thoughts are about H_2O, but only contingently so; those very same thoughts could have been about XYZ instead. An externalist has to say there are *essential* differences in the contents of some duplicates' thoughts.

My gloss on "externalist" employs the notion of subjects being "internally the same." It isn't straightforward what that amounts to. It'd be nice if our definition of externalism didn't presuppose any commitments about materialism. But it's not obvious that we can characterize "how you are internally" in terms of phenomenology either. Some philosophers (Dretske 1995 Ch 5; Tye 1995 §5.4 and 2000 §3.5) say that sensory phenomenology can itself be externalist. Our definitions ought to permit such a view. Additionally, I think it can also be part of your phenomenology that you are now occurrently thinking such-and-such—where that content may intuitively be an externalist one. For these reasons, I find it difficult to specify what it is for subjects to be "internally the same," in a way that preserves the familiar classifications. I'll just ask you to use whatever rough and ready understanding you have of this.

If we accept any kind of externalism, then we confront a puzzle that arises about our ability to tell what we're thinking. This comes from what's commonly known as **McKinsey's Argument**.[2] It rests on two ideas.

First, we think you can tell the contents of your thoughts just by introspection. For example, you can tell just by introspection that you're thinking that water puts out fires:

McK-1 You're thinking a thought with the content *Water puts out fires*.

As we proceed, it will be handy to have a way of identifying thoughts, without prejudging anything about how they're individuated, or indeed, whether they're even externalist. One way to identify a melody is to *produce* it: you can say "Hey, you know that jingle [here you hum: *la-la-lalalaa*…]?" Analogously, I think you can identify a thought content by thinking or entertaining it. You can say: "the thought content [here you *think* the relevant content, perhaps by rehearsing to yourself the sentence '`Water puts out fires`']." I'm using the notation:

the content *Water puts out fires*

to express this way of demonstrating a thought content, by thinking it.[3]

The second idea driving McKinsey's Puzzle is that you can *also* tell that some of your thoughts are externalist, purely by armchair philosophical reflection. And it would seem that if a given thought is externalist, then it's only available to be had by subjects in certain sorts of environments (more on this in a moment). Hence, it looks like you can establish a priori something of the form:

McK-2 If you're thinking a thought with the content *Water puts out fires*, then…[here some claim about your environment, e.g., it actually does or did contain samples of water].

---

[2] After McKinsey 1991; see also Brueckner 1986 and Brown 1995.

[3] Notice that this is not a *description*; it doesn't just mean "whatever content I think by saying these words to myself." It's a way of genuinely apprehending the content, and thereby enabling yourself to refer to it. This will be important later.

Putting the two together, it looks like you can conclude, purely on the basis of introspection and a priori philosophy:

McK-3 Your environment is the relevant way [e.g., it actually does or did contain samples of water].

And while few doubt that we *do know* things like McK-3, it's extraordinary that we should be able to know them *purely on basis of introspection and a priori philosophy*. That's the puzzle. It's counter-intuitive that you should be able to tell what your environment is like just on the basis of this kind of reasoning.[4]

In discussing this puzzle, some authors refer to our supposed introspective knowledge of McK-1 and our supposed philosophical knowledge of McK-2 as all being "a priori."[5] The name doesn't matter much. But there are important epistemic differences between introspection and the usual paradigms of the a priori. Calling them all "a priori" can encourage confusions; some of which we'll be disentangling later. Hence, I will reserve the name "a priori" for justification that comes from logical understanding, philosophical reasoning, and so on. Things like *I am now thinking about a prime number* I'll say instead are justified through your introspective experience or awareness of your occurrent mental life.[6] I'll use the umbrella term "by reflection" to cover the lot. So in my terminology, the surprising result posed by McKinsey's Puzzle is that you can tell what your environment is like *purely by reflection*.

One response to the McKinsey Puzzle *endorses* the result that you can establish things like McK-3 by reflection alone. That is one way of construing **Putnam's**

---

[4] As we'll discuss later, the puzzle isn't confined to your *knowing* or having some special *authority* about the presence of water. It's already counter-intuitive that you should be able in this way to acquire any justification at all to believe there really is water in your environment.

[5] See, for example, McKinsey 1991, Boghossian 1997, and McLaughlin and Tye 1998. See also Kitcher 1980 §V.

[6] I'll avoid any substantial assumptions about what that amounts to. BonJour 1998, pp. 7ff also denies that introspective justification is a priori.

Argument in Chapter 1 of *Reason, Truth, and History*. Putnam argues that people who have always been brains in vats can't refer to or think about those vats. (At least, not *by* using the word '`vat`.') So, in order for us to have thoughts about vats (using '`vats`'), our environment has to be a certain way: it has to be such that we haven't always been brains in vats. As it happens, we can tell by introspection that *we do* have thoughts about vats. (In fact, such thoughts are *necessary*, to be entertaining the skeptical hypotheses we are entertaining.) So it follows that we haven't always been brains in vats.[7]

Other philosophers resist that result. They say we can only establish things like McK-3 through empirical investigation. So the puzzling reasoning has to be blocked somehow. One way to block it is to go **Incompatibilist** about externalism and your ability to know the contents of your thoughts by reflection alone. The Incompatibilist says it can't be true *both* that a thought has an externalist content *and* that you're able to tell your thought has that content, just on the basis of introspection.[8] Some Incompatibilists take the incompatibility to discredit externalism. Others take it to discredit your ability to tell what externalist thoughts you're having by reflection alone.

I think these are all over-reactions to the puzzle. We can find a more sober response that steers between them. We *can* tell what we're thinking by reflection alone, even when what we're thinking is externalist. However, this *doesn't* give us a route to reflective knowledge—or even reflective justified belief—about what our environment is like. So I will argue.

As it happens, I'm *not* sure that reflective knowledge of our environment is to be avoided at all costs. After all, some epistemologists maintain that thoughts like ***My senses***

---

[7] This is only one interpretation of Putnam's argument. Some interpretations employ further premises about what brains in a vat *do* refer to and think about, when they use the word '`vat`'; or what *I* would refer to with '`vat`' if I were a brain in a vat. Other interpretations focus on disquotational knowledge about one's language, rather than introspective knowledge of what one is thinking.

Sawyer 1998, Warfield 1995 and 1998, and Tymoczko 1989 defend similar responses to the McKinsey Puzzle.

[8] This is McKinsey's, Brueckner's, and Brown's own response to the puzzle (Brown takes it back in her 2004). See also Boghossian 1997. There are also other arguments for Incompatibilism, besides McKinsey's; two notable sources are Woodfield 1982 and Boghossian 1989.

are reliable are justified a priori (albeit defeasibly). Here's another possible route to reflective knowledge of your environment: (i) notice, on the basis of introspection, that you're having an experience as of hands, and that you have no evidence that your senses are misleading you; (ii) apply your favorite a priori epistemology of perception to get the conclusion *So I am justified in believing that I have hands*; (iii) help yourself to a defeasible but rational ampliative inference from *I am justified in believing P* to *P*. Voilà: now you've got a purely reflective (albeit roundabout) justification to believe you have hands. It may even suffice for knowledge.[9] This merits careful discussion—especially step (iii)—but it's not obvious that the reasoning is illegitimate. Neither does it trade on any assumptions peculiar to externalism. So my examination of the McKinsey argument won't take it for granted that reflective knowledge of our environment is flat-out impossible. I just want to get the details straight. As these turn out, I think the McKinsey argument will *not* give us a route to such knowledge.

II

Let's attend to the bits we left unspecified at the end of McK-2. We supposed that, if a given thought is externalist, then it'd only be available to be had by subjects in certain sorts of environments. Is that supposition right?

It'll be useful to think about an example drawn from the history of chemistry.[10]

Mendeleev presented his first periodic table in 1869. At that time, and then again more carefully in 1871, he postulated the existence of four missing elements. He called one of them '`ekaboron`.' In 1879, Lars Fredrick Nilson, unaware of Mendeleev's predictions, spectrographically identified a new element in some Scandinavian minerals. He managed to chemically isolate an oxide of this element; and he dubbed the element '`scandium`.' This turned out to be Mendeleev's ekaboron. At least, it occupied

---

[9] Compare "The Explainer" in Hawthorne 2002.

[10] For background, see http://web.lemoyne.edu/~giunta/EA/CONTENTS.HTML and http://homepage.mac.com/dtrapp/periodic.f/periodicity.html.

ekaboron's position in the periodic table; and Mendeleev had closely predicted many of its properties, such as its atomic weight, its valence, and the density of its oxide.

Now, in reality, Mendeleev thought elements were individuated by their atomic weight rather than by what we call their "atomic number." Protons weren't discovered until 1918. But it will simplify our discussion to pretend that Mendeleev did stipulate how many protons ekaboron has (namely, 21) in advance. Additionally, I will assume that ekaboron *was* scandium, that is, that Mendeleev was able to think and talk about this element before it was discovered in nature and chemically isolated. That assumption may well be challenged; and I'm not certain it's true. But for this discussion we'll assume it. Finally, let's imagine that Mendeleev engaged in correspondence with another chemist in 1871, discoursing at length about ekaboron but *not* telling her what the element's fundamental individuating properties were.

We now have three different subjects thinking about scandium. First, there's Mendeleev, who stipulates that with '`ekaboron`' he's referring to the element with 21 protons. Second, there's Mendeleev's friend, who acquires competence with the name '`ekaboron`,' but who doesn't know how the substance is chemically individuated. Third, there's the discoverer Nilson, who independently encounters the element in nature. Each of these subjects says to himself, '`Ekaboron [scandium] is a silvery metal`.' How many different thought contents will we have?

Different externalists will answer that question differently. Some will say that since all the thoughts concern a single substance, there is only a single content. Some will say that since the substance is cognitively presented to the subjects in three different ways, there are three contents. Some might argue that Mendeleev and Nilson think different contents, but Mendeleev's correspondent acquires the ability to think the same contents that Mendeleev thinks; so altogether there are only two contents. I'd *like* our inquiry to apply to all these views, so I hope to avoid taking sides here as much as possible. That forces us to proceed carefully in formulating the McKinsey reasoning.

Consider next a Sort-of-Twin Earth. The Mendeleev there is internally different: he's introduced '`ekaboron`' to refer to the chemically similar element with atomic number 39. But he says all the same things about it in his correspondence; and his friend is an internal duplicate of the Earthly friend. There's also an internal duplicate of Nilson

there, who's happened upon a different metal, yttrium, that turns out to have atomic number 39 and to correspond to the twin-Mendeleev's predictions.

Now how many thought contents do we have? Everyone will say Mendeleev and his twin are thinking different thoughts; after all, they're not even internal duplicates. Externalists will want to say that Nilson and his twin are thinking different thoughts, too, despite being internal duplicates. Some externalists will want to say that Mendeleev's and twin-Mendeleev's correspondents are also thinking different thoughts.

Are the different thoughts had by Nilson and his twin *only available* to subjects whose environments contain samples of the respective elements? That depends on whether their thoughts are the same as Mendeleev's and twin-Mendeleev's thoughts. For plausibly Mendeleev's thoughts *don't* constitutively depend on the environmental presence of the element he's postulating. Mendeleev would have had the same thoughts even if Earth turned out to contain no traces of scandium. So if Nilson's thought has the same content as Mendeleev's thought, then it's not true that that content is only thinkable by subjects whose environments really contain scandium. What may be true is that the content is not thinkable *in the way Nilson thinks it* unless one has genuinely encountered scandium. But it's tricky to say what these "ways of thinking" amount to, without taking sides about issues that externalists disagree about.

And what about the correspondents? Even if we agree that their thoughts differ, and so are externalist, it's plausible that, like Mendeleev, they too would have had the same (deferential) thoughts even if their environment had contained no traces of scandium. What their thoughts seem constitutively to depend on is not samples of scandium but rather communication with Mendeleev and his twin. But here too matters are sticky, since we don't want to take sides on whether their thoughts do or don't have the same contents as Mendeleev's and Nilson's thoughts. The other guys' thoughts don't require the existence of any correspondence.

Here's my attempt to finesse these issues. Let's focus on Nilson and his twin. They each reason in the following way:

(1) I'm now thinking the content *Scandium is a silvery metal.*

(2) I'm understanding this concept *scandium* to be governed by my ostensive introduction, rather than by deference to any authority; and I'm not making any definite assumptions about how it's chemically individuated.

(3) Anyone who's able to think the content *Scandium is a silvery metal* in the way described in (2) must inhabit an environment that does or did contain samples of scandium.

(4) So my environment does or did contain samples of scandium.

Note that (1)–(4) specify *a form* of argument. When Nilson thinks it through, he thinks premises about scandium. When his twin thinks it through, he thinks premises about yttrium. It's prima facie plausible that these subjects should know—or at least be defeasibly justified in believing—premises (1) and (2) just on the basis of introspective reflection. And it's prima facie plausible that they should also know—or at least be defeasibly justified in believing—premise (3) on the basis of a priori, Putnam- and Kripke-style philosophical reasoning. The apparent result is that they acquire some wholly reflective justification to believe (4). That's surprising. (It remains surprising even if it's allowed that their justification to believe the premises, and hence the conclusion, is empirically defeasible.)

Mendeleev's correspondent would instead reason to a different conclusion: perhaps, "So my environment does or did contain other subjects."[11]

III

The McKinsey-style reasoning is a form of *modus ponens*. So one way to block it would be to articulate and defend constraints on when *modus ponens* reasoning is legitimate. Some responses to McKinsey's Puzzle take that form.

One move is to deny Closure. Perhaps Nilson is able to rule out the relevant alternatives to each of (1)–(3), but *more* possibilities are relevant alternatives to (4), and he's not in a position to rule them out. Maybe that means he's able to know (1) through

---

[11] For this version of the argument, see Burge 1982, Brown 1995, McLaughlin and Tye 1998 pp. 312ff, Falvey 2000, and Brown 2001.

(3), without being able to know (4). Maybe. But denying Closure is not so popular these days. Even among epistemologists who employ the framework of "relevant alternatives," most would rather *keep* some form of Closure.[12] In any event, I want to restrict our attention to responses to McKinsey's Puzzle that *don't* require us to deny Closure.

Crispin Wright and Martin Davies have formulated a different sort of constraint, in their discussions of "transmission-failure."[13] Their analyses of the McKinsey argument reward careful consideration; and they interact in interesting ways with the diagnosis I'll be giving. But I'll have to reserve discussion of them for another occasion.

In order to avail ourselves of the McKinsey-style reasoning, we need to have reflective justification to believe all its premises *simultaneously*. Sometimes *modus ponens* reasoning fails because in acquiring justification to believe one of the premises, one loses justification to believe another. For example, I may start out thinking, quite reasonably, that you are not a rap star. The only time I ever heard *you* rap was when we did karaoke together. So I justifiably think *If you're a rap star, then so am I*. But then you confront me with evidence of your covert musical career. I'm not now justified in affirming the antecedent and concluding that I'm a rap star too. This is because in acquiring justification to believe *You are a rap star*, I *lost* my grounds for believing the conditional *If you're a rap star, then so am I*.[14]

We should take care that no such funny business is going on with the McKinsey-style reasoning. At the moment, there is no special reason to think it *is*. In the McKinsey-style reasoning, your justification to believe conditionals like:

McK-2 If you're thinking a thought with the content *Water puts out fires*, then your environment is so-and-so.

