
BROWN ON SELF-KNOWLEDGE AND DISCRIMINABILITY¹

BY

SANFORD C. GOLDBERG

Abstract: In her recent book *Anti-Individualism and Knowledge*, Jessica Brown has presented a novel answer to the self-knowledge ‘achievement problem’ facing the proponent of anti-individualism. She argues that her answer is to be preferred to the traditional answer (based on Burge, 1988a). Here I present three objections to the claim that her proposed answer is to be preferred. The significance of these objections lies in what they tell us about the nature of the sort of knowledge that is in dispute. Perhaps the most important lesson I draw from this discussion is that, given the nature of knowledge of one’s own thoughts, discriminability (from relevant alternatives) is not a condition on knowledge as such.

1. Introduction

In her (2004), Jessica Brown argues that the discriminability objection to the compatibility of anti-individualism and privileged access is not adequately addressed by the traditional argument for compatibilism (based on Burge, 1988a). Brown thus proposes to replace the traditional argument with her own preferred argument, which makes use of the distinction between relevant and irrelevant (knowledge-undermining) alternatives. Here I argue, first, that the case she offers for preferring her compatibilist argument to the traditional Burge-inspired one is weakened by the fact that her reconstruction is not faithful to the spirit of the Burge-inspired compatibilist argument; second, that the discriminability requirement on knowledge is not happily applied to the domain of judgements aiming at the sort of self-knowledge at issue; and, finally, that even if it were otherwise acceptable, Brown’s own preferred compatibilist position (unlike the traditional position) fails to preserve a core thesis compatibilists should want to preserve. My narrow conclusion is that proponents of AI should

Pacific Philosophical Quarterly 87 (2006) 301–314

© 2006 The Author

Journal compilation © 2006 University of Southern California and Blackwell Publishing Ltd.

prefer the traditional response over Brown's proposed alternative. But the more important lesson goes beyond my criticism of Brown: given the nature of the sort of knowledge one has when e.g. one knows that one oneself is thinking that *p*, discriminability (from relevant alternatives) cannot be a necessary condition on knowledge as such.²

2. *Brown's dialectic*

Anti-individualism (henceforth 'AI') is the thesis that "the mental natures of many of an individual's mental states and events are dependent for their individuation on the individual's social and physical environments" (Burge, 1986, p. 697). (I use Burge's general formulation of AI, since Brown holds (2004, p. 135) that the discrimination objection applies to all versions of AI.) As formulated by Brown, the doctrine of privileged access (henceforth 'PA') is the claim that "a subject can have *a priori* knowledge of her thought contents" (Brown, 2004, p. 36). The compatibility of these two doctrines has been a much-discussed issue in the recent literature.³ One way in which some have thought to challenge the compatibilist claim is by raising what Brown calls the "achievement problem", which is the problem of making clear how it can be that one can have "privileged access to one's thoughts" even though "what thoughts one has . . . depends on environmental facts of which one may be ignorant" (2004, p. 234).⁴

The traditional reply to the achievement problem is owed to Burge, 1988a.⁵ This reply appeals to the self-verifying nature of the relevant judgements – judgements of the form '(With this very thought) I am presently thinking that *p*'. (I will designate the members of the relevant class of first-person present-tense judgements 'FPPT judgements'.) On Brown's reconstruction, the appeal to the self-verifying nature of FPPT judgements is made in order to establish the *reliability* of these judgements in both actual and counterfactual environments. Since this reliability is independent of the truth of AI, the result (on this reconstruction) is that, on the assumption of a reliabilist conception of knowledge, FPPT judgements amount to knowledge whether or not AI is true. Since FPPT judgements are not based on any empirical grounds – at least the judgements are not based on anything beyond what is available to reflection – the knowledge in question is *a priori* in the minimal sense that it is available from the armchair. Hence the achievement problem is met, according to (Brown's construal of) the traditional reply.

But Brown points to what she regards as an important weakness with this reply, namely, that it will fail to satisfy anyone who thinks that there is a *discriminability condition* on knowledge. To show this, she offers the "discriminability argument". The two-step strategy behind such

an argument is, first, to establish that if AI is true then for any subject *S* currently thinking that *p*, there is some counterfactual situation, indiscriminable (by *S*) from the actual situation, in which *S* is *not* thinking that *p*; and, second, to use this first result, together with the discriminability requirement on knowledge, to conclude that if AI is true then (whether or not *S*'s FPPT judgement is reliable) *S* fails to know that she is presently thinking that *p* (and so fails to know *a priori* that she is presently thinking that *p*).

On this score Brown points out that (on the assumption of AI) the situation of reliability-without-discriminability obtains with respect to judgements regarding one's own occurrent thoughts, in cases with a *slow switched* subject who remains in ignorance of her switching regimen (Brown, 2004, pp. 45–59). Suppose that, prior to the start of her switching regimen, Sally would make an FPPT judgement by way of an assertion of:

(W) (With this very thought) I am presently thinking that water is wet.

Since at this point Sally has been a life-long resident of Earth, she thereby succeeds in expressing and self-ascribing a WATER-thought (= a thought involving the concept WATER). But, assuming AI is true, there comes a point after the switching regimen where Sally's use of 'water' will express something other than WATER (say, TWATER) – in which case the very same (morphological) sentence-type in her mouth will express a TWATER-thought. Since Sally is ignorant of her switching regimen, she will fail to be able to distinguish between situations in which she is self-ascriptively thinking a WATER-thought, and those in which she is not (e.g. because she is self-ascriptively thinking a TWATER-thought). So on the assumption that discriminability is a necessary condition on knowledge, the self-ascriptive judgement she expresses with (W) does not amount to knowledge of the thought she is thinking – