

## Chapter 6

### **Privileged self-knowledge and externalism:**

#### **A contextualist approach**

(Originally published in *Pacific Philosophical Quarterly* 83: 235–252, 2002)

#### *Preview*

Many people argue that privileged self-knowledge is incompatible with semantic externalism. I develop a contextualist approach to self-knowledge, and examine what this approach should lead us to say about the apparent incompatibility. Though such contextualism compels us to re-think the notion of privilege associated with self-knowledge, it can contain the damage wreaked by the externalist doctrine.

## Chapter 6

### Privileged self-knowledge and externalism:

#### A contextualist approach

##### I. Introduction.

In this paper I develop a contextualist conception of our privileged, first-person authoritative self-knowledge of what we think and what we mean. I then examine what this conception should lead us to say about the incompatibility between the notions of privileged self-knowledge and semantic externalism.

Many people hold that the knowledge we have of the contents of our thoughts as we think, and of what we mean as we speak, is somehow privileged. It is as if we have a special access to what we think and mean, different from the access other people have. At the very least, as people say, this knowledge is *a priori* in the sense that the first person, in contrast to the third person, can have it without having to investigate the environment and form inferences on that basis.<sup>1</sup>

Adapting David Lewis's notion of 'elusive knowledge' (1996),<sup>2</sup> I wish to explore what I think is a plausible contextualist conception of this domain of knowing.

There is much discussion about whether the notion of privileged self-knowledge about what we think and mean is compatible with semantic externalism which says that the contents of what we think and mean depend on our environment.<sup>3</sup> One aspect of the tension concerns how we can

know things that depend on the environment without having to investigate the environment.

Another aspect concerns the worry that, if I have self-knowledge, and know enough about the philosophical doctrine of semantic externalism, then I can know contingent truths by *a priori* means. I assume that the arguments that lead to these worries are sound, but show that, on the contextualist conception of self-knowledge, it is relatively easy to contain the damage they wreak.

The contextualist approach to the special case of self-knowledge suggests itself because the above two worries are similar to some general sceptical arguments concerning ordinary kinds of knowledge, worries with which contextualism deals well. It needs to be argued, however, that a response that is plausible in the general case is likewise plausible in the special case of self-knowledge. In particular, the contextualist response must allow us to retain a satisfactory notion of the *privilege* of self-knowledge. I argue that contextualism plausibly allows this, if we view the notion of privilege, not as pertaining to a somehow higher quality of knowing, but rather as pertaining to a reasonable lowering of the standards of knowing.

## **II. A contextualist conception of privileged self-knowledge.**

In order to present the contextualist conception of self-knowledge I shall adapt Lewis's general contextualist notion of 'elusive' knowledge (Lewis 1996). I begin by outlining Lewis's general account and then I explain how it applies in the particular case of privileged knowledge of thought and meaning.

Here is Lewis's definition of knowledge. '*S* knows that *P* iff *S*'s evidence eliminates every possibility in which not-*P* - Psst! - except for those possibilities that we are properly ignoring.'

(Lewis 1996, p. 554). I do not know that wobbegongs are bottom-dwellers if my evidence does not eliminate the possibilities in which they are not bottom-dwellers. Unfortunately, my evidence cannot eliminate the possibility that an evil demon is deceiving me and wobbegongs are not bottom-dwellers. But we can now, with Lewis, rescue my knowledge by adding the proviso that this is a possibility that I can properly ignore.

Lewis sensibly provides a set of rules for what can and cannot be properly ignored (Lewis 1996, p. 554-560). We cannot properly ignore the possibility that actually obtains (the rule of actuality), or any possibility that the subject believes to obtain (the rule of belief), or any possibility that saliently resembles those possibilities (the rule of resemblance). We can properly, but defeasibly, ignore the possibility that the processes whereby we obtain information (perception, memory, testimony) are not reliable (the rule of reliability), and we can defeasibly ignore the possibility that some standard methods of non-deductive inference are not reliable. We can also ignore possibilities which it is common knowledge that those around us normally ignore (the rules of method).

Finally, any possibility that we actually attend to cannot be properly ignored (the rule of attention), thus if someone directs our attention to a sceptical possibility, then we cannot ignore it, at least as long as we are in that particular context. We cannot save ourselves from scepticism by saying that we *could* ignore the sceptical possibility when it is actually raised. That just amounts to not taking on the challenge of scepticism in the first place. But we can save ourselves from scepticism, in practice, by noticing that it is only in the context of doing epistemology that our knowledge is destroyed. In other contexts we may properly ignore the sceptical possibilities.

Knowledge, on this view, is elusive in the sense that as soon as we begin to examine it, it may vanish. That is, as soon as we raise the possibility that not-*P*, then *S*'s knowledge that *P* may elude him or her.

Lewis draws attention to two important, related aspects of this story. Firstly, it is a consequence of this account of knowledge that a limit case of knowledge is knowing by ignoring. "I have to grant, in general, that knowledge just by presupposing and ignoring *is* knowledge; but it is an *especially* elusive sort of knowledge, and consequently it is an unclaimable sort of knowledge" (Lewis 1996, p. 562). It is especially elusive because, as soon as someone mentions a case of such knowledge we direct attention to the ignored possibility, and then it is no longer actually, and properly, ignored.

Secondly, my knowledge of the same thing may be good or poor knowledge. My knowledge that *P* is very poor, in fact unclaimable, if my evidence for *P* does not eliminate any not-*P* possibilities, and I just ignore them. My knowledge that *P* may be claimable, but still pretty poor if my evidence only eliminates a few not-*P* possibilities and I ignore a lot. My knowledge is good, or stable, if my evidence eliminates a lot of not-*P* possibilities and I ignore fewer. That makes it less likely that we end up in a context where our attention is drawn to a hitherto ignored not-*P* possibility.