---

[12] For discussion, see Stine 1976, Brueckner 1985, Cohen 1988, Vogel 1990, DeRose 1995 esp. §10, Klein 1995, and Hawthorne 2005.

[13] See especially Wright 2000 and 2003, and Davies 1998, 2000, 2003a and 2003b.

[14] Compare Harman 1973 Chapter 9; Ginet 1980; and Sorenson 1988.

comes from reflecting on a priori thought-experiments. Why should learning that the antecedents of these conditionals are true—that you are thinking thoughts with the indicated contents—defeat or undermine your grounds for believing the conditional? At the moment, we have no ground for saying it would. But we will return to this possibility later.

For the time being, then, the way looks clear for you to legitimately combine your justification to believe premises (1) through (3), and thereby acquire justification to believe the surprising conclusion (4).

IV

But is your justification for believing all the premises really wholly reflective? We need to think about this more carefully.

So far we've been thinking about Nilson and his twin. They inhabitant different environments—one where scandium is distributed in certain Scandinavian minerals, the other where yttrium is so distributed—and as a result they end up thinking different thoughts. But their environments have it in common that there's *some* substance of the appropriate sort that they're interacting with. Let's call any environment of that sort a **hospitable environment**. Nilson will also have duplicates whose environments are **inhospitable**. These will be places like Boghossian's Dry Earth.[15] Nilson's unlucky duplicate there will just be *hallucinating* handling some minerals, isolating a metallic oxide, and so on. From the inside, everything will seem to him just as it seems to the real Nilson.[16] But outside, there's no substance there for his thoughts to latch onto. What *will* the Nilson in that inhospitable environment be thinking, when he says to himself, '`Scandium is a silvery metal`'?

There are a variety of answers one might give here.

Perhaps he's thinking superficial descriptive thoughts: thoughts true in case some new metallic element he's just identified in such-and-such minerals is a silvery metal.

---

[15] See Burge 1982, 114ff; and Boghossian 1997.

[16] Though see fn. 1 for some difficulty with this.

Now, no doubt *he will* be thinking such descriptive thoughts; and so too will the real Nilson. What's controversial is whether they're the thoughts that unlucky Nilson has *in* saying to himself '`Scandium is a silvery metal`.' It's reasonable that they might be. On this proposal, Nilson would be thinking descriptions that nothing in his environment *manages to satisfy*; but the *descriptive thoughts* would be available to him, despite the inhospitality of his environment.[17]

A different proposal is that, in saying to himself '`Scandium is a silvery metal`,' Nilson is unwittingly having thoughts about his own experiences. This is what projectivists say about our color concepts.[18] As it turns out, they say, our environment is inhospitable to those concepts. There are no qualities "out there" of the right sort for our color concepts to latch onto. Instead, when we think about colors, we're really thinking about our own experiences; we just wrongly *project* qualities of those experiences onto the outside world.

Another proposal is that Nilson is unwittingly thinking about a *fictional* metal—in the same way that our thoughts about Tolkein's *mithril* are thoughts about a fictional metal. Or perhaps he's thinking about a necessarily uninstantiated substance.[19] Or perhaps his thoughts have a special, "gappy" content that is incapable of ever being true.[20]

---

[17] Boghossian 1997 argues that the Dry Earthers' ***`water`*** concept cannot be a "compound, decompositional concept." I agree. Our complex description '`the new metallic element I've just identified in these minerals`' shouldn't be understood as sharing its *logical form* with the Dry Earthers' term '`water`.' I'm just using it to specify their term's *intension*. That may be the best we can do; we might not *have* a term with the same intension *and* logical form. Boghossian seems to argue further that, if we don't ourselves have a way *to say* what concept is expressed by the Dry Earther's term '`water`,' then there is no fact of the matter what its content and intension are. Here I'm unconvinced. See McLaughlin and Tye 1998 §VI for discussion.

[18] See, e.g., Boghossian and Velleman 1989.

[19] Consider Shoemaker's "NI-intentional properties" in his 1990; and see also Stoneham 1999.

[20] On these "gappy" contents, see Adams, Fuller and Stecker 1993; Adams and Stecker 1994; Braun 1993 and 2005; Salmon 1998; and Reimer 2001a and 2001b.

Or perhaps, following McDowell and Evans on demonstrative thoughts, Nilson doesn't manage to have *any* contentful thought in saying to himself '`Scandium is a silvery metal.`'[21] He may well be having various thoughts at the same time: perhaps some of the descriptive thoughts we've already identified. But none of those will be what he's thinking *in* saying those words to himself. Nothing will be. It will just falsely seem to him that he's thereby thinking anything contentful.

As I said, there are variety of things one can say here. Keep in mind that these are all proposals about what thoughts *unlucky* Nilson is having. We shouldn't have doubts about the extension of *our own* thoughts, or the real Nilson's thoughts, with respect to the inhospitable environment. Our thoughts are about scandium, a substance that happens not to be present there. Nilson's unlucky counterpart isn't in a position to be thinking *those* thoughts. We're trying to figure out what *he is* thinking, in their place.

Some philosophers will maintain that the contents unlucky Nilson thinks are *also* contents the real Nilson thinks in saying to himself '`Scandium is a silvery metal.`' Internalists say *that's all* that the real Nilson thinks. But an externalist can say the real Nilson thinks these contents too—*in addition to* thoughts about the particular natural kind he's interacting with. Externalists disagree about whether the real Nilson's thoughts do have any internally common factors of this sort. I will not here take a stand on the question.

If the unlucky Nilson *is* thinking things that the real Nilson isn't *also* thinking in saying '`Scandium...`' to himself, then by our definition of "externalist," the unlucky Nilson's thoughts are externalist too. For he has internal duplicates (the real Nilson and his duplicate on Twin Earth) who fail to think those same contents.

However, there's an interesting question nearby that is still unsettled: Are the contents unlucky Nilson is thinking *available* to the real Nilson? Perhaps the real Nilson needs to use different words to formulate those thoughts—e.g., perhaps he wouldn't be thinking them in saying to himself '`Scandium...`' though he would be thinking them in

---

[21] See McDowell 1977, 1984, and 1986; and Evans 1981 and 1982. Consider also McGinn 1989's "strong externalism."

saying to himself '`The new metallic element I've just identified in these minerals...`'

On some of the proposals about what unlucky Nilson is thinking, it's plausible that his contents *would* be available in this way to the real Nilson. On other proposals, this may be less plausible. To take one example, consider the proposal that unlucky Nilson is thinking about a fictional substance. Perhaps the relevant fiction is not cognitively available to the real Nilson. Kripke argued that the thoughts we express with '`Sherlock Holmes`' are about an essentially fictional detective. A world where some real person lives at 221B Baker Street, is called '`Holmes`,' and does such-and-such wouldn't be a world where *our* Holmes really exists. That real detective may be beyond our referential reach.[22] Conversely, one might argue that *our* fictional Holmes is beyond the referential reach of the real detective and his neighbors. They're not appropriately acquainted with our fiction. So too, if unlucky Nilson is thinking about a fictional element when he thinks to himself '`Scandium...`,' he may be thinking a content that we're not ourselves in a position to think. At any rate, that's a view that could be argued.

Similarly, if unlucky Nilson is thinking a *gappy* content, and gappy contents are sufficiently fine-grained, then this too may be a content that's unavailable to us using any words. That's another view that could be argued.

So, collecting these possibilities together: perhaps unlucky Nilson is thinking contents that the real Nilson is *already* also thinking; or perhaps he's thinking contents that the real Nilson *could* think, using other words; or perhaps he's thinking contents that are strictly unavailable to the real Nilson. We also mooted a different possibility: on a McDowell/Evans-inspired view, unlucky Nilson isn't thinking any contentful thought at all—at least not *in* saying to himself '`Scandium is a silvery metal`.' We'll consider this last possibility carefully in later sections; set it aside for now.

We need to keep this variety of options in mind when judging the epistemic status of the conditional McKinsey premise:

---

[22] See Kripke 1980, pp. 157-8.

(3) Anyone who's able to think the content *Scandium is a silvery metal* in the way I do must inhabit an environment that does or did contain samples of scandium.

Unlucky Nilson will be thinking a premise of this form to himself. What he's thinking will probably be false. The content *he's* demonstrating by thinking to himself '`Scandium is a silvery metal`' *is* available to subjects (like himself) whose external environments never did contain any new metallic element, any of his own experiences, any essentially fictional or necessarily uninstantiated metal, or whatever it is he refers to with '`scandium`.'

How does that bear on the epistemic status of the lucky Nilsons: the one in our environment and the one on Twin Earth? Presumably, they're not in any position to know a priori that they're the lucky ones. For all they know a priori, *they* may be in inhospitable environments. If they *are* in such environments, the premises of form (3) that *they're* thinking would be false too. So it's doubtful that they *do* know these premises a priori. They're only in a position to know a priori premises of *this* form:

(3*) *If* my environment is hospitable, *then* anyone who's able to think the content *Scandium is a silvery metal* in the way I do inhabits an environment that does or did contain samples of scandium.

But that only enables them to draw conclusions about what their environment is like if it's hospitable. And it's not that surprising that one should be able to know *things like that* purely by reflection. Alternatively, perhaps they're in a position to know a priori:

(3**) The content *Scandium is a silvery metal* seems or *purports* to be a content that, when thought in the way I think it, is only available to subjects in environments that do or did contain samples of scandium.

But that only enables them to draw conclusions about what their environment purports to be like. And that too may well be intuitively allowed to be knowable by reflection. The McKinsey result only manages to be *puzzling* when it stays close to the conclusion we originally formulated.

We've just rehearsed one popular way of defusing McKinsey's Puzzle.[23] I don't think it can be a *complete* solution—as we'll see, we haven't yet covered all the bases—but it's right as far as it goes. It's often an a priori open possibility for subjects that they're in inhospitable environments, and that *if they are* in such environments, the thoughts they're having *aren't* restricted to subjects in environments containing the relevant stuffs (or genuine experts). So they won't be in a position to know the conditional McKinsey premise a priori, after all.

Is this a form of Incompatibilism? Are we saying that the premises of McKinsey's argument can't all be true?

No. As I understand Incompatibilism, it's the thesis that it can't be true *both* that your thought has an externalist content *and* that you're able to know your thought has that content, just by reflection. We're not saying that. On the current account, for instance, the real Nilson's thought—the thought he's reflectively aware of having—*does have* an externalist content. The conditional McKinsey premise really is true of him. He's just not in a position to know a priori that that's so.

But isn't this still Incompatibilist in a sense? Aren't we granting that if externalism is true, you *can't* after all know by reflection alone *which* thought you're having—is it the thought about scandium? the thought about yttrium? or an unlucky thought about a merely fictional stuff? You can't tell.[24]

It is part of this account that you can't know some philosophically interesting things about your thoughts—namely, whether premise (3) is true of them—by reflection alone. I think it's plausible that whatever ***scandium*** concept you have can survive *learning* whether (3) is true (or false) of it.[25] We are here giving up on having complete reflective knowledge of our thoughts' nature. But that's something an internalist might give up too. (For example, perhaps it's part of the nature of even internalist thoughts that they be

---

[23] See Gallois and Hawthorne 1996, and McLaughlin and Tye 1998, esp. §IV.

[24] See here Boghossian 1994.

[25] Compare: before learning the facts about jade, we believed ***Jade is a single natural substance***; after learning the facts, we denied that very same belief.

materially realized in such-and-such ways; but that isn't knowable just by reflection.) Even when you fail to know something's complete nature, you might still know *what* that thing is. For example, even if you fail to know the complete nature of the metal that composes your ring, you might still know what metal it is. You might *know that* it's gold your ring is made of. McKinsey's argument threatens your having even this kind of epistemic achievement with respect to your thoughts, on the basis of reflection. It threatens your being able to tell by reflection *that* you're thinking *Water puts out fires*, or *that* you're thinking *Scandium is a silvery metal*. The present account defuses this threat. It says: Yes, you can know you're thinking those contents. You just don't know a priori what constraints the thinking of them places on a subject's environment.

Consider this analogy. You're a newcomer to Metropolis. On Monday, you meet newspaperman Jimmy Olson. On Tuesday you meet his colleague Clark Kent. On Wednesday you witness crimes being stopped by a red-caped superhero the locals call '`Superman`.' You don't realize you already met him in disguise the day before. Now are you in a position to know a priori that Clark Kent = Superman? Most philosophers will say no. A few will say that the proposition that Clark Kent = Superman is identical to the proposition that Superman = Superman, and that *you are* in a position to know that a priori. However, *all* theorists should agree that you're not yet in a position to know that *Jimmy Olson* ≠ Superman. That's just not something you could know without doing some investigation and gathering certain kinds of evidence. However, it's no obstacle to your knowing who you're talking to, when you're talking to Jimmy Olson. You can know you're talking to Jimmy without knowing everything interesting about him: his passport number, that he's distinct from Superman, and so on. (*On Monday*, you were doubly unable to know that Jimmy Olson ≠ Superman: in the first place because you hadn't done the necessary investigation, and in the second place because you hadn't yet encountered Superman, even in disguise, and no one had mentioned him to you. So you weren't yet capable of thinking *anything* about him.)

Your situation with respect to *water* and *scandium* thoughts is much the same. You can know by reflection that you're thinking this content, *Scandium is a silvery metal*, without knowing everything interesting about it—that it's about a substance with 21 protons, that it's not about yttrium, that it's about a real substance rather than a

fictional one, and so on. (Here, too, you're *doubly* unable to know your thoughts don't concern yttrium: in the first place because you haven't done the necessary investigation, and in the second place because you're not yet in a position to think anything about yttrium.)[26]

V

Let's return to the McDowell/Evans view that we set aside. On that view, the unlucky Nilson isn't thinking any thought when he says to himself '`Scandium is a silvery metal`.' What if the real Nilson knew a priori that *that's* the correct account of the cognitive life of subjects in inhospitable environments? Then he'd know a priori that anyone who *does* manage to think contentfully *is* in a hospitable environment. Hence, he apparently *would* be able to know a priori that:

(3) Anyone who's able to (genuinely) think the content *`Scandium is a silvery metal`* in the way I do (or seem to) must inhabit an environment that does or did contain samples of scandium.

So the surprising McKinsey result is still with us.[27] And we wouldn't get off the hook just because Nilson *failed to know* that the McDowell/Evans account is correct. If he merely has some undefeated a priori *justification to believe* it's the right account, that's bad enough. For then he'd have some a priori justification to believe (3) is true, and as a

---

[26] Compare Falvey and Owens 1994's contrast between introspective knowledge of content and introspective knowledge of comparative content.

These issues are why I was so careful about how we're identifying contents in our discussion. If all you were in a position to know was *`Whatever content I think by saying such-and-such is so-and-so`*, then the complaint that you don't know what you're thinking might have some teeth. As it is, though, I think you're able to think these contents, apprehend and demonstrate them in doing so, and know by reflection *that* you're thinking them, despite not knowing whether premise (3) is true of them.

[27] Boghossian 1997 stresses this. Wright and Davies also doubt that we can count on the strategy from §iv, above, always being available. They characterize themselves as giving responses to the "worst-case scenarios," where (3) *is* knowable a priori.

result, some reflective justification to believe the McKinsey result. That seems just as counter-intuitive as saying he could reflectively *know* the McKinsey result. It's counter-intuitive that Nilson should be able to get any justification at all about his environment *in this reflective way*. And that prospect is still on the table, so long as he has any undefeated a priori justification to think the McDowell/Evans view might be correct.

We need to examine this possibility closely. In handling it, I'm going to make use of apparatus that I've argued for in other papers.[28] So I'll begin by summarizing the points I want to import.

In those other papers, I argued that just because a certain type of thought is true whenever it's entertained, it does not follow that subjects thereby have a priori justification to believe it's true. On a McDowell/Evans-style view, for instance, thoughts of the form:

(5) Jack exists.

and

(6) I perceive *this wall* [here you perceptually demonstrate a wall].

will be true whenever they're successfully entertained.[29] If Jack didn't exist, or you weren't perceiving a wall, then your referential attempts would fail and on this view you wouldn't be thinking contentful thoughts of the specified forms. But plausibly (5) and (6) *aren't* the sorts of thing you can know a priori. (In our present discussion, it's under dispute whether you can know them *by reflection*: that is, relying both on a priori sources and your introspective awareness. But in those papers, and for the moment here, I'm just talking about whether these claims are knowable strictly a priori.)

---

[28] Pryor 2006a and 2006b.

[29] Here some issues we tried to side-step in §ii re-arise. Millians will argue that thoughts with the same *content* as (6) can be entertained falsely: e.g., if you refer to the wall using a name acquired from other speakers, and aren't then perceiving it. Many Millians will grant that that would amount to entertaining (6) in a different "way" than I'm envisaging in the text. If so, my talk of "thoughts of such-and-such a form" should be read to include not just the thought's content, but also a relevant way of thinking it.

On the other hand, I argued that "hedged" thoughts like:

(7h) If Jack exists, then he is self-identical.

may be knowable a priori. Not because they're true whenever they're entertained; but rather through your understanding of the logical relations they involve. Here matters get somewhat subtle. On the McDowell/Evans view, you won't be able to entertain (7h) unless Jack exists, too. But I argued for a distinction between *what has to be the case for a thought to have a given content*, and *what are the proper logical entailments of the thought*. I think the relevant logic for these thoughts is a free logic on which (7h) *doesn't* logically entail (5). The virtue of such a view is that it can do justice to our pre-theoretic impression that:

- experience plays some role in justifying (5); and
- it doesn't seem rational to believe (5) to degree 1.

Whereas it seems that:

- there *should* be some purely logical truth, along the lines of (7h), that we could rationally believe just on the basis of our logical understanding; and
- that we could have more rational confidence in than we have in (5).

So on the view I proposed, (7h) did not logically entail (5). However, for a McDowell/Evans-inspired theorist, you wouldn't be able to *contentfully think* (7h) unless (5) were true. Still, I argued, if (5) *were* true, and (7h) thinkable, then your logical understanding of (7h) could give you a priori justification to believe *it*, without giving you a priori justification to believe that *you're thinking it*; nor that (5) is true. You can be in a position to think the thought, and have a priori justification for it, without being in a position to tell a priori that that is so. It takes *more* to justify you in believing you *really have* the thought than it takes to justify the thought itself.

When it's an open question for you whether Jack exists, it may seem peculiar to try to think thoughts like (7h), without yet knowing that you'll succeed. (It's somewhat like prefacing an email with "If you're still reading email at this address …" That insures you no better against not being read.) However, you don't really have better alternatives. Your ability to *withhold belief* in (7h), or *be agnostic* towards (7h), is just as threatened by the prospect of Jack's not existing as full belief towards (7h) would be. There is such a condition as *having no degree of belief at all* towards a thought: e.g., if you've never

considered the thought, or are incapable of considering it. But when you *are* considering (7h)—as it happens, successfully—then you'll *have to* have some cognitive attitude towards it that wouldn't be available were Jack not to exist. You just won't be in a position to know a priori that you do.

Now, experiences will be necessary for you to be able to think either of (5) or (7h). But sometimes the experiences that enable you to think a thought are part of what justify you in believing it; and other times they merely enable the thinking. Compare: in order to have the concept *chromatic*, and believe:

(8) Chromatic colors are colors.

perhaps you need to have had visual experiences. But plausibly visual experiences play no role in justifying your belief. On the other hand, in order to be able to think the demonstrative thought

(6) I perceive *this wall* [here you perceptually demonstrate a wall].

you also need to have perceptual experiences. And here the experiences *are* plausibly part of what justify your belief. I maintain that *Jack exists* is a posteriori, like (6), whereas *If Jack exists, then he's self-identical* is a priori, like (8). From the premise *that* you're thinking this latter thought, you may be able to infer a priori that Jack exists. But from the thought itself, you cannot.

I've been talking so far about a priority, strictly construed. What I've said leaves it open whether we can know *Jack exists* and *I perceive this wall* by *reflection*. One view will say you need ordinary *perceptual* justification to believe the wall exists, *in order to* be able to understandingly think the demonstrative thought (6). Similarly, you'll need ordinary perceptual evidence about Jack to even understand the hypothesis that he exists. A different view will say that you can be justified in believing these thoughts solely from your introspective awareness of seeming to successfully think them, and your a priori knowledge that thoughts of these types need referents. That is, you could have McKinsey-style, wholly reflective, justification to believe them. My sympathies lie with the former view. But at this stage, we still need to seriously consider both alternatives.

VI

Armed with that background, we're now in a position to complete our diagnosis of the McKinsey-style reasoning.

In some ways, the apparatus I've introduced *aids* the McKinsey-style reasoning. For we suggested that on a McDowell/Evans-style view, the premise:

(3) Anyone who's able to (genuinely) think the content *Scandium is a silvery metal* in the way I do (or seem to) must inhabit an environment that does or did contain samples of scandium.

will be knowable a priori. And you might balk at that, thinking that if for all you know a priori, you're in an inhospitable environment and thinking nothing contentful by saying those words to yourself, then you *can't* know this premise a priori—even if, as it happens, you're lucky and *are* contentfully thinking the premise. The apparatus I've introduced makes it possible to finesse that worry. We'll replace (3) with the "hedged" claim:

(3h) If the content *Scandium is a silvery metal*, which I hereby seem to be thinking, really exists, then anyone who is able to genuinely think it in the way I seem to must inhabit an environment that does or did contain samples of scandium.

If the McDowell/Evans-style view is correct, there's still no contentful thought of this form for unlucky Nilson to have. But if the lessons I summarized are right, then such a hedge will give us a form of premise that subjects in hospitable environments will be able to think *and* may well have a priori justification for. Unlike the hedges (3*) and (3**) we considered earlier, this hedge doesn't yet obviously handicap the McKinsey Argument.

Premise (2) can be hedged in a similar way. We'll discuss premise (1) momentarily.

There are various things Nilson knows by reflection that fall short of premise (1), and are *not* enough, given what else he knows by reflection, to infer that his environment contains samples of scandium. For instance, he can know by reflection *I seem to be thinking a thought by rehearsing to myself the sentence* '`Scandium is a silvery metal.`' He can know by reflection *I seem to be having a thought that attributes*

being a silvery metal to some substance. These are general, non-externalist thoughts. Their contents don't even purport to be constitutively dependent on scandium; so they are thoughts that even the unlucky Nilson in his inhospitable environment can be in a position to think, and to know by reflection. As I said, they're too weak to provide knowledge by reflection of a McKinsey-style conclusion.

However, if they're *all* we can know about our thoughts by reflection, we should be disappointed. On a McDowell/Evans-style view, I think *they are* all that unlucky Nilson can know. But the real Nilson, and his duplicates in hospitable environments, can do better.

Let's consider a hedged version of premise (1):

(1h) If the content *Scandium is a silvery metal* really exists, then I am now thinking it.

The original premise (1) would amount to the combination of (1h) and:

(1∃) The content *Scandium is a silvery metal* really does exist.

On a McDowell/Evans-style view, neither (1h) nor (1∃) will be contentfully thinkable unless you're in a hospitable environment. But if you *are* lucky enough to be able to think them, then I think it's plausible that (1h) should be justified for you just on the basis of your introspective awareness of your occurrent mental life. You don't need first to *establish that* your environment is hospitable. And it's (1h) *itself* you have justification for: not just a metalinguistic claim, like *If (1h) expresses a thought, then what it expresses is true*. (1h) itself constitutes a more intimate and satisfying reflective knowledge of what you're thinking than what we considered before.

What about (1∃)? Is it *also* knowable by reflection? That's not so clear. We're supposing it to be an open epistemic possibility that you're having a McDowell/Evans-style illusion of contentful thought. That suggests that *you won't* be able to know (1∃), not until you acquire some evidence that you're really one of the lucky subjects in a hospitable environment. I'm not saying you need justification to believe your environment is hospitable in order to *think* *Scandium is a silvery metal*. Neither do you need it in order to be *justified in believing* any *scandium* thought: the thought *If scandium exists, it's self-identical* may be justified a priori. What I'm saying is that—to the extent

that McDowell/Evans-style illusions are an open possibility—you need antecedent justification to believe your environment is hospitable to be justified in believing that *you're genuinely thinking* ***scandium*** thoughts. That sacrifices *some* reflective knowledge about your own mind, but it's a sacrifice I think we can be comfortable with. We can leave your justification to believe (1h) in place. Like the claim ***If Jack exists, then he's self-identical***, I think (1h) is something you can have justification to think *without* yet having justification to believe you *can think* it. The source of your justification to believe (1h), and the degree to which you're justified in believing it, may be different from the source and degree to which you're justified in believing you do successfully think (1h).

For the moment, then, here's what I propose. Lucky Nilson is (perhaps without knowing it) in a position to think premises (1h), (1∃), and so on. His introspective awareness of his thoughts justifies him in believing (1h), that he's *purporting* to think a certain content: namely this one [here he mentally produces the content ***Scandium is a silvery metal***]. Nilson doesn't know by reflection *whether* he's succeeded in thinking anything. But in fact *he has* succeeded, and introspection does justify him in believing that *that* is what he's thinking, if he's thinking anything. Because Nilson can't tell by reflection alone that he is successfully thinking, he's not in a position to infer via McKinsey reasoning that his environment is a hospitable one. That takes further, empirical evidence.

The relevant empirical evidence will be readily had: the same experiences that enabled Nilson to think about scandium, and apply the name '`scandium`' to it, will also justify him in believing it really is present in his environment.[30] So his thought ***My environment contains scandium*** will be like the thought we considered earlier, ***I perceive this wall***. Assuming you have no defeating or undermining evidence,[31] no *more*

---

[30] Brewer 2000, pp. 428-9 says this too. And he seems to think that's the full story: our justification for believing (1) is wholly and univocally empirical. As you'll see in a moment, I think the story is less straightforward.

[31] What if you *do* have defeating or undermining evidence? Consider a case where scandium does exist, and you see it, but you have overwhelming evidence that you're hallucinating. In such a case, you might think, falsely, that ***That stuff*** [ostending what you take to be a hallucination, but is in fact some

experiences may be needed, to justify you in believing the thoughts, than it takes to be able to think them. Nonetheless, your experiences in these cases are playing an essential justifying role; and so your justification will be perceptual rather than reflective.

Now, the account I've just sketched is not mandatory. Though we may be comfortable saying you can't *know* (1∃) by reflection alone, we may feel uneasy saying you don't even have reflective justification to believe (1∃). As I suggested at the end of §v, a different view is that your introspective awareness of seeming to successfully think a thought does give you some justification to believe *there is* a content that you're thinking. On that view, you *would* have some reflective justification to believe premise (1∃). And if you continued to have reflective justification to believe (2) and (3h), then it looks like you might achieve reflective knowledge, or reflective justification to believe, that your environment contains scandium, after all.[32]

But let's proceed carefully. There's plausibility to the idea that introspection gives you some justification to believe (1∃). There's also plausibility to the idea that you can have reflective justification to believe that the McDowell/Evans view, and hence premise (3h), are correct. But in order for the McKinsey-style reasoning to carry through, you need reflective justification to believe *the conjunction* of these premises. As we saw in §iii, sometimes justification to believe one premise in a modus ponens argument defeats or undermines your justification to believe another. I think something like that is happening here.

---

scandium] ***doesn't exist***. If you believed the McDowell/Evans-style view to be correct, you might also think, falsely, that you just then failed to think any contentful demonstrative thought. You may even have all things considered *justification* to believe those things: to believe of scandium that it doesn't exist, *and* that your attempts to think so are unsuccessful. You may know, too, that *if* your attempts to think so *were* successful, they'd have to be false. This is delightfully perverse, and all merits closer discussion. But I see no paradox.

[32] This is more or less the position Wright thinks we're in with respect to *some* versions of McKinsey's Argument: see Wright 2003 §V.

If having a McDowell/Evans-style illusion of contentful thought is an open possibility for you, then presumably any justification you have to believe (1∃) will be *defeasible*. One kind of evidence that would contribute towards its defeat would be evidence that the McDowell/Evans view is correct and you *really are in* an inhospitable environment. The more confident you became that that was so, the less confident you could rationally be in (1∃). That's one kind of defeating evidence. The analogue in the perceptual case would be getting evidence that you're hallucinating things other than as they are. In the perceptual case, there are also other kinds of defeating evidence. For example, there's evidence merely that you're dreaming, or evidence merely that your eyes don't work. These undermine your entitlement to think *you're perceiving* your environment, without (directly) justifying you in believing that things are other than they appear. Now, consider evidence that your eyes are so configured that it would look to you as though you have hands regardless of whether you really do have hands. This isn't (directly) evidence that you lack hands. Neither is it (directly) evidence that you *are* right now failing to perceive. But plausibly, it too will to some degree undermine your perceptual justification to believe you have hands.

I think the McDowell/Evans view undermines your introspective justification to believe (1∃) in the same way. The more confident you become that the McDowell/Evans view is correct, and that inhospitable environments induce illusions of contentful thought, the less entitled you are to think that introspection *tells you* that an episode of seeming-to-think is an episode of really-thinking.[33] This is so even *in advance* of your getting evidence to think *you are* in an inhospitable environment.

So I will agree that your introspective awareness of seeming to successfully think a thought gives you *some* justification to believe *there is* a content that you're thinking. That justification is ordinarily seamless: you don't have to take any inferential step from *This is what I seem to be thinking* to *This is what I really am thinking*. No more than you have to take any inferential step from *This is how things look to me* to *This is how*

---

[33] Wright 2000 §VI also entertains (but in the end rejects) the possibility that your theoretical justification to believe the externalist premise might "compromise" your justification to believe the introspection premise.

things are. In both cases, though, additional evidence can undermine your justification to believe the stronger claim, and leave you merely with justification to believe the weaker.

This undermining will usually be a matter of degree. Usually your situation will be one of having *some* degree of rational confidence that you are genuinely aware of real thoughts, and a real perceptual environment, and *some* degree of rational confidence that introspection alone, or your unaided senses, lack the authority to tell you that's so. As the balance of evidence changes, you will seesaw between these alternatives.[34]

Applying this to our present discussion means you will seesaw between having reflective justification to believe not just (1h) but (1∃) too, on the one hand, and having reflective justification to believe (3h), and that it's beyond the authority of introspection *to tell you* that you're having a genuine thought, on the other. You'll usually have some degree of reflective justification on each side. But that just means you're superposed between two epistemic situations, neither of which by itself justifies you in believing all the premises the McKinsey-style reasoning requires. I don't see any persuasive reason to think you ever have reflective justification to believe *the conjunction* of (1∃) and (3h). Until you do, you're in no position to draw inferences about your environment, by reflection alone.

VII

Let me close with one final wrinkle. Consider a mathematician who carefully reasons through a proof that has some subtle undetected flaw. Is the mathematician justified in believing the conclusion? Well, perhaps he has some inductive justification: he knows

---

[34] This is not precisely the same as the scenario we discussed in §iii. There you started with justification to believe a conditional, that was undermined by the evidence you got for its antecedent. Here I expect you start out with reflective justification to believe the antecedent—(1∃)—and that gets undermined by the reasons you get to believe the conditional—(3) or (3h). The cases are interesting similar.

Notice that it's not (3) or (3h) *itself* that's doing the undermining. *Empirical* evidence that your environment is hospitable, and hence that your *scandium* thoughts are only available to other subjects whose environments contain scandium, too, would not undermine your ability to tell by reflection that you're having genuine thoughts. Rather, it's *the particular philosophical ground the McDowell/Evans view gives you* for believing (3h) that's the villain.

that proofs in that journal, and proofs he finds compelling, have a good track-record. But set that aside. Does he have any a priori or at least *reflective* justification to believe the conclusion? Does his experience of seeming to deduce the conclusion give him any justification for it?

I don't know what's the best thing to say about that. I can be persuaded either way. Suppose we decide to say that reasoning through a flawed proof *can* sometimes give one reflective justification to believe its conclusion. We might say that one has reflective justification, but it's not *impeccable*. It has some flaw the discovery of which would undermine the justification. But we may want to allow that, if you're reasonably ignorant of the flaw, you do have reflective justification.

If that's what we decide to say, then we should probably say the same about McKinsey's argument, too. I've argued that the McKinsey-style reasoning *is* flawed. However, subjects may not have seen these flaws. They may find the reasoning pretty compelling, and perhaps, given their unenlightened state, they're even reasonably entitled to do so. If we allow the mathematician to have reflective justification to believe his conclusion, then we ought also to allow these unenlightened subjects to have reflective justification to believe McKinsey-style conclusions about their environments. What they'll lack is *impeccable* reflective justification. That's what I've intended to be discussing throughout.[35]

ACKNOWLEDGMENTS: Versions of this paper have been circulating since 2000. I've benefited from useful comments at Rochester, Penn, Stanford, Brown, Utah, Barcelona, Kentucky, St Andrews, UNAM, CUNY, Indiana, and a conference in Tlaxcala. Special thanks to Miguel Fernández, John Collins, Henry Jackson, and Skip Larkin for their detailed responses. I got further useful advice from the URG and the Eminees reading groups in Cambridge; from my grad seminars at Princeton and NYU; and from Richard

---

[35] Compare §6 of Pryor 2004. I was aided in coming to this view by discussions with Stephen Schiffer of his 2005.

Heck, Alison Simmons, Mark Johnston, Scott Soames, Gideon Rosen, and Stephen Schiffer. I appreciate all their good advice, despite not being able to implement all of it.

Bibliography

Adams, F., Fuller, G. and Stecker, R. (1993) "Thoughts without Objects" *Mind and Language* 8, 90-104.

Adams, F. and Stecker, R. (1994) "Vacuous singular terms" *Mind and Language* 9, 387-401.

Boghossian, P. (1989) "Content and self-knowledge" *Philosophical Topics* 17, 5-26. Reprinted in Ludlow and Martin, eds. (1998), 149-73.

Boghossian, P. (1994) "The transparency of mental content" *Philosophical Perspectives* 8, 33-50.

Boghossian, P. (1997) "What the externalist can know a priori" *Proceedings of the Aristotelian Society* 97, 161-75. Also appears in C. Wright, M. Smith and C. Macdonald, eds., (1998) *Knowing Our Own Minds: Essays in Self-Knowledge*. Oxford Univ. Press, Oxford, 271-84.

Boghossian, P., and Peacocke, C., eds. (2000) *New Essays on the A Priori*. Oxford Univ. Press, Oxford.

Boghossian, P. and Velleman, D. (1989) "Color as a secondary quality" *Mind* 98, 81-103.

BonJour, L. (1998) *In Defense of Pure Reason*. Cambridge Univ. Press, Cambridge.

Braun, D. (1993) "Empty names" *Nous* 27, 449-69.

Braun, D. (2005) "Empty names, fictional names, mythical names" *Nous* 39, 596-631.

Brewer, B. (2000) "Externalism and a priori knowledge of empirical facts" in P. Boghossian and C. Peacocke, eds., *New Essays on the A Priori*. Oxford Univ. Press, Oxford, 415-32.

Brown, J. (1995) "The incompatibility of anti-individualism and privileged access" *Analysis* 55, 149-56. Reprinted in Ludlow and Martin, eds. (1998), 185-93.

Brown, J. (2001) "Anti-individualism and agnosticism" *Analysis* 61, 213-24.

Brown, J. (2004) *Anti-Individualism and Knowledge*. MIT Press, Cambridge, Mass.

Brueckner, A. (1985) "Skepticism and epistemic closure" *Philosophical Topics* 13, 89-117.

Brueckner, A. (1986) "Brains in a vat" *Journal of Philosophy* 83, 148-67.

Brueckner, A. (1992) "What an anti-individualist knows a priori" *Analysis* 52, 111-18. Reprinted in Ludlow and Martin, eds. (1998), 197-205.

Burge, T. (1982) "Other bodies" in A. Woodfield, ed., *Thought and Object*. Oxford Univ. Press, New York, 97-120.

Cohen, S. (1988) "How to be a fallibilist" *Philosophical Perspectives* 2, 91-123.

Davies, M. (1998) "Externalism, architecturalism, and epistemic warrant" in C. Wright, M. Smith and C. Macdonald, eds., *Knowing Our Own Minds: Essays in Self-Knowledge*. Oxford Univ. Press, Oxford, 321-61.

Davies, M. (2000) "Externalism and armchair knowledge" in P. Boghossian and C. Peacocke, eds., *New Essays on the A Priori*. Oxford Univ. Press, Oxford, 384-414.

Davies, M. (2003a) "The problem of armchair knowledge" in S. Nuccetelli, ed., *New Essays on Semantic Externalism and Self-Knowledge*. MIT Press, Cambridge, Mass., 23-55.

Davies, M. (2003b) "Externalism, self-knowledge, and transmission of warrant" in M. Frápolli and E. Romero, eds., *Meaning, Basic Self-Knowledge, and Mind: essays on Tyler Burge*. CSLI Publications, Stanford, 105-130.

DeRose, K. (1995) "Solving the skeptical problem" *Philosophical Review* 104, 1-52.

Dretske, F. (1995) *Naturalizing the Mind*. MIT Press, Cambridge, Mass.

Evans, G. (1981) "Understanding demonstratives" in H. Parret and J. Bouveresse, eds., *Meaning and Understanding*. Walter de Gruyter, New York. Reprinted in G. Evans (1985) *Collected Papers*. Oxford Univ. Press, Oxford, 291-321.

Evans, G. (1982) *The Varieties of Reference*. J. McDowell, ed. Oxford Univ. Press, Oxford.

Falvey, K. (2000) "The compatibility of anti-individualism and privileged access" *Analysis* 60, 137-42.

Falvey, K. and Owens, J. (1994) "Externalism, self-knowledge and skepticism" *Philosophical Review* 103, 107-37.

Gallois, A. and Hawthorne, J. (1996) "Externalism and skepticism" *Philosophical Studies* 81, 1-26.

Ginet, C. (1980) "Knowing less by knowing more" *Midwest Studies in Philosophy* 5, 151-61.

Harman, G. (1973) *Thought*. Princeton Univ. Press, Princeton.

Hawthorne, J. (2002) "Deeply contingent a priori knowledge" *Philosophy and Phenomenological Research* 65, 247-69.

Hawthorne, J. (2005) "The case for closure" in M. Steup and E. Sosa, eds., *Contemporary Debates in Epistemology*. Blackwell, Malden, Mass., 26-43.

Kitcher, P. (1980) "A priori knowledge" *Philosophical Review* 89, 3-23.

Klein, P. (1995) "Skepticism and closure: why the evil genius argument fails" *Philosophical Topics* 23, 213-36.

Kripke, S. (1980) *Naming and Necessity*. Harvard Univ. Press, Cambridge, Mass.

Lewis, D. (1979) "Attitudes de dicto and de se" *Philosophical Review* 88, 513-43. Reprinted in D. Lewis (1983) *Philosophical Papers*, vol. 1. Oxford Univ. Press, New York, 133-56.

Ludlow, P., and Martin, N., eds. (1998) *Externalism and Self-Knowledge*. CSLI Publications, Stanford.

McDowell, J. (1977) "On the sense and reference of a proper name" *Mind* 86, 159-85. Reprinted in J. McDowell (1998a) *Meaning, Knowledge and Reality*. Harvard Univ. Press, Cambridge, Mass., 171-98.

McDowell, J. (1984) “De re senses” in C. Wright, ed., *Frege: Tradition and Influence*. Blackwell, Oxford, 98-109. Reprinted in J. McDowell (1998a) *Meaning, Knowledge and Reality*. Harvard Univ. Press, Cambridge, Mass., 214-27.

McDowell, J. (1986) “Singular thought and the extent of inner space” in P. Pettit and J. McDowell, eds., *Subject, Thought and Context*. Oxford Univ. Press, Oxford, 137-68. Reprinted in J. McDowell (1998a) *Meaning, Knowledge and Reality*. Harvard Univ. Press, Cambridge, Mass., 228-59.

McGinn, C. (1989) *Mental Content*. Blackwell, Oxford.

McKinsey, M. (1991) “Anti-individualism and privileged access” *Analysis* 51, 9-16. Reprinted in Ludlow and Martin, eds. (1998), 175-84.

McLaughlin, B. and Tye, M. (1998) “Externalism, Twin Earth, and self-knowledge” in C. Wright, M. Smith and C. Macdonald, eds., *Knowing Our Own Minds: Essays in Self-Knowledge*. Oxford Univ. Press, Oxford, 285-320.

Nuccetelli, S., ed. (2003) *New Essays on Semantic Externalism and Self-Knowledge*. MIT Press, Cambridge, Mass.

Pryor, J. (2004) “What's wrong with Moore's argument?” *Philosophical Issues* 14, 349-78.

Pryor, J. (2006a) “Hyper-reliability and a priority” *Proceedings of the Aristotelian Society* 106. <http://www.jimpryor.net/research/papers/Hyper-Reliability.pdf>

Pryor, J. (2006b) “More on Hyper-reliability and a priority” <http://www.jimpryor.net/research/papers/More-Hyper.pdf>

Putnam, H. (1981) *Reason, Truth and History*. Cambridge Univ. Press, Cambridge.

Raffman, D. (1998) “First-person authority and the internal reality of beliefs” in C. Wright, M. Smith and C. Macdonald, eds., *Knowing Our Own Minds: Essays in Self-Knowledge*. Oxford Univ. Press, Oxford, 363-9.

Reimer, M. (2001a) “A ‘Meinongian’ solution to a Millian problem” *American Philosophical Quarterly* 38, 233-48.

Reimer, M. (2001b) "The problem of empty names" *Australasian Journal of Philosophy* 79, 491-506.

Salmon, N. (1998) "Nonexistence" *Nous* 32, 277-319.

Sawyer, S. (1998) "Privileged access to the world" *Australasian Journal of Philosophy* 76, 523-33.

Schiffer, S. (2005) "Paradox and the A Priori" in T. Szabó Gendler and J. Hawthorne, eds., *Oxford Studies in Epistemology*, vol. 1.

Segal, G. (1989) "The return of the individual" *Mind* 98, 39-57.

Shoemaker, S. (1990) "Qualities and qualia: what's in the mind?" *Philosophy and Phenomenological Research* Suppl. Vol. 50, 109-31. Reprinted in S. Shoemaker (1996) *The First-Person Perspective and Other Essays*. Cambridge Univ. Press, Cambridge, 97-120.

Sorensen, R. (1988) "Dogmatism, junk knowledge, and conditionals" *Philosophical Quarterly* 38, 433-54.

Stine, G. C. (1976) "Skepticism, relevant alternatives, and deductive closure" *Philosophical Studies* 29, 249-61.

Stoneham, T. (1999) "Boghossian on empty natural kind concepts" *Proceedings of the Aristotelian Society* 99, 119-122.

Tye, M. (1995) *Ten Problems of Consciousness*. MIT Press, Cambridge, Mass.

Tye, M. (2000) *Consciousness, Color, and Content*. MIT Press, Cambridge, Mass.

Tymoczko, T. (1989) "In defense of Putnam's brains" *Philosophical Studies* 57, 281-97.

Vogel, J. (1990) "Are there counter-examples to the closure principle?" in M. D. Roth and G. Ross, eds., *Doubting: Contemporary Perspectives on Skepticism*. Kluwer, Dordrecht, 13-27.

Warfield, T. (1995) "Knowing the world and knowing our minds" *Philosophy and Phenomenological Research* 55, 525-45.

Warfield, T. (1998) “A priori knowledge of the world: knowing the world by knowing our minds” *Philosophical Studies* 92, 127-47.

Woodfield, A. (1982) “Introduction” in A. Woodfield, ed. *Thought and Object*. Oxford Univ. Press, New York.

Wright, C. (2000) “Cogency and question-begging: some reflections on McKinsey’s paradox and Putnam’s proof” *Philosophical Issues* 10, 140-63.

Wright, C. (2003) “Some reflections on the acquisition of warrant by inference” in S. Nuccetelli, ed., *New Essays on Semantic Externalism and Self-Knowledge*. MIT Press, Cambridge, Mass., 57-77.

Wright, C., Smith, M. and Macdonald, C., eds. (1998) *Knowing Our Own Minds: Essays in Self-Knowledge*. Oxford Univ. Press, Oxford.