Ryerson University
Digital Commons @ Ryerson

Philosophy Publications and Research | Philosophy

1-1-2002

Externalism, Memory, and Self-Knowledge

K J. Kraay
Ryerson University, kraay@ryerson.ca

Follow this and additional works at: http://digitalcommons.ryerson.ca/philosophy

Part of the Philosophy Commons

Recommended Citation
Kraay, K J., "Externalism, Memory, and Self-Knowledge" (2002). *Philosophy Publications and Research.* Paper 10.
http://digitalcommons.ryerson.ca/philosophy/10

This Article is brought to you for free and open access by the Philosophy at Digital Commons @ Ryerson. It has been accepted for inclusion in Philosophy Publications and Research by an authorized administrator of Digital Commons @ Ryerson. For more information, please contact bcameron@ryerson.ca.

K. J. KRAAY

EXTERNALISM, MEMORY, AND SELF-KNOWLEDGE

ABSTRACT. *Externalism* holds that the individuation of mental content depends on factors external to the subject. This doctrine appears to undermine both the claim that there is *a priori* self-knowledge, and the view that individuals have privileged access to their thoughts. Tyler Burge's influential *inclusion theory of self-knowledge* purports to reconcile externalism with authoritative self-knowledge. I first consider Paul Boghossian's claim that the inclusion theory is internally inconsistent. I reject one line of response to this charge, but I endorse another. I next suggest, however, that the inclusion theory has little explanatory value.

Externalism holds that mental content is individuated by factors external to the subject: "individuating many of a person or animal's mental kinds – certainly including thoughts about physical objects and properties – is necessarily dependent on relations that the person bears to the physical, or in some cases social, environment" (Burge 1988, 650). This doctrine appears to undermine two extremely plausible claims. First, externalism seems inconsistent with the view that there is *a priori* self-knowledge. (If mental content is individuated partly by factors external to *S*, then it seems that *S* must investigate the external world in order to know her own thoughts.) Second, if externalism rules out self-knowledge, it is also inconsistent with the thesis that individuals have privileged access to their thoughts. (Without self-knowledge, privileged access is impossible.)[1] Tyler Burge's *inclusion theory of self-knowledge* purports to reconcile externalism with authoritative self-knowledge (Burge 1998).[2] He identifies a class of "*cogito*-like beliefs"; beliefs of the form "I think (herewith) that writing requires concentration" or "I am thinking that water is a liquid". Burge considers these to be paradigmatic instances of self-knowledge, and he reasons that in such cases,

> … one simultaneously thinks through a first-order thought (that water is a liquid) and thinks about it as one's own. The content of the first-order (contained) thought is fixed by non-individualistic background conditions. And by its reflexive, self-referential character, the content of the second-order judgment is logically locked (self-referentially) onto the first-order content which it both contains and takes as its subject matter. Since counterfeit contents logically cannot undermine such self-knowledge, there should be no temptation

Erkenntnis **56:** 297–317, 2002.
© 2002 *Kluwer Academic Publishers. Printed in the Netherlands.*

to think that, in order to have such knowledge, one needs to master its enabling conditions (659–660).

Burge's inclusion theory has been called "...not only the most promising but also the most widely accepted externalist theory of privileged self-knowledge" (Bernecker 1996, 262), but it has nevertheless had its critics. In Part 1, I consider Boghossian's charge that the inclusion theory is internally inconsistent. I suggest that replies by Peter Ludlow, Sven Bernecker, and Anthony Brueckner fail, but I urge that a different analysis of slow switching can lead to a successful defense of the inclusion theory against Boghossian's charge. In Part 2, however, I suggest that the inclusion theory has minimal explanatory value.

1. BOGHOSSIAN'S MEMORY ARGUMENT

One notable criticism of Burge is Boghossian's "memory argument" (Boghossian 1989, 22–23). Boghossian observes that on the inclusion theory, it is possible for *S* to know that he is having a water-thought at time t_1, but if he subsequently undergoes a slow switch to Twin Earth in which his water concept is replaced by a *twater* concept, then at t_2, *S* (who has by stipulation forgotten nothing) will be unable to say what the content of his thought was at t_1.[3] Boghossian concludes that "[t]he only explanation ...for why *S* will not know tomorrow what he is said to know today, is not that he has forgotten, but that he never knew. Burge's self-verifying judgments do not constitute genuine knowledge" (op. cit., 23). Ludlow (1995c, 157) formulates this argument as follows:

(1) If *S* forgets nothing, what *S* knows at t_1 *S* knows at t_2.
(2) *S* forgets nothing.
(3) *S* does not know that *P* at t_2.[4]
(4) $\therefore$ *S* did not know that *P* at t_1.

Boghossian's argument has recently spawned a torrent of criticism.[5] Ludlow, Bernecker, and Brueckner try to defend the inclusion theory against Boghossian's criticism while remaining faithful to his assumption that slow switching entails the *loss* of concepts.[6] John Gibbons and Burge criticize Boghossian by rejecting this very assumption: they suggest that slow switching entails the *addition* of new concepts, not the *loss* of old ones. In Sections 1.1–1.3, I reject the former strategy and endorse the latter.

1.1. *Ludlow and Bernecker's The Conceptual Replacement View*

Ludlow chides Boghossian for relying on an understanding of memory incompatible with externalism: "[t]he consistent social externalist is bound to say that the content of a memory is fixed at the time recollection takes place – for it is the embedding circumstances of that memory which are crucial to the fixing of its content" (1995c, 158).[7] Armed with this application of externalism, Ludlow reasons as follows:

(5) On externalism, mental content in slow-switching cases is individuated by current social or physical environment.

(6) Memories have mental content.[8]

(7) ∴ In slow-switching cases, memorial content is individuated by the *current* social or physical environment.

(8) *S* does not forget anything as a result of the slow switch.

(9) It is not the case that what *S* knows at t_1, *S* knows at t_2.

(10) ∴ Boghossian's (1) is false.

Brueckner notes that "Ludlow's position is rather curious. If I indeed do remember at t_2 what I was thinking at t_1, then presumably I thereby know at t_2 (via memory) what I was thinking at t_1. But then (1) would not be refuted and (3) would be false, contrary to Ludlow's intended strategy" (1997, 6). As will be seen, Ludlow does claim that post-switch memories (of the relevant sort) amount to knowledge, because he claims that they are reliable sources of knowledge about pre-switch thoughts. Hence, Brueckner is right that Ludlow's argument fails to undermine (1), and that it should be construed as a criticism of (3).[9] But Ludlow's argument can only be deemed a successful attack on (3) if his account of memory is acceptable, and so this must be evaluated. Hofmann (1995) objects that on Ludlow's account of memory, beliefs about times prior to the switch will be rendered false: "... if Peter recollects that he had had, at t_1, some arugula experience, then also this memory has turned by now into a falsehood, since it has become a memory-thought about tarugula".[10] Ludlow responds by suggesting that

> [i]f memories are individuated by their contents, then there are in fact different memories at t_1 and t_2 ... A single memory does not flip-flop back and forth from being true to being false. Rather, a single memory will be replaced by another memory having another content. Such memories, although possibly transient, would not be unreliable as a [sic] sources of knowledge. To the contrary, there is no reason at all why they cannot be completely reliable in the environmental conditions in which they occur (1995b, 73–74, emphasis added).[11]

To investigate what Ludlow means when he claims that such beliefs – beliefs ostensibly about t_1 that employ t_2-concepts – are reliable sources of knowledge, consider his example:

> [S]uppose that at time t_0 I come to know that water is wet. At time t_1, before I shift environments, I may recall that water is wet. Later, at t_2, due to undetected environmental changes, I may have the recollection that twater is wet. Is this second episode of memory less reliable than the first? It is difficult to see why (1995b, 74).

As it stands, this example has little to do with the inclusion theory, since neither of the thoughts Ludlow describes are *second-order* beliefs of the sort Burge considers paradigmatic examples of self-knowledge. To rectify this, suppose that *S* has the following three thoughts at the times indicated:

t_0: I am thinking that water is wet.
t_1: I recall t_0: I am thinking that water is wet.
t_2: I recall t_0: I am thinking that twater is wet.

Ludlow's claim, then, is that the t_2-thought is just as reliable a source of knowledge as the t_1-thought. But is it? I take it that a reliable source of knowledge is one that is (at least) *non-accidentally truth-indicative*.[12] If so, it is difficult to see whether the judgements made at t_2 and t_1 are reliable. Certainly, *part* of each judgement is true: water, after all, *is* wet, and so is twater. We may even presume (if we wish) that these are *reliable* parts of the relevant beliefs: let us suppose that *S* knows that she has reliable faculties that prompted her to form the judgement that water is wet and (some time later) the judgement that twater is wet. Moreover, the judgements "I am thinking that water is wet" and "I am thinking that twater is wet" are reliably true on the inclusion theory. (In fact, they are maximally reliable according to Burge, since they are infallible.) But none of this matters, for Ludlow's claim concerns the judgements made at t_1 and t_2 *in their entirety*. Are these entire judgements non-accidentally truth-indicative? If *S* has a good memory and no compelling reason to doubt it, then perhaps the t_1 judgement is reliable. It is true, after all, that at t_0 *S was* thinking that water is wet. But, even if *S* has a good memory and no reason to doubt it, it is difficult to see how the judgement made at t_2 can be thought reliable. After all, at t_0 *S* was contemplating the wetness of *water*, not the wetness of *twater*, and so the "memory" that *S* has at t_2 is simply false, and (hence) cannot be a reliable source of knowledge.

It seems, then, that Hofmann (1995) is right to allege that "...if [the circumstances of recollection] determine memory content, then memory turns into an empty, absurd faculty. That is so, since memory can no longer do what it is supposed to do, namely, to recall the very same thoughts one

earlier on had entertained".[13] But Bernecker thinks such a criticism misguided: "The problem with this objection to Ludlow's notion of memory is the assumption that a memory state *necessarily* contains the contents and concepts of the relevant earlier state ... The transfer of contents and concepts across time might be a sufficient condition for memory but it falls short of being a necessary condition" (1998, 341).

If a t_2-memory is *of* a t_0-state, then the concept entertained at t_0 must be available at t_2. Bernecker's denial that a t_0-concept must be available for a t_2-memory thus amounts to a denial of the claim that the only plausible conception of memory is one wherein a memory must be *of* something. Bernecker and Ludlow both seek to replace this understanding of memory with one in which memories are memories-*about* rather than memories-*of*. (Presumably, this means that they concede that a twater-thought cannot count as a memory of water.) Bernecker says that "... the job of memory, rather than to replay previously recorded contents, is to provide information *about* past states relative to the present environmental conditions" (1998, 341, emphasis added), while Ludlow suggests that "[a]ccording to the externalist conception of memory that I have proposed, it is not the job of memory to record contents, but rather to provide information *about* past episodes relative to current environmental conditions" (1995b, 74–75, emphasis added).

The shift from memory-*of* to memory-*about* thus purports to reconstruct a coherent account of memory without requiring that the t_0-concept be available at the time of remembering. Unfortunately, this rather desperate move is utterly implausible. Neither author gives any indication of how the latter understanding of memory differs from the former, save to assert that the latter lacks the tacitly admitted deficiencies of the former. And, in any event, they are mistaken whether "about" is taken *intensionally* or *extensionally*, the t_2-"memory" cannot be *about* t_0. Taken intensionally, the twater-thought is about the concept *twater*, not the concept *water*. And taken extensionally, the twater-thought it about the chemical compound XYZ which is clear, potable and covers most of Twin Earth, not the chemical compound H_2O that is clear, potable, and covers most of earth.[14] Thus, the Ludlow–Bernecker account entails that all memorial beliefs in slow-switching cases are false: it entails that memory is impossible on slow switching.[15]

The Ludlow–Bernecker defense suffers from a further drawback. Even if Ludlow and Bernecker were able to establish that a twater-thought can count as a memory *about* water, it would still be fundamentally unclear how this kind of memory provides information "... about the past relative to current environmental conditions" (Ludlow 1995b, 74–75). After all, it

is not as though S is in a position to learn anything about the difference between the two earths by the act of remembering, since (by hypothesis) the two earths are phenomenologically indistinguishable.

In a recent paper, Ludlow appears to tacitly concede this point. Instead of holding that memories provide information "... about the past *relative* to current environmental conditions" (1995b, 74–75, emphasis added), he suggests that memories "... ought to provide us information about past episodes which are *relevant* to current conditions (1999a, 167, emphasis added). The difference is crucial. Ludlow is now prepared to explicitly concede that a subject's t_2-belief about falling in some twater at t_1 "... is literally false ...", since the subject fell into water at t_1, not twater. But, he suggests that such false beliefs are *relevant* to Twin Earth because they are *useful*: "If at t_1 I believed it was possible to drown in water, memory will deliver a belief that it is possible to drown in twater – and a good thing too! Twater is no less wet or dangerous than water" (1999a, 167).