So much by way of introducing Lewis's notion of elusive knowledge. I now explain why I think a contextualist conception of privileged self-knowledge is appropriate. The primary reason is that it gives us a straightforward, non-intellectualist and non-introspectionist account of wherein the privilege of self-knowledge consists. The privilege consists precisely in it not being a

requirement for self-knowledge about meaning and thought that *S* investigates the environment or form inferences. For we can know what we think and mean largely by presupposing and ignoring. There is no special privileged epistemic route to what we think and mean, the privilege consists in the *absence* of certain requirements. In other words, we are generally, and in most ordinary contexts, prepared to attribute knowledge of thought and meaning to subjects even though their evidence does not exclude any not-*P* possibilities.<sup>4</sup>

At this point, then, we have to say that there is nothing epistemically superior about privileged self-knowledge. On the contrary, it is, in Lewis's terminology, pretty poor knowledge. All it takes to destroy it is someone somehow directing attention to the possibility that *S* doesn't have it.

If self-knowledge is privileged in the sense that *S* can have it without excluding any not-*P* possibilities, then we should expect it to be an especially elusive sort of knowledge. Yet this doesn't seem to be the case, and moreover, we are generally very sure that we know what we think and mean. Hence, if I want to propose a contextualist conception of privileged self-knowledge, then I need to explain how it can be that such supposedly especially elusive knowledge appears to us to be very stable. The answer I favour is that we only rarely are in a context where we need to claim, or doubt, our own or other's self-knowledge, but that when we are in such a context we can target someone's self-knowledge merely by directing attention to a not-*P* possibility. If this is right, then, though self-knowledge may elude us very easily, it doesn't elude us very often.

I think there is a very good explanation for why we rarely find ourselves in contexts where our self-knowledge is destroyed. Thinking and speaking are, to a large extent, tools for representing the world and for transmitting information about the world. What matters to us is whether the tools function properly, that is, whether what we think and what we mean enable us to represent and transmit information. Since we are interested in representing correctly and in transmitting useful, veridical information, we may often find ourselves in a context where doubts arise about a particular piece of evidence, and where consequently our knowledge of the relevant facts of the matter eludes us in the way Lewis describes. But rarely is the tool itself subject to such scrutiny. If you say or think that the guards do their rounds at 4 o'clock, and somehow doubts arise, then we first try to get the relevant facts of the matter right, and mostly it is only if the doubts stubbornly persist that we end in a context where there is doubt about your self-knowledge. No doubt we refrain from creating contexts of semantic doubt for the good reason that conscious thought and communication would be incredibly cumbersome if we had to constantly worry about eliminating possibilities in which *S* doesn't have self-knowledge.<sup>5</sup>

When doubts arise about whether we know what we think and mean, our self-knowledge is destroyed. But this kind of context rarely persists for normally such doubts are quickly resolved. Normally, the subject whose self-knowledge is under scrutiny only needs to re-affirm what he or she thinks or means. That quickly restores *S*'s self-knowledge and enables the thinker or the interlocutors to pass into a context where the relevant facts of the matter again can come under consideration. We do not care about restoring self-knowledge by spending time and energy eliminating a lot of not-*P* possibilities because it is not in itself especially disastrous to fail to know what one thinks or means. It only matters in so far as it has an adverse impact on our capacity for representing or communicating.<sup>6</sup>

There are of course various ways in which we can end in more persistent or chronic contexts of semantic doubt. If *S* says “put the wheels on the garage!” and we know that *S* is in a fairly advanced stage of Alzheimer’s, then we may refrain from attributing knowledge of what she means to her. Likewise if *S* has just received a blow to the head, or is very drunk or otherwise delirious. These are relatively rare cases where our doubts make us go directly to considering *S*’s self-knowledge without first considering questions about the relevant facts of the matter.

Contexts of semantic doubts can arise in more indirect ways than as a result of, e.g., receiving a blow to the head. To illustrate, I mention two kinds of semantic characters who, with their particular linguistic behaviour, make it more likely that someone’s self-knowledge is destroyed. Firstly, there are semantic conservatives (or language mavens).<sup>7</sup> They refuse to change their language use in step with the rest of the community and they reproach others for the way they use their language, that is, for not knowing the ‘real’ meaning of what they say. Secondly, there are semantic sycophants, they change what they mean according to their company and often say things like “that’s exactly what I meant” when in fact they didn’t. Semantic conservatives and sycophants are infuriating because they create contexts of semantic doubt so often that it is hard to have a conversation about the relevant facts of the matter with them. As we shall see, some philosophers of language and mind are able to do much the same.

Under this contextualist conception of privileged, first-person authoritative self-knowledge of what we think and mean it is not a special, epistemically superior, sort of knowledge. The reason it appears special (and invites philosophical theorising) is something about our practices of attributing this kind of knowledge and the particular contexts in which it usually comes under

threat. Its peculiarity says more about the purposes of our thinking and conversing than about our epistemic capacities. We can now also easily explain the intuition about the asymmetry of first-person authoritative self-knowledge and third-person access to what *S* thinks and means. This is the intuition that a third person, *S'*, cannot know what *S* thinks or means in the same privileged way as *S* herself can. The contextualist can simply point out that this is just because, when it comes to attributing knowledge of what others mean or think, we more often end in contexts where more evidence is required. The difference between first- and third-person access is more a difference of degree than a difference in sort of knowledge.

### **III. The compatibility of privileged self-knowledge and semantic externalism.**

I now want to discuss what the above contextualist conception of privileged self-knowledge should lead us to say about a group of much discussed problems about the compatibility of self-knowledge and semantic externalism.

These problems arise from a perceived tension between: (a) the idea that knowledge of thought and meaning is privileged, or *a priori* in the sense that one does not have to investigate one's environment or form inferences in order to have it. And (b) the philosophical thesis of semantic externalism which says that content and meaning depend on one's environment. Thus, for example, 'water' as expressed by *S*<sub>1</sub> may mean water on Earth, but mean twater on Twin Earth as expressed by *S*<sub>2</sub>, where *S*<sub>1</sub> and *S*<sub>2</sub> are indiscernible (see Putnam 1975). Semantic externalism also comes in a social form such that the content of what is thought or said depends on the communal use of the concept in question (see Burge 1979). (a) and (b) are in tension on two fronts usefully distinguished by Martin Davies (2000). On the one hand there is the *achievement* problem: how can *S* achieve knowledge about thought and meaning without having to investigate

the environment, given that meaning depends on contingent facts about *S*'s environment? On the other hand there is the *consequence* problem: if *S* knows *a priori* that (1) *S* thinks or means that water is wet, and *S* knows, via *a priori* philosophical speculation, that (2) if *S* thinks or means that water is wet, then water (or other speakers) exists, then *S* can come to know *a priori* that (3) water (or other speakers) exists. The problem is that we ought not be able to know such contingent truths via an *a priori* route.<sup>8</sup>

I shall consider these problems in turn. In each case I focus on what I think is the most potent version of the problem. I assume that the problems are serious and relevant, and that the challenge that they pose must be met by any defender of externalism and privileged self-knowledge. The force of the contextualist conception of self-knowledge is that it respects the challenge yet manages to contain the damage wreaked by the problems.

#### *The achievement problem.*

Initially, externalism poses no special problem for the achievement of self-knowledge about meaning. In this I agree with Heil (1988), Wright (1992), Davies (2000) and others.

Philosophical speculation tells us that when *S* thinks or means that water is wet, what she thinks or means depends on the environment. We must assume that *S* can think or mean this without having to investigate the environment, for the achievement problem was intended to be about how *S* can achieve knowledge of this fact *a priori*. But, now, if we represent this knowledge by saying that *S* knows that she thinks or means that water is wet, then the same piece of philosophical speculation tells us that the embedded occurrence of 'water is wet' also depends on the environment, and that here, likewise, it can have this externalist content or meaning without *S* having to investigate the environment. Hence, we are led to say that, if we have a general

account of self-knowledge in place, then semantic externalism poses no further, special problems (on the condition that we do not need any empirical knowledge in order to think or mean that water is wet in the first case).

*Diachronic self-knowledge and Twin Earth shuffles.*

There might however be a special case of achieving self-knowledge which is challenged by semantic externalism. For consider what happens if, where time  $t_2$  is later than  $t_1$ ,  $S$  at  $t_2$  says “at  $t_1$  I thought that water is wet”.<sup>9</sup> The problem with this case is that it is a possibility that  $S$  between  $t_1$  and  $t_2$  has moved from a twater context to a water context and thus at  $t_1$   $S$  thought that twater is wet. By hypothesis, such Twin Earth shuffles go undetected by  $S$ . Must we not then say that, since  $S$ 's evidence does not eliminate the possibility that she has been subject to a Twin Earth shuffle, she cannot know what she thought? We have been assuming that she does have self-knowledge, and Twin Earth shuffles are clearly possible. So we cannot deny the possibility of either. If self-knowledge is, as I have argued, elusive, then we have a way out: the possibility that  $S$  has been subject to a Twin Earth shuffle is properly ignorable.<sup>10</sup> So, in so far as  $S$  hasn't actually been subject to a Twin Earth shuffle, and in so far as no-one directs attention to the possibility that she has,  $S$  does know what she thought. On the other hand, while in a context where someone, for example a philosopher of language and mind, directs attention to the possibility of Twin Earth shuffles,  $S$ 's self-knowledge is in fact destroyed.

*Self-knowledge and social externalism.*

It might be accepted that we can properly ignore the farfetched possibility of Twin Earth shuffles. But semantic externalism also comes in the form of social externalism according to which ordinary changes in a linguistic community's usage patterns can influence what a speaker

thinks or means.<sup>11</sup> Call such changes Social Semantic shuffles. In the original story they also go undetected by *S* (though of course we could also conceive of shuffles that are detectable). The possibility of such Social Semantic shuffles poses the same kind of challenge to diachronic self-knowledge as Twin Earth shuffles. Someone might object that we cannot adopt the same response to this challenge because it is much less plausible that we can properly ignore such possibilities since they seem quite normal and widespread.<sup>12</sup> If Social Semantic shuffles are normal and widespread, then they appear to be relevant for *S*'s claim to have self-knowledge, and we ought therefore to require that *S* has evidence for her self-knowledge which eliminates them. But if we have to keep track of the usage patterns of our peers (and our earlier selves) in order to know what we each think and mean, then this knowledge ceases to be privileged.

The problem is that if we ignore the possibility that there are undetected Social Semantic shuffles, then we seem to be violating Lewis' rule of resemblance which says that we cannot properly ignore a possibility which saliently resembles a possibility which cannot itself be properly ignored (by rules other than the rule of resemblance, Lewis 1996, p. 556). This rule is intended to deal with Gettier-like cases where *S* eliminates not-*P* possibilities and *P* actually obtains, but where *S* doesn't know that *P*.<sup>13</sup> In the present case, the problem would seem to be that the possibility of undetected Social Semantic shuffles resemble the actual world in respect of *S*'s evidence, which is the same in the two possibilities. Actually, however, it doesn't violate the rule because Lewis, at a similar point in his general account of knowledge, introduces an *ad hoc* exception to the rule of resemblance, viz. that it does not apply to resemblance in respect of *S*'s evidence (Lewis 1996, 556-7). Although this is not entirely satisfactory (since it amounts to not letting the sceptic have his or her full say after all) we could perhaps deal with the problem of Social Semantic shuffles in that way.

However, this will not do because the rule of resemblance itself is a threat to the notion of privileged knowledge of meaning. For if we attribute knowledge of meaning according to that rule, then we must require that *S* eliminates nearby possibilities in which what *S* says or means falls victim to a (perhaps detectable) Social Semantic shuffle. For, if the actual world is affected by many Social Semantic shuffles, then such nearby possibilities resemble actuality in respect of Social Semantic shuffles (and, if it is a detectable shuffle, not in respect of *S*'s evidence; so this case could not be covered by the *ad hoc* clause). And it seems *S* can only eliminate such possibilities by inference and investigation of her environment. Hence, even if we deal with the problem about undetectable Social Semantic shuffles via the *ad hoc* exemption to the rule of resemblance, our knowledge of what we think and mean still will not be privileged.<sup>14</sup>

This appears to leave us with a dilemma for now we see that the real culprit is the rule of resemblance. With the rule of resemblance, self-knowledge is not privileged. Without it, we cannot deal with Gettier-like cases. I think this dilemma is only apparent: in giving an account of self-knowledge we can do without the rule of resemblance because the intuitions that drive the Gettier-like counterexamples do not apply in the case of knowing thought and meaning. If this is right, then there is after all something special about self-knowledge: in contrast to other domains of knowing we are happy to attribute knowledge of thought and meaning to *S* even though her evidence does not eliminate nearby relevant possibilities.

Of course it would be *ad hoc* if we get rid of the rule of resemblance just in order to save the privilege of self-knowledge, but I think we can provide good independent reasons for leaving the rule out. Here is a Gettier-like case for self-knowledge. Assume the actual world unbeknownst to

*S* is marred by very frequent Social Semantic shuffles. In that case, by the Rule of Actuality, *S* will very often fail to have privileged knowledge of what she thinks and means. This is a consequence we have to accept. Now assume that in spite of all the Social Semantic shuffles that occur, *S* happens to get one of the few unaffected meanings right. At  $t_2$  *S*, say, thinks that at  $t_1$  she thought that wobbegongs are carpet sharks, and she is right. Without a rule of resemblance, we have to attribute knowledge of meaning to *S* because without the rule, *S* can properly ignore the actuality-resembling possibility that the thought that wobbegongs are carpet sharks is affected by Social Semantic shuffles. Clearly, our intuitions about knowledge tell us that this kind of response would not be acceptable for attributing knowledge of real barns in the land of bogus barns. But it does seem acceptable for attributing knowledge of the thought that wobbegongs are carpet sharks because *S* might be able to use that thought to represent the world and express that thought in communication in order to transmit information about wobbegongs, and that, as I suggested in Section II, is what we care about when it comes to attributing self-knowledge about thought and meaning.<sup>15</sup>

Here is a somewhat anecdotal case to illustrate this thought. *S* returns home to where her old dialect used to be spoken. But, unbeknownst to her, her old peers have changed: though their vocabulary is mostly the same, almost all words have acquired new meanings. By chance, the very first sentence *S* utters contains only words unaffected by these massive semantic changes. The sentence is, say, “the linguistics department used to be around the corner on N Street.” It seems wrong (indeed, almost arrogant in the manner of semantic conservatives) to insist that *S* doesn’t know what that sentence means, even though, had she uttered almost any other sentence, she would not have known: what she cares about is getting to the linguistics department where she no doubt will find out about all those other sentences whose meanings she doesn’t know.

There is a more principled argument for why we must reject the rule of resemblance. If we have the rule of resemblance, then *S* must engage in observation and inference in order to eliminate not-*P* possibilities. Engaging in such observation and inference involves having thoughts and perhaps saying things (e.g., asking questions). But if actuality is marred by Social Semantic shuffles, then the thoughts that *S* has, and the meaning of what she says, as she is engaging in these very observations and inferences, will themselves be under threat. In her very attempt to eliminate not-*P* possibilities in accordance with the rule of resemblance she would thus be led to go back and reapply the rule to a new set of thoughts and meanings; and in order to eliminate not-*P* possibilities for that set she would have to reapply it to a further set once again, and so on. As a consequence, she could never satisfy the rule.<sup>16</sup> So she could never get to know what she thinks and means, if she is not allowed to violate the rule of resemblance. If there is to be knowledge of thought and meaning at all—whether privileged or not—it had better not be constrained by the rule of resemblance.<sup>17</sup>

The problem with the rule of resemblance was that attributing knowledge according to it entails that it cannot be privileged knowledge. If we motivate the rejection of the rule of resemblance along the above lines, then there is no *ad hocery* since these points are not aimed at establishing that this knowledge can be privileged, they turn on whether we can have such knowledge at all and on the role this knowledge plays for us.

We should not forget that all this is consistent with the point that if someone actually directs attention to the possibility that there are Social Semantic shuffles, then *S*'s self-knowledge is destroyed because this is a possibility which her evidence does not eliminate. But as long as no-

one does that, and as long as *S* gets it right, then the occurrence of frequent shuffles does not impede her self-knowledge. The absence of the rule of resemblance makes it easier for us to attribute knowledge of thought and meaning to *S*, but this knowledge may still very easily elude her.

This gives us a conception of what it is to be a knowledgeable thinker and speaker. It is not, as in the case of being knowledgeable about barns, to be reliable in the sense that one eliminates nearby relevant possibilities. Rather, one is a knowledgeable thinker and speaker in so far as one does not end in contexts where doubts are raised about one's knowledge of what one thinks and means. I think we can add to this that, in order to be a knowledgeable thinker and speaker, one must be able to react properly when one in fact is in a context of doubt, that is, one must be able to quickly get representation and communication on line again.

If this is the right conception of our privileged knowledge of thought and meaning, then we must say that there is something special about this privilege after all. It is not privileged in the sense that it is epistemically superior, rather it is 'privileged' in the sense that the standards of knowing thought and meaning are lower than the standards for other domains of knowing.

*The consequence problem.*

Here is the consequence problem. Consider this valid argument:

- (1) *S* thinks or means that water is wet
- (2) If *S* thinks or means that water is wet, then water exists

Therefore:

(3) Water exists

Premise (1) is knowable by privileged self-knowledge which is *a priori* in the somewhat imprecise, but customary negative sense of not depending on empirical investigation;<sup>18</sup> (2) is knowable by *a priori* philosophical reasoning. But (3) ought not to be knowable on the basis of two pieces of *a priori* knowledge. The problem is that normally we would require someone to have engaged in empirical investigation, and eliminated some possibilities in which water does not exist, in order to claim knowledge that water exists. This is the consequence problem: if we are committed to privileged self-knowledge and to semantic externalism, then, as a consequence, we are committed to the possibility of knowing such contingent truths by *a priori* means.

There are various ways to take this problem. One can argue it shows that privileged self-knowledge and semantic externalism are incompatible (Boghossian 1997). One can argue that we ought to bite the bullet: we can in fact know *a priori* that water exists (Sawyer 1998, see also Miller 1995). Or one can argue for a principle which limits the transmission of warrant from the premises to the conclusion (Davies 2000, see also Wright 1985). The response we are led to by the contextualist conception of elusive self-knowledge differs from these three responses, though it does share some of their features. This response will allow us to meet the challenge posed by the consequence problem while being able to efficiently contain the damage it wreaks.

I do not think we really need to construe a limitation principle for the transmission of warrant. For often the warrant *S* has for (1) will as a matter of fact be destroyed before she gets to draw the conclusion (3). This is because we may assume that (2) often actually directs our attention to a not-*P* possibility. (2) is after all a philosophical exercise involving Twin Earth cases. In that

case a context is created in which *S*'s knowledge concerning (1) eludes her. And then the move to (3) is blocked (this kind of move is familiar from the general discussion of contextualism and scepticism, I take my cue from Lewis 1996, p. 564, see also DeRose 1995; part of the point being made in this paper is that we can use this move in the debate about externalism and self knowledge, only if a contextualist construal of privileged self-knowledge as such is plausible at all).

Perhaps, however, the philosophical speculation which sustains semantic externalism does not have to rely on Twin Earth cases or other scenarios that might direct our attention to not-*P* possibilities and thus destroy the warrant for (1). (Perhaps some arguments for the object-dependence of singular thoughts could be construed in this way, see for example Evans 1982, Ch. 9.4). In that case perhaps *S* can retain the warrant for (1) for long enough to arrive at the conclusion (3). This would appear to be bad enough for then it seems that *S* can indeed know contingent truths on the basis of *a priori* philosophical speculation and privileged self-knowledge.

I think this is unavoidable, and that we should bite the bullet. But we should not overdramatise this. The conclusion that water exists must have its warrant transmitted from the two premises. So its warrant cannot be any stronger than the warrants for the premises. But we have already seen that the warrant for (1) is very weak indeed. In order to know (1) *S* need not have eliminated any possibilities by inference or empirical investigation. Hence, *S* can know (3) without having eliminated, by inference or empirical investigation, any possibilities in which water does not exist. This knowledge is highly unstable: in most cases where one would genuinely claim it, one would also immediately create a context in which one attends to the

possibility that water does not exist, i.e. a context in which this knowledge eludes *S* immediately. Since this knowledge claim is so very easy to defeat, we should not be worried that speakers can possess it on the basis of self-knowledge and speculation about externalism. We ought only to be worried should the warrant for (3) turn out to be so strong that *S* can get further epistemological mileage out of having this knowledge.

Notice that under this conception of elusive self-knowledge the conclusion that water exists cannot provide *S* with confirmation in the case of doubt about the existence of water (Davies, 2000, mentions this as a problem for the ‘bite the bullet’ response advocated by Sawyer 1998). For if *S* is actually in doubt about the existence of water, then she has created a context in which a warrant as weak as the warrant for self-knowledge is not going to do any epistemic work whatsoever. Moreover, if *S* believes that speculation about semantic externalism can help to settle doubt about the existence of water, then it is highly likely that a context will be created in which her warrant for (1) is destroyed by directing attention to sceptical possibilities, even before she gets to try to transmit it to (3).

So far this gives us the following answer to the consequence problem: either the warrant for (1) is destroyed by (2) before we arrive at (3), or we do indeed arrive at (3), but our knowledge of it is highly unstable and nothing to worry about.

This line of response to the consequence problem doesn’t pay any special attention to the *a priori* status of self-knowledge. As such it seems especially plausible to me because I think that the consequence problem, on the customary reading of the notion of *a priori* knowledge, has got

less to do with the privileged status of self-knowledge than with worries about the strength of one's evidence in support of a knowledge.