This new view cannot be construed as an attack on Boghossian's (3). Rather, it is to be understood as a criticism of (1), as follows. Slow switching brings about a change in the set of concepts available to S. Thus, if at t_2 S no longer has a given t_1-concept available, then she cannot be said at t_2 to *know* that concept. This suggests that it is *not* the case that if S forgets nothing, what she knows at t_1 she knows at t_2, contrary to (1). It is true that S cannot *remember* her t_1-water-thoughts, but this is not because she *forgot* them, but rather because the membership of her conceptual set changed as a result of the switching: water was replaced with twater.

In my view, this is an effective response to (1). However, it comes at a great cost. This denial of (1) sacrifices the very possibility of memory in all slow-switching cases where the t_1-concept is no longer available.[16] And, in so doing, it forfeits memorial knowledge.[17] As will be seen in 1.3, these undesirable consequences are avoidable, since it is plausible to suppose that concepts are not lost on slow switching.

1.2. *Brueckner's Denials of (2): Conceptual Replacement View*

1.2.1. *Brueckner's Denial of (2): S forgets something*

Ludlow (1995b) and Bernecker (1998) think that Twin-Earth mental content can count as memory "about" earth. Brueckner (1997) agrees that concepts are lost on slow switching, but he rejects this account of memory. In fact, rather than trying to argue that t_2-content can constitute memory, Brueckner takes precisely the opposite tack: he argues that conceptual switching entails *forgetting*. Brueckner reasons that when contents have shifted, "I misrepresent a past event (a thinking of a *chicory*-thought) in virtue of applying the wrong concept [*twicory*] to it while lacking the right

one (the one that correctly applies to the event)" (7). Brueckner thinks that Boghossian was therefore unfair to *stipulate* that there is no forgetting in slow-switching cases. If the conceptual shift due to slow switching constitutes forgetting, then Boghossian's (2) is false.

In the text, Brueckner equates *forgetting* with *memory failure*. However, in an endnote he concedes that these terms may not be synonymous. He suggests that to *forget* P may imply an inability to "...confidently and correctly answer appropriate questions regarding P" (11, n. 22). Thus, if S knew P at t_1, she can fail to remember P at t_2 "by virtue of failing to satisfy a causal condition on remembering" (*ibid.*). But, if Q tells S about P at t_3, then, according to Brueckner, S may still fail to remember P at t_3, but (being able to confidently and correctly answer questions concerning P) it is not the case that S has *forgotten* P.

This distinction between *forgetting* and *memory failure* is dubious. It seems clear that at t_3, Q either does or does not rekindle a memory in S. If Q rekindles a memory in S at t_3, then it seems reasonable to say that S forgot (or had memory failure), and then subsequently remembered (did not forget) P. However, Brueckner cannot have this in mind, for he says that S's confidence at t_3 "[derives] *entirely* from a friend's recent recounting of the events" *(ibid.*, emphasis added). So, Q does not rekindle the memory of P in S at t_3. But if Q does not rekindle the memory of P in S at t_3, then it seems at best implausible to claim that S has not forgotten P, merely in virtue of his (recently regained) ability to confidently and correctly answer questions concerning P. This ability may indeed be a necessary condition for S's not having forgotten, but it does not suffice: surely the satisfaction of the causal conditions on remembering is also necessary for S's not having forgotten at t_3.

Although his proleptic distinction is implausible, Brueckner is right to worry about his assimilation of memory failure and forgetting. In my view, the correct distinction between memory failure and forgetting is that the latter is a proper subset of the former. Every instance of forgetting is incontrovertibly an instance of memory failure, but the converse does not hold. For example, consider massive memory failure brought about by a botched lobotomy. Surely we would not say that the patient has merely "forgotten" much of what she knew before the operation; rather, we would note that her memory failed because of the disastrous surgery. More support for this distinction comes from the fact that we are far more likely to hold people culpable for the things they forget, while we tend to be more forgiving when people are unable to remember because their memory failed due to factors beyond their control. If this distinction is worthwhile, then the memory failure brought about by conceptual shift

is a paradigmatic instance of the kind of memory failure that does not constitute forgetting, and so Brueckner's argument against (2) fails.

1.2.2. *Brueckner's Denial of (1): S remembers but does not know*
Brueckner (1997, 8) considers a case of multiple slow switching as a result of which *S* loses and subsequently regains the concept "chicory". When this concept is regained, *S* utters

(M) I remember that I was thinking at t_1 that chicory is bitter.

Suppose that *S* did indeed think at t_1 that *chicory* is bitter. Since (M) is now true, Brueckner is prepared to consider it a memory, but not knowledge. (M) does not constitute knowledge because whether or not S has been informed about the switching, one of the conditions for knowledge is not satisfied. If *S* is *unaware* of her switch yet expresses a true belief about a past *chicory*-thought, (M) is a justified, true, memorial belief, but it does not count as knowledge because it is a *relevant alternative* that S has been switched. If, on the other hand, *S* is *aware* that switching took place (but does not know which world she is in) yet still expresses a true belief about a past *chicory*-thought, then Brueckner argues that (M) constitutes true memorial belief, but not knowledge, this time because the *justification* condition for knowledge is not met (1997, 8–10).

On this account, *S* can be said to have a memory because in uttering (M) *S* truly represents the t_1-belief, and there is a certain causal connection between the t_1-thought and the t_2-memory of it. But, Brueckner admits that "[t]he causal link is unusual, in virtue of the comings and goings of one of the concepts involved in the memorial representation. If this implies that the causal conditions for memory are not satisfied, then for that reason ... there is memory failure in the current case" (12, n. 24). This causal link certainly is highly unusual: Brueckner's defense of the possibility of memory covers only such cases of multiple slow switching where, fortuitously, the right concept is available. So Brueckner's account of memory implausibly makes that faculty extremely small, and utterly dependent on luck. More seriously, this account can offer no way to distinguish between memories and mere (luckily) true statements about the past. It seems reasonable to suppose that the difference between *memories* and luckily true statements about the past ought to be a *causal* difference. Since Brueckner's theory cannot account for this difference, it is fair to say that the causal conditions for memory are indeed not satisfied, and so, contrary to Brueckner, there is memory failure in this case. This is a significant consequence. If (as Ludlow suggests) slow-switching cases are prevalent, then, on Brueckner's account, memory failure is equally so.[19]

Each of these attempts to defend the inclusion theory from Boghossian's argument thus fails. Ludlow (1995b) and Bernecker (1998) cannot consistently deny (1), since they hold that a switched content can be a source of knowledge "about" t_0. Nor can their argument be successfully recast as a denial of (3), for their analysis of memory is absurd. Ludlow (1999a) successfully rejects (1), but only by forfeiting a significant sort of memory. Brueckner's denial of (2) implausibly equates forgetting with memory failure, and Brueckner's denial of (1) depends on a significantly circumscribed view of memory – a view that has been shown untenable. Moreover, each defense entails that memorial knowledge is *impossible* in slow-switching cases. Ludlow (1995b) and Bernecker (1998) claim that t_2-content can be a reliable source of knowledge "about" t_0, but this argument fails. Worse still, Brueckner is prepared to forfeit memorial knowledge *ab initio* in both his defences, since his denial of (2) claims that *S* forgets t_1-content on slow switching, and his denial of (1) holds that in certain rare cases, *S* can remember t_1-content, but not know it.

These undesirable results all flow from the assumption that slow switching results in the *loss* of concepts. But, as Boghossian (1989) noted, this position is not essential to externalism.[20] I therefore turn to a defence of the inclusion theory that relies on the view that concepts are added but not lost on slow switching. Since this defence is plausible and does not entail the absolute impossibility of memory (and hence memorial knowledge) in all slow-switching cases, it is preferable.

1.3. *Defensive Strategy B: The Conceptual Addition View*

Recently, Burge has claimed that he never meant to suggest that concepts are lost on slow switching: "Displacement was never part of the switching cases, at least in my understanding of them. Cohabitation was always the assumed case. I did not and do not consider the displacement model (as a general model for switching cases) a plausible account. I did not discuss and criticize the displacement model in the 1988 paper, but largely because I thought it clearly implausible" (1998, 364, n. 13).

Burge (1998) and Gibbons (1996) have independently offered accounts of memory according to which slow switching results only in conceptual *addition*, not conceptual *loss*.[21] Gibbons observes that "[w]hile it is easy to see how causal contact with a new type of substance can give you a new concept, it is not at all clear how it can take one away" (295).[22] Burge notes that "[m]erely being in the second environment, with concepts appropriate to that environment, does not prevent [*S*] from retaining and thinking concepts appropriate to the first. Nor does it automatically prevent his retaining knowledge that he had before" (357).

Ludlow objects that such 'preservative memory' is inconsistent with externalism: "... [e]ven if there were a mechanism which could, as it were, freeze the contents of an initial mental episode and carry it in memory indefinitely ... it would be a way of preserving the content of a thought which we could no longer have (if, as the externalist supposes, the contents of our current thoughts are determined by our new environment)" (1995b, 75).[23] This response begs the question against externalists like Burge and Gibbons, who plainly do not suppose what Ludlow supposes they suppose. Burge and Gibbons do not think that externalism entails the view that the content of *every* thought is determined by the new environment: Burge, for example, holds that "[p]reservative memory normally retains the content and attitude commitments of earlier thinkings, through causal connections to the past thinkings ... the memory content is fixed by the content of the thinking that it recalls" (1998, 357).

To be fair it must be noted that Burge's claim here is empirical, and undefended by scientific evidence. But, on the other hand, Ludlow wants to show that the externalist *should* reject preservative memory, and this cannot be done by merely insisting that all of "... our earlier contents must surely fade out of the picture" after a slow switch because "... we are embedded in a new nexus of causal and social relations" (1999a, 164). This too is an undefended empirical claim, and so perhaps this aspect of the skirmish cannot be settled without scientific adjudication. So far, then, the debate founders in a clash of intuitions between rival accounts of externalist memory.

Ludlow (1999a) levels a new charge against Burge–Gibbons 'preservative memory'. He suggests that preservative memory can convert seemingly sound arguments into unsound ones:

> Suppose, for example, that at t_2 I fall in some twater and think a thought which I express as "I am thinking that water is wet." I then reason out loud as follows: "I am thinking that water is wet, and I was thinking that water is wet at t_1, therefore I have thought that water is wet at least twice." Here is a case where the content of the first premise of my reasoning is causally connected to my falling into twater at t_2, and the content of the second premise is causally connected to my falling into water at t_1. If those are the contents at work in my inference, then my inference is (contrary to appearances) unsound. (165)

If preservative memory converts sound arguments into unsound arguments, this is a significant liability. But externalists need not be committed to the view that inferences of this sort are both apparently sound *and* really sound. Externalists can plausibly argue that sound arguments are comprised of certain sorts of *propositions*, and that the utterances under review do not represent the right sort of proposition. In particular, externalists can hold that the first premise is a proposition about *twater* and the second premise is a proposition about *water*, and so the argument *really* is un-

sound, contrary to appearances. On this view, Ludlow's example is no more problematic than the following (apparently sound but really unsound) verbal inference: "I am looking at Hesperus, but this morning I was looking at Phosphorus, so I have seen at least two stars today".[24]

Both the conceptual replacement view and the conceptual addition view are, presumably, underdetermined by empirical data. But the latter view avoids the numerous unpalatable consequences of the former, and survives Ludlow's most recent objection. Thus, it is clear that externalism need not be committed to a view of slow switching that entails the loss of concepts available before the switch. If this is so, then Boghossian's (3) can plausibly be denied, since it assumes just such a view of slow switching. If Boghossian's (3) can plausibly be denied, then Boghossian's argument does not show that the inclusion theory of self-knowledge is incoherent.[25]

2. EXPLANATORY POWER

I will now argue that while the inclusion theory has not been shown incoherent, it can be shown to lack significant explanatory power. To do so, the *scope* of the theory must be examined. Burge (1988) did not discuss this issue in detail; he only claimed that *cogito*-like thoughts constitute paradigmatic instances of self-knowledge (649, 658). Recently, Burge has suggested that the conceptual self-awareness seen in *cogito*-like judgments "... is not an unusual phenomenon among people with normal second-order abilities ... *cogito*-like judgments constitute a significant segment of our everyday mental activity" (1996, 93). However, he has more recently noted that "... self-verifying judgments are just a small sub-class of the self-knowledge to which we have special authority" (1998, 355, n. 4). It is difficult to know what to make of claims regarding the prevalence of *cogito*-like thoughts. Naturally, third-person observational data on this point appears impossible. Nor are introspection or self-reflection useful guides in this matter, for presumably, the more one introspects or reflects on one's thinking, the more *cogito*-like thoughts one is likely to think.

Nonetheless, we need not avoid the question of scope. Boghossian argues forcefully that Burge's class of (putatively infallible) *cogito*-like judgments is too broad. He notes that the paradigm does not help explain our knowledge of *standing states*. There is, after all, a logical gap between believing that writing requires concentration and *judging* that one believes that writing requires concentration. Error is thus logically possible in this domain, and so such judgments cannot be termed infallible.[26] More seriously still error is also manifestly possible for judgments about mental states *just past*.[27] *S* may think that *just now* she thought *P* though it is

false that she just now thought P. Finally, Boghossian restricts the scope of the inclusion theory to *first-person second-order present-tense assertions*. He reasons that it is possible to judge that one fears that writing requires concentration without actually so fearing. Accordingly, only *first-person second-order present-tense* assertions about "a mere thinking or entertaining of a proposition" comprise Burge's paradigm (1989, 21).[28]

Boghossian's first two limitations appear to have been uncontroversial. However, it is sometimes thought that his third limitation is too severe. Infallible knowledge of *content* in all *first-person second-order present-tense* assertions is one thing, but it seems implausible to suggest that only one kind of first-order propositional *attitude* (assertion) can be known authoritatively. Just as (according to Donald Davidson 1987, 446) skepticism concerning self-knowledge is the transposed image of Cartesian skepticism ("... our beliefs about the world are mostly true, but we may easily be wrong about what we think"), the inclusion theory now seems to invert common sense: we infallibly know the (externally determined) content of our *cogito*-like thoughts, but we have no self-knowledge of (and hence no privileged access to) all of our *own attitudes* save, in certain cases, assertion. Clearly, this appears to be a significant limitation on the explanatory power of this theory. Two responses to this charge are possible. One is to deny that the theory has this consequence, and the other is to concede that it has this consequence, but to deny that it is serious. Each will now be considered and rejected.

2.1. *Denial*

It might be thought that the inclusion theory *can* render infallible knowledge of propositional attitudes (other than assertion) in the same way as it renders infallible knowledge of content. Such an argument might proceed as follows. It is sometimes difficult to see that first-order attitude becomes *part of* second-order content because descriptions or indicators of such states do not always explicitly indicate their propositional attitude. Consider the first-order judgment "arugula is tasty". Clearly, the content of this state is expressed within the quotation marks, and no propositional attitude is explicitly indicated. However, if this content is thought or expressed in a certain way, it could indicate a particular propositional attitude on the part of the thinker or declarer. Let's assume that this attitude is *doubt*. However, if S reflects on her mental state, she may think or utter "I doubt that 'arugula is tasty' ". We now have a second-order judgment about the first-order thought. The sentence "I doubt that 'arugula is tasty' " itself *is* the content of the second-order judgment. What makes this content a second-order judgment is, in turn, *its* being entertained or uttered in the

appropriate kind of way. Then, just as with the first-order judgment, the relevant propositional attitude is not contained within the quotation marks, it is rather *added* to the words uttered or thought entertained, just in case the words are uttered or thoughts entertained in the appropriate kind of way. This means that on the inclusion theory, although first-order propositional attitude is clearly different from first-order content, the propositional attitude of the first-order judgment infallibly *becomes part of the content of the second-order judgment*.

Unfortunately, this response commits the fallacy of composition. Part of the first-order thought is its content, and the inclusion theory has established that second-order content necessarily covaries with first-order content. However, just because this is true of *part* of the thought (the content) is no reason to suppose that it is true of *all* of the thought (content and attitude). The inclusion theory holds that second-order content cannot come apart from first-order content, but since the attitudes are not properties of the content, there is no reason to suppose that they too covary infallibly.[29]

To put the point another way, this response merely establishes that first-order attitude is necessarily a component of the second-order thought. However, this fact alone offers no reason to suppose that all first-order attitudes must be known infallibly or authoritatively in the way that first-order content is.

2.2. *Concession*

The attempt to deny this limit on the explanatory power of the inclusion theory is thus unsuccessful. However, it is possible to concede that the inclusion strategy has this limitation while denying that such a limitation is problematic. This appears to be Bernecker's most recent opinion. In 1996, Bernecker considered the problem of privileged knowledge of attitudes to be a serious threat to Burgean compatibilism.[30] He appears to have since changed his mind, however, since his 1998 "Self-Knowledge and Closure" uses considerations about inclusion theory's inability to account for privileged access to attitudes to *defend* that very theory from Boghossian's charge that it cannot account for the fallibility and incompleteness of self-knowledge. Before considering Bernecker's 1998 response and why it fails, a few words are in order about this criticism and why it is misguided.

Boghossian claims that "[t]he most important consideration, however, against an insubstantial construal of self-knowledge derives [from the undeniable claim] that self-knowledge is both fallible and incomplete. ... To put this point another way, it is only if we understand self-knowledge to be a cognitive achievement that we have any prospect of explaining its admit-

ted shortcomings" (1989, 19). Boghossian is apparently objecting to the inclusion theory's rendering of *all* self-knowledge as infallible. However, just pages later, Boghossian significantly circumscribes the set of Burge's "paradigm" judgments. As noted above, Boghossian argues convincingly that the paradigm cannot account for standing mental states, mental states in the immediate past, and mental states involving attitudes other than bald assertion. Clearly then, Boghossian's criticism is misguided, for the Burgean compatibilist is free to argue that the inclusion theory does not render all self-knowledge infallible and complete, but only self-knowledge of first-person second-order present-tense assertions.[31]

Bernecker does not take this route to refute Boghossian, but rather relies on the very point that he established in 1996: that the inclusion theory cannot explain privileged access to attitudes. He reasons that

> [s]ince self-knowledge consists in the identification of the attitude as well as the content, the inclusion theory doesn't provide a complete account of privileged self-knowledge. ... The inclusion theory therefore cannot be extended to provide a solution to privileged self-knowledge of attitudinal components, and there is no indication that there is some other externalist account to be had of this kind of knowledge. Thus, there is at least one respect in which the inclusion theory can allow for privileged self-knowledge to be fallible and incomplete. (1998, 346–347)

Bernecker realizes that this might not be thought a conclusive defense of the inclusion theory, for he notes that it remains open to Boghossian to respond by suggesting that this fact ought to count *against* the inclusion theory rather than *for* it. To this imagined response, Bernecker suggests that "... it is only through the study of externalism that a reasonable notion of self-knowledge emerges. Privileged access to the attitudinal components of one's thoughts is one of the Cartesian superstitions that the inclusion theory forces us to abandon" (1998, 347).