My point is that there could be a kind of consequence problem even if self-knowledge were not privileged. Assume self-knowledge about a limited set of concepts is not privileged, i.e. that we have to somehow monitor our previous thoughts and our use of words to know what we think and mean (maybe I walk around with a tape recorder and must record and play back what I say in order to have self-knowledge, or perhaps I have to ask my peers what I mean, or consult a diary to know what I think). In that case there could still be a consequence problem, for if externalism is true, then we would know by philosophical reasoning and by observing our own behaviour (listening to the tape, asking our peers or consulting the diary) that water exists. This would be a consequence problem for normally we would not allow such evidence to justify the claim that water exists. This problem is not keyed to any qualms about knowing contingent truths through *a priori* means. What makes this a consequence problem is simply worries about the strength of the justificatory case for the conclusion that water exists. I say it is the same in the case where the knowledge claim is based on privileged self-knowledge: it is just a worry about how strong the justificatory case is. On the contextualist conception of self-knowledge, we can see that the justificatory case is very weak indeed, and there will in that case not be anything to worry about.

The customary negative definition of *a priori* knowledge is after all rather imprecise, so perhaps we should welcome this attempt to sideline its relevance for self-knowledge. I should therefore end by briefly considering the question whether self-knowledge might be *a priori* in another sense. We have to face that question in any case, because, in the context of a contextualist

epistemology we should be wary of the definition of *a priori* knowledge as ‘knowledge not based on investigation of the environment’. Under this definition the contextualist would be forced to say that *all* truths are knowable *a priori*. For all truths can in principle be known on the basis of presupposing and ignoring alone. This is clearly not acceptable, so the contextualist had better have a different requirement on *a priori* knowledge.

Perhaps we can say that *S* can know *P a priori* only if *S* can know *P* without knowing anything about which world is in fact the actual world. A complete definition of *a priori* knowledge along these lines would have independent plausibility because it would allow us to respect Kripkean insights by treating *a priori* and necessary truths in two different, but related dimensions.<sup>19</sup> Such a definition would be congenial to the contextualist if knowing by ignoring and presupposing differs from knowing without knowing anything about which world is in fact the actual world. This is borne out if, as I think is plausible, knowing what needs to be ignored and presupposed in some cases can depend on knowing which world is in fact the actual world.

Should we construe the privilege of self-knowledge as *a priori* in this sense? Not necessarily, for we have seen that the privilege is associated with a lowering of the standards of knowing, and there is no reason to think this holds of *a priori* knowledge in general. In that case, we have no reason to claim that premise (1) of the consequence problem is knowable *a priori* (in the new sense), and then the problem would not arise.

On the other hand, we could have other reasons to think that the privilege of self-knowledge is *a priori* in precisely the sense that it depends on making assumptions about how things would be if  $w_1$  were actual, or  $w_2$ , or ..., without knowing which  $w_n$  is in fact actual. This would not be

wholly unwelcome because it would make self-knowledge not only *a priori*, but also contingent. But it should be noticed that this approach takes issue with semantic externalism for it makes knowledge of relational content secondary to knowledge of non-relational content: *S* has privileged knowledge of the relational content of 'water is wet' by virtue of knowing how its content depends on context (Jackson 1998). As such this approach would deal with the incompatibility between self-knowledge and externalism by denying the latter.

#### **IV. Concluding remarks.**

On a contextualist approach to our knowledge of what we think and mean we can make sense of the particular privilege we associate with this kind of knowledge. It is privileged in the sense that the standards for attributing it are rather low and because we, for good pragmatic reasons, do not often find ourselves in a context in which our knowledge of meaning is destroyed.

Under this conception we can contain the damage from the arguments to the effect that self-knowledge and semantic externalism are incompatible. We can achieve privileged, but elusive knowledge of relational content by properly ignoring Twin Earth and Social Semantic shuffles.<sup>20</sup> And there is no danger associated with being able to arrive at knowledge of certain contingent truths on grounds of privileged knowledge because self-knowledge is very elusive and cannot be used to settle doubts about these contingent truths.

It is an interesting question what it is about semantic facts that enables us to apply such low standards when we attribute knowledge about what we think and mean. *Why* is it so easy for the first person to have such knowledge? A conjecture about the metaphysics of semantic facts can perhaps help here. Assume, as is common, that semantic facts are somehow dependent on our

use of thoughts and expressions. Thus we could say, for example, that content depend on conceptual roles.<sup>21</sup> In that case we can say that in deploying thought and the expressions of our language we produce and reinforce content. But if we, by thinking and speaking, are ourselves producers and reinforcers of content, then it is no wonder that it is easy for us to know what we think and mean.<sup>22</sup>

Of course this cannot be the end of the story. For the story about how content depends on use comes with the proviso that not all use constitutes content. Without the proviso, mistakes would not be possible (see Kripke 1982, Ch. 2). The story about content must come with an account of how our use is to some degree also subordinate to our earlier use and to our peer's use (see Pettit 1990, 1999). This is what I have tried to accommodate in the above account: the contexts in which our knowledge of meaning is destroyed will likely be contexts in which we register some discrepancy in use which we cannot resolve by focusing on the relevant facts of the matter.<sup>23</sup>

#### References:

Boghossian, Paul. (1989). "Content and Self-Knowledge," *Philosophical Topics* 17, pp. 5-26.

———. (1997). "What the Externalist can Know A Priori," in *Knowing Our Own Minds*, C.

Macdonald, B.C. Smith, C. Wright (eds.). (Oxford: Oxford University Press), pp. 271-84.

Brown, Jessica. (1995). "The incompatibility of anti-individualism and privileged access," *Analysis* 55:3, pp. 149-56.

Burge, Tyler. (1979). "Individualism and the Mental," *Midwest Studies in Philosophy* 5, *Studies in Metaphysics*, Peter A. French, Theodore E. Uehling, Howard K. Wettstein. (eds.), (Minneapolis: University of Minnesota Press).

- . (1988). “Individualism And Self-Knowledge,” *Journal of Philosophy* 85., pp. 649-665.
- . (1996). “Our Entitlement to Self-Knowledge,” *Proceedings of the Aristotelian Society* XCVI, pp. 91-116.
- Cohen, Stewart. (1987). “Knowledge, Context, and Social Standards,” *Synthese* 73, pp. 3-26.
- DeRose, Keith. (1995). “Solving the Sceptical Problem,” *The Philosophical Review* 104, pp. 1-50
- Davidson, Donald. (1987). “Knowing One’s Own Mind,” *Proceedings and Addresses of the American Philosophical Association* 60, pp. 441-458.
- Davies, Martin and Humberstone, I. L. (1980). “Two Notions of Necessity,” *Philosophical Studies* 38, pp. 1-30.
- Davies, Martin. (2000) “Externalism and Armchair Knowledge,” in P. Boghossian and C. Peacocke (eds), *New Essays on the A Priori*. (Oxford: Oxford University Press), pp. 384-414.
- Evans, Gareth. (1982). *The Varieties of Reference*. (Oxford: Oxford University Press).
- Heil, John (1988). “Privileged Access,” *Mind* 97, pp. 238-51.
- Horwich, Paul. (1998). *Meaning*. (Oxford: Oxford University Press).
- Hohwy, Jakob. (2001). “Semantic Primitivism and Normativity,” *Ratio* 14, pp. 1-17.
- Jackson, Frank. (1998). *From Metaphysics to Ethics*. (Oxford: Oxford University Press).
- Kripke, Saul. (1982). *Wittgenstein on Rules and Private Language*. (Oxford: Blackwell).
- Lewis, David. (1996) “Elusive Knowledge,” *Australasian Journal of Philosophy* 74, pp. 549-67.
- Ludlow, Peter. (1995a) “Externalism, self-knowledge and the prevalence of slow switching,” *Analysis* 55:3, pp. 45-9.
- . (1995b) “Social Externalism, self-knowledge, and memory,” *Analysis* 55.3, pp.157-159.
- . (1997). “On the relevance of slow switching,” *Analysis* 57:4, pp. 285-6.

- Ludlow, Peter and Martin, Norah (eds.). (1998). *Externalism and Self-Knowledge*. (Stanford: CSLI Publications).
- McKinsey, Michael. (1991) "Anti-individualism and Privileged Access," *Analysis* 51, pp. 9-16.
- Miller, Richard, W. (1997) "Externalist self-knowledge and the scope of the a priori," *Analysis* 57, 67-75.
- Pettit, Philip. (1990). "The Reality of Rule-Following," *Mind* 99, pp. 1-21.
- (1999): "A Theory of Normal and Ideal Conditions," *Philosophical Studies* 65.
- Pinker, Stephen. (1994) *The Language Instinct*. (Penguin).
- Putnam, Hilary. (1975) "The Meaning of 'Meaning'," in *Language, Mind, and Knowledge, Minnesota Studies in the Philosophy of Science* 7, (Minneapolis: University of Minnesota Press).
- Sawyer, Sarah (1998) "Privileged Access to the World," *Australasian Journal of Philosophy*, 76, pp. 523-33.
- Stalnaker, Robert (1978) "Assertion," in P. Cole (ed.), *Syntax and Semantics* 9, (New York: Academic Press), pp. 315-32.
- Warfield, Ted. A. (1997). "Externalism, privileged self-knowledge and the irrelevance of slow switching," *Analysis* 57:3, pp. 282-4.
- Wright, Crispin (1985) "Facts and Certainty," *Proceedings of the British Academy* 71, pp. 429-72.
- . (1989) "Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics," in *Reflections on Chomsky*, Alexander George (ed.). (Oxford: Blackwell).
- . (1992) "On Putnam's Proof that we are not Brains in Vats," *Proceedings of the Aristotelian Society*, pp. 67-99.

---

<sup>1</sup>See, e.g., the papers in Ludlow and Martin 1998 (e.g., McKinsey 1991, Brown 1995).

<sup>2</sup>For other organisations of contextualism, see Cohen 1987, DeRose 1995.

<sup>3</sup>See references below. Good places are Ludlow and Martin 1998, Boghossian 1997, Davies 2000.

<sup>4</sup>Wright 1989 draws on the same intuition about our self-knowledge, though not in a context of defending contextualism about self-knowledge.

<sup>5</sup> Probably it would be more than just cumbersome: if meaning and content somehow depend on use, then it is hard to see how it could be a requirement that *S* eliminates the possibility that she doesn't use her words and thoughts correctly. How can meaning and content ever get off the ground if we must constantly ask questions about the use that constitutes it? Indeed, how could we then begin asking questions at all? (see Davidson 1987). I return to this issue below.

<sup>6</sup>Under this conception, knowledge of thought and meaning is not so poor knowledge that it is unclaimable or unascrivable. The reason is related to the considerations above: when we engage in any form of critical reasoning about something we need to ascribe knowledge of thought and meaning to the person doing the reasoning (here I follow Burge 1996). If this is right then, whenever we try to resolve doubt about something (e.g., about whether the guards are doing their rounds at 4 o'clock) we need to claim or ascribe knowledge of thought and meaning. This doesn't mean however that knowledge of thought and meaning eludes the subject because the contexts in which it is claimed or ascribed are normally contexts where attention is directed to not-*P* possibilities relevant to the facts of the matter (the guards doing their rounds), not relevant to knowledge of thought and meaning. The point is that when we claim or ascribe knowledge of thought and meaning it is mostly not in contexts where we need to defend this knowledge, rather it is in contexts where we need knowledge of thought and meaning in order to defend other knowledge claims.

---

<sup>7</sup>See Pinker 1994, p. 373.

<sup>8</sup>Various forms of this argument can be found in the literature, see e.g. McKinsey 1991, Boghossian 1997, Davies 2000, Peacocke 1999, Ch. 5.2.