This response is hardly stirring, and it does little credit to the inclusion theory. If self-knowledge of propositional attitudes is essential to self-knowledge, then Bernecker has given up the quest for an account of self-knowledge. And if self-knowledge of propositional attitudes is essential to privileged access, then Bernecker has forfeited privileged access. But it is overwhelmingly plausible that we have significant amounts of both self-knowledge and first-person authority. This, after all, was the chief consideration that led to the development of the inclusion theory. Bernecker's concession is analogous to endorsing skepticism upon discovering that a certain technical anti-skeptical argument fails, rather than seeking a better argument for the (overwhelmingly plausible) conclusion. Finally, Bernecker is to be faulted from a tactical perspective, for, as noted, he fails to avail himself of a simpler response to Boghossian's criticism.

The inclusion theory of self-knowledge seeks to counter the charge that authoritative self-knowledge is impossible on externalism. In showing these two positions compatible, the inclusion theory establishes the *possibility* of infallible, self-verifying knowledge of certain mental content. (This is only a possibility because it is logically possible that *S* go through life without ever having a *cogito*-like thought.) Moreover, the inclusion theory also establishes the possibility of privileged access to this knowledge of mental content, since this infallible self-knowledge is only possible from the first-person perspective.

Yet, as has been argued, the inclusion theory only demonstrates the possibility of self-knowledge of a very restricted class of mental *content* – it is silent concerning authority over all propositional attitudes save *assertion* in limited cases. Thus, it appears that the inclusion theory can defend only an extremely limited amount of first-person authority and a very limited account of self-knowledge – infallible self-knowledge of first-order *content* in a very restricted class of *first-person, second-order, present-tense assertions.* Given how plausible is the notion that we have a considerable amount of both self-knowledge and special authority over a wide variety of our thoughts – both past and present – the inclusion theory must be considered extremely limited in explanatory power.[32] This charge might be countered if more could be said concerning whether and how *cogito*-like thoughts are "paradigmatic" cases of self-knowledge. Though Burge is fond of insisting that these thoughts are paradigmatic instances, almost nothing has been written on what this means.[33] This is not to deny that such an account can be given, but only to note that it is *required* if the inclusion theory is to have significant explanatory value.

ACKNOWLEDGEMENTS

Previous versions of this paper were presented at the *1999 Western Canadian Philosophy Association Conference* and at the *8th Annual Harvard–M.I.T. Graduate Student Philosophy Conference.* I am grateful for the valuable comments and criticisms provided by David Checkland, Michael Hymers, Carl Still, and two anonymous *Erkenntnis* referees.

NOTES

[1] Privileged access entails self-knowledge, but self-knowledge does not entail privileged access. (Imagine a race of beings who know both their own minds and each other's minds equally well.)

[2] The term is due to Bernecker (1996, 265).

[3] Burge (1988) explains slow switching as follows:

> Suppose that one underwent a series of switches between actual earth and actual twin earth so that one remained in each situation long enough to acquire concepts and perceptions appropriate to that situation. Suppose occasions where one is definitely thinking one thought, and other occasions when one is definitely thinking its twin. Suppose also that the switches are carried out so that one is not aware that the switches are occurring. The continuity of one's life is not obviously disrupted ... Now suppose that, after decades of such switches, one is told about them and asked to identify when the switches took place. The idea is that one could not, by making comparisons, pick out the twin periods from the 'home' periods. (652–653)

[4] "...it is quite clear that tomorrow he will not know what he thought at t_1. No self-verifying judgment concerning his thought at t_1 will be available to him then" (Boghossian 1989, 23). Indeed, no judgment at all concerning this thought will be available, since (by hypothesis) the t_1-concept was lost on the slow switch.

[5] One exception is Goldberg (1997). Goldberg argues that the key intuition of the memory argument is correct, and tries to recast the overall argument without relying on questionable considerations about memory.

[6] Two background points must be noted here. First, Boghossian notes that "...we may imagine that after a series of such switches, *S* ends up with both earthian and twin-earthian concepts: thoughts involving both arthritis and tharthritis are available to him" (1989, 13). While he thinks that this way of understanding slow switching "is perfectly coherent – and a lot more interesting" he says that he will follow tradition and assume conceptual displacement rather than this alternative. Boghossian's memory argument thus relies on an understanding of slow switching that even he does not think essential to externalism. One other preliminary qualification is needed. I will portray Bernecker as though he always assumes the conceptual displacement view, but this is a little misleading. In 'Self-Knowledge and Closure', he briefly considers the possibility that for some time immediately after a switch, S may have a concept that refers to both water and twater (1998, 341). However, I take this point to be a relatively uncontroversial view of what happens at times immediately after switching. The view Boghossian considers more interesting holds that both concepts are available long after the time of switching, and Bernecker's defense of Ludlow does not seem to countenance this possibility.

[7] Michael Tye (1998) urges a similar view of memory. Since Jane Heal (1998) effectively rebuts Tye, I will not comment on Tye's account here.

[8] This formulation assumes that a memory is comprised of (at least) two components: mental content and propositional attitude.

[9] This is ironic, because Ludlow (1995c, 157) claims that "The natural place to focus an attack [on Boghossian] is on premiss (3), and indeed, a number of efforts have been aimed at undermining just this premiss, but for the most part it seems to me that these attempts have fallen short ... I want to suggest an alternative line of attack against Boghossian's argument from slow-switching and memory ... I will argue that the weak premiss in Boghossian's argument is premiss (1)".

[10] As quoted in Ludlow (1995b, 73).

[11] Ludlow misinterprets Hofmann as identifying a t_1-memory with a t_2-memory and arguing for a change in the truth-value of this memory between t_1 and t_2. In fact, Hofmann

is explicitly concerned with a t_2-memory of a t_1-experience. This does not, however, affect his substantive claim in this passage.

[12] I thank Michael Hymers for helping me clarify this argument.

[13] As quoted in Ludlow (1995b, 73). Brueckner (1997, 6) makes the same point. He asks us to consider

(M) I remember that I was thinking at t_1 that chicory is bitter.

and he says that

> ...according to Ludlow's externalism about memory, the content of my thought at t_1 (a content involving the concept of chicory) is different from the content expressed by the final that-clause in my t_2 utterance of (M) (a content involving the concept of twicory). In other words, I did not think at t_1 the thought that my t_2 utterance of (M) represents me as having thought. At t_1, I thought that chicory is bitter, but my utterance at t_2 represents me as having thought that twicory is bitter. Thus, my utterance of (M) is false and does not express any genuine memory. So it is hard to see how Ludlow's externalism about memory shows that from t_1 to t_2 'I forgot nothing'. The externalism about memory establishes exactly the opposite. It establishes that at t_2, I fail to remember what I thought at t_1. Thus if at t_1 I knew what I was thinking (as Ludlow maintains), then by t_2 I forgot something I knew at t_1.

Brueckner's claims about memory loss and forgetting are assessed in Section 1.1.2.1.

[14] Another way of putting the point might rely on the following analogy. A book purporting to be a book *of* poems must of course contain poems, not twoems. Presumably, Ludlow and Bernecker would agree. But just as surely, a book *about* water cannot be a book about twater, contrary to what Ludlow and Bernecker seem to think.

[15] In his recent book, Ludlow attempts to shore up his position as follows: "...suppose we hold that the timelessness of truth is preserved, but that the contents of our statements shift over time ...in [this] case, the episode of my thought E that occurs at t_0 has water content when I am at t_0 but twater content when I am at t_1. Thus, the thought might not be identified with a single content, but rather with a series of contents at different times" (1999b, 155). Accordingly, Ludlow urges the following reconstruction of Boghossian's memory argument:

(0) If a first-order thought E has content P at t_1, then it has content P at t_2.
(1) If S forgets nothing, then what S knows at t_1, S knows at t_2.
(2) S forgot nothing.
(3) At t_2, S does not know the content of E to be P.
(4) Therefore, at t_1, S did not know the content of E to be P.

Ludlow urges that "...the previously hidden premise (0) is seriously flawed, since E may well have the content P at time t_1 but something else altogether (say, Q) at t_2" (156).

This argument contains a tendentious (and heterodox) assimilation of 'statements' and 'thoughts'. Externalists do assert that the contents of statements may shift: the basic idea is that the linguistic expression "Water is thirst-quenching" can express different contents at different times. In contrast, externalists deny that the same thought E can have different contents at different times: *water*-thoughts and *twater*-thoughts are *different thoughts*, though they can be distinguished neither by phenomenology nor by ordinary language-use.

Ludlow is therefore not entitled to ascribe (0) to Boghossian. Presumably, his motivation for doing so is to develop a view according to which some unique thought E may be 'about' t_1, although it 'has' a t_2-concept. But such a view begs the question against Boghossian, and, in any event, I have sketched reasons for thinking that this understanding is no more plausible than its predecessor.

[16] Brueckner (1997) considers cases where, after multiple switches, the t_1-concept is (luckily) available to S at some later time. Brueckner is prepared to elevate such cases to the status of memory, and I object below in 1.2.2.

[17] Ludlow hints that this objection loses its force once a distinction is drawn between narrow and wide type-identifications of thoughts (1999a, 166). Roughly, such a distinction holds that a water-thought and a twater-thought are type-identical in terms of their non-relational contents ("... they are instantiated by the same data-structure, or perhaps ... they play the same functional role"), but type-distinct in terms of their relational contents. In simpler terms, the idea is that there is only *partial* memory failure on slow switching, since memory delivers content that is phenomenologically indistinguishable from pre-switch content. Though Ludlow never says so, I take it that this move is supposed to diminish the significance of his concession that all t_2-'memories' about t_1 are false. I disagree. This move is just another way of emphasizing the point that false memorial beliefs can be useful. While it is undoubtedly true that some fictions are useful, it is surely a drawback that Ludlow's account of memory renders all post-switch 'memories' false. Accounts of memory that avoid this drawback will therefore, *ceteris paribus*, be preferable.

[19] Ludlow (1995a) argues for the prevalence of slow switching.

[20] See note 6 above.

[21] Heal argues this point in a different fashion, suggesting that there is no reason to think that a slow switch is ever complete, where a complete switch is one in which S has lost all cognitive contact with Earth (1998, 106ff).

[22] Burge makes the same point a different way, noting that "[c]oncepts mark abilities; just moving around and acquiring new concepts will not in general obliterate such abilities, especially given that one still has uses for the old concepts and a perfectly good memory" (1998, 365, n. 13).

[23] This charge is repeated in Ludlow 1999a, 167.

[24] Burge (1998, 366–368) insists anyone who reasons in this manner makes a mistake of memory identification, not one of reasoning. Moreover, he asserts (in a different context) that such mistakes are non-culpable, because "[m]emory failures that cause demonstrations to fail are failures of background conditions necessary to the proper function of reasoning" (1993, 464).

[25] An anonymous *Erkenntnis* referee suggests that Ludlow and Bernecker might be better off to concede that concepts may be preserved on Putnamian Twin-Earth switching, but to deny that concepts can be preserved on earthbound switches between social groups. It is not clear to me that such a retreat would be helpful to Ludlow and Bernecker. Twin-Earth switching is admittedly more fanciful than earthbound social switching, but the two are relevantly similar with respect to the postulated effects of slow switching. That is, any evidence that suggests that concepts are *lost* on the former kind of switch will also be evidence that concepts are lost on the latter kind of switch. The switches are, after all, slow, which means that putative effects on S's conceptual array are brought about only after a significant period of interaction with the new environment or society. Features of this very interaction are held to affect membership in one's conceptual set, and there is no relevant

difference between the kind of interaction that takes place on Twin Earth and the kind of interaction that occurs in a post-switch social group.

[26] It might be objected that the logical possibility of error is insufficient to establish the possibility of fallibility, for though error may be logically possible, it may be physically or epistemically impossible. It does seem difficult to imagine actual cases of error here, but in order for this objection to be telling, a plausible account of such physical or epistemic impossibility must be given.

[27] Burge (1996) concedes that "... the special features of *cogito* cases do depend on present tense" (114).

[28] Rockney Jacobsen further restricts the class of infallible assertions, for he rightly points out that not all *first-person second-order present-tense assertions* are infallible (Jacobsen, R.: 'Self-Quotation and Self- Knowledge', *Synthese* **110**, 1997, 419–445). Jacobsen offers three examples: "I assert the same as you", "I assert what Kant asserted", and "I assert the same as I asserted yesterday" (427). Of these he notes that "... it does no harm to concede their fallibility: since I may have misheard what you said, misunderstood what Kant said, or misremembered what I said yesterday, and since you, Kant, or my earlier self, may not have asserted (or even uttered) anything at all, I may be mistaken in asserting that I assert the same as you, or Kant, or my earlier self" (427).

[29] As Gibbons puts the point,

> As an account of self-knowledge, [the inclusion theory] is incomplete, not only in terms of its sketchiness, but in principle. We have here, at most, an account of knowledge of content. We answer the question of how we know what we believe, a question about knowledge of content, in terms of content inheritance. We cannot answer the question of how we know that we believe something, a question about knowledge of the attitudes, in the same terms. The latter question is more difficult and has received much less attention. (1996, 294)

[30] Bernecker (1996) claims that "... though compatibilism explains knowing that it is P I believe, it doesn't explain how I can have privileged knowledge that the state I occupy is a state of believing rather than, say, a state of doubting, or a state of expecting, etc. But if I don't authoritatively know *that*, I cannot be said to possess privileged self-knowledge, for self-knowledge consists in the identification of the attitude as well as the content" (263).

[31] Of course, my point against Boghossian is unfair if Boghossian's position is that all knowledge must be fallible – even narrowly circumscribed subclasses of self-knowledge. Since I see no reason for anyone to hold this position, I do not impute it to Boghossian.

[32] An anonymous *Erkenntnis* referee suggests that it is unfair to demand significant explanatory power of the inclusion theory, since its stated goal is merely to secure the compossibility of externalism and self-knowledge. If any compossibility is established (even over a very narrow class of judgements) this goal has been met. I concede that the inclusion theory meets this goal, but I hold that this goal is too limited: much more work is required to show that externalism is compatible with the vast majority of our self-knowledge. I do not show, of course, that we actually have a vast domain of self-knowledge, but I take it that such a view is held by most participants in this debate. To the extent, then, that the inclusion theory fails to reconcile this view with externalism, it is radically incomplete.

[33] Burge has recently tried to defend a broad range of self-knowledge in a different way. He urges in 'Our Entitlement to Self-Knowledge' that this entitlement "... does not depend on the empirical content of the judgements. It does not depend on checking whether our

judgements meet certain conditions. It depends on the judgements' being of a kind essential to critical reasoning. Critical reasoning presupposes that people are entitled to such judgments. Since we are critical reasoners, we are so entitled" (1996, 115–116). In this paper, he seems to concede that little follows from the fact that there is a small class of self-verifying judgements: "... *cogito*-like thoughts are in many ways special cases. If we are to understand critical reasoning, the entitlement that I have discussed must apply more broadly. It must include judgements about beliefs, intentions, wants, as well as occurrent thoughts" (114). For criticisms of this paper, see Peacocke (1996) and MacDonald (1998).

REFERENCES

Bernecker, S.: 1996, 'Externalism and the Attitudinal Component of Self-Knowledge', *Noûs* **30**, 262–275.

Bernecker, S.: 1998, 'Self-Knowledge and Closure', in P. Ludlow and N. Martin (eds), *Externalism and Self-Knowledge*, CSLI Publications, Stanford, pp. 333–350.

Boghossian, P.: 1989, 'Content and Self-Knowledge', *Philosophical Topics* **17**, 5–26.

Brueckner, A.: 1997, 'Externalism and Memory', *Pacific Philosophical Quarterly* **78**, 1–12.

Burge, T.: 1988, 'Individualism and Self-Knowledge', *The Journal of Philosophy* **85**, 649–663.

Burge, T.: 1993, 'Content Preservation', *The Philosophical Review* **102**, 457–488.

Burge, T.: 1996, 'Our Entitlement to Self-Knowledge', *Proceedings of the Aristotelian Society* **46**, 91–116.

Burge, T.: 1998, 'Memory and Self-Knowledge', in P. Ludlow and N. Martin (eds), *Externalism and Self-Knowledge*, CSLI Publications, Stanford, pp. 351–370.

Davidson, D.: 1987, 'Knowing One's Own Mind', *Proceedings of the American Philosophical Association* **60**, 441–458.

Gibbons, J.: 1996, 'Externalism and Knowledge of Content', *The Philosophical Review* **105**, 287–310.

Goldberg, S.: 1997, 'Self-Ascription, Self-Knowledge, and the Memory Argument', *Analysis* **57**, 211–219.

Heal, J.: 1998, 'Externalism and Memory II', *Proceedings of the Aristotelian Society Supplemental Volume 72*, pp. 95–109.

Hofmann, F.: 1995, 'Externalism and Memory', Manuscript, Department of Philosophy, University of Tübingen.

Jacobsen, R.: 1997, 'Self-Quotation and Self-Knowledge', *Synthese* **110**, 419–445.

Ludlow, P.: 1995a, 'Externalism, Self-Knowledge, and the Prevalence of Slow-Switching', *Analysis* **55**, 45–49.

Ludlow, P.: 1995b, 'Social Externalism and Memory: A Problem?', *Acta Analytica* **14**, 67–76.

Ludlow, P.: 1995c, 'Social Externalism, Self-Knowledge, and Memory', *Analysis* **55**, 157–159.

Ludlow, P.: 1999a, 'First Person Authority and Memory', in M. De Caro (ed.), *Interpretations and Causes: New Perspectives on Donald Davidson's Philosophy*, Dordrecht, Kluwer Academic Publishers, pp. 159–170.

Ludlow, P.: 1999b, *Semantics, Tense, and Time: An Essay in the Metaphysics of Natural Language*, The MIT Press, Cambridge.

MacDonald, C.: 1998, 'Externalism and Authoritative Self-Knowledge', in C. Wright, B. Smith and C. MacDonald, *Knowing Our Own Minds*, Clarendon Press, Oxford, pp. 123–154.

Peacocke, C.: 1998, 'Our Entitlement to Self-Knowledge: Entitlement, Self-Knowledge, and Conceptual Redeployment', in P. Ludlow and N. Martin, *Externalism and Self-Knowledge*, CSLI Publications, Stanford, pp. 265–303.

Tye, M.: 1998, 'Externalism and Memory I', *Proceedings of the Aristotelian Society Supplemental Volume 72*, pp. 77–94.

Manuscript submitted December 20, 2000
Final version received May 29, 2001

Department of Philosophy
University of Toronto
215 Huron Street
Toronto, ON
Canada M5S 1A1
E-mail: kkraay@chass.utoronto.ca