<sup>9</sup>Such diachronic cases are also discussed by Boghossian 1989, and Wright 1992. See also Burge 1988, 1996.

<sup>10</sup>Boghossian (1989, p. 12-13), speaking in terms of relevance, makes the similar point that as long as Twin Earth is not a relevant possibility it can be ignored.

<sup>11</sup> See Ludlow 1995a.

<sup>12</sup>Boghossian (1989, p. 13-14) makes this kind of objection, arguing that slow switches between Earth and Twin Earth may be relevant alternatives. See also Ludlow 1995a, 1995b. This issue is often, but in my contextualist view misleadingly, discussed in terms of the nature of memory. Thus Ludlow organises the general principle behind Boghossian's argument like this: (1) If *S* forgets nothing, then what *S* knows at  $t_1$ , *S* knows at  $t_2$ . (2) *S* forgot nothing. (3) *S* does not know that *P* at  $t_2$ . Therefore (4) *S* did not know that *P* at  $t_1$  (Ludlow 1995b). For a contextualist (and others), the principle in (1) is false: even though *S* forgets nothing the standards of knowing may have changed such that, at  $t_2$ , *S* simply needs more evidence in order to know that *P*. On this reading the nature of memory is irrelevant — what matters is the strength of one's justificatory case relative to a given context. This also creates problems, but, as we shall see in the next paragraph, of a different sort.

<sup>13</sup>Lewis mentions Goldman's example (originally due to Carl Ginet) of the land of bogus barns: 'Unbeknownst to me, I am travelling in the land of bogus barns; but my eye falls on one of the few real ones. I don't know I am seeing a barn, because I may not properly ignore the possibility that I am seeing yet another of the abundant bogus barns. This possibility saliently resembles

---

actuality in respect of the abundance of bogus barns, and the scarcity of real ones, hereabouts.’  
(Lewis 1996, p. 557).

<sup>14</sup> Could we not just expand the *ad hoc* clause so it deals with this kind of case too? I think that would be unwise: if we expand the clause so that the rule of resemblance doesn’t apply to even relevant possibilities like Social Semantic shuffles, then it is hard to see how the rule can be invoked to deal with Gettier-cases in the first case. Presently, however, I shall argue that, for the domain of self-knowledge, we should do without the rule of resemblance altogether.

<sup>15</sup>Of course, we do not attribute knowledge of meaning to *S* if we have reason to believe that *S* is not a competent speaker in the first place. Perhaps *S* is trying to break into a new language, in which case we could perhaps be justified in doubting *S*’s self-knowledge about the thought that wobbegongs are carpet sharks, even though *S* happens to get it right. But in the case under consideration the assumption must be that *S* is a competent speaker. Actually, it seems there must be an upper limit to how widespread the Social Shuffles are. If they become too frequent, then meaning and content cease to be stable enough for error and misrepresentation to make sense.

<sup>16</sup> Notice that this is not saying that if we cannot ignore a possibility *w* because it resembles actuality, then we cannot ignore any possibility which resembles *w*. This would violate the proviso on the rule of resemblance that it only applies to possibilities that resembles possibilities that cannot be ignored by rules *other* than the rule of resemblance (Lewis 1996, p. 556). Rather, it says that as soon as one has new thoughts and says new things, then one has to go back and consider possibilities that resemble actuality in respect of Social Semantic shuffles.

<sup>17</sup>Some of these intuitions are also exploited, more systematically, in Davidson 1987. Of course, there are contexts in which it does not seem to be true that we leave out the rule of resemblance, e.g., when we deal with legal or scientific documents, rife with stipulative definitions. In those

---

cases it seems that we do rule out nearby possibilities. But, as is familiar, stipulative definitions (and ostensive definitions as well) only work in a context where no-one raises further questions about the content of the definitions themselves.

<sup>18</sup> Since this sense of the *a priori* has become standard in the debate about self-knowledge, I shall stick to it in most of my discussion. However, at the end of this section I shall briefly discuss the consequences of adopting another understanding of the *a priori*.

<sup>19</sup> Cf. Stalnaker 1978, Davies and Humberstone 1980, Jackson 1998.

<sup>20</sup> There is some discussion about what it would mean for self-knowledge and externalism to be compatible (Ludlow 1997, Warfield 1997). My argument is that they are compatible inasmuch as there are possible worlds in which there is self-knowledge and externalism is true (thus I think I side with Warfield). A stronger condition on compatibility would require that there could be no possible worlds in which the truth of externalism precludes privileged self-knowledge, or *vice versa*. A world with constant Social Semantic shuffles would preclude self-knowledge about all thoughts and expressions in that world. But this kind of world is the kind of sceptical world which contextualism is designed to deal with, so, from the contextualist's point of view at least, the issue is not keyed to the stronger condition on compatibility (and anyway, I don't think such a world would allow for enough stability in content to allow for mistakes and misrepresentation, so it would be a world without meaning and thought at all, and *a fortiori* without knowledge of meaning and thought — but not through any fault of externalism).

<sup>21</sup> For a recent proposal of meaning as use, see Horwich 1998.

<sup>22</sup> In contrast, if we conceive of such facts as use-independent primitive or Platonic entities, it would be very hard to explain how we could so easily know semantic facts. For more on such primitive semantics, see Hohwy 2001.

---

<sup>23</sup>An earlier version of this paper was read at the Australasian Association for Philosophy Annual Conference, Auckland 1997. I wish to thank the audience on that occasion as well as Richard Holton, Frank Jackson, Adrienne Lehrer, Daniel Nolan, Tim Oakley, Philip Pettit, Michael Smith, and Daniel Stoljar for helpful comments and discussions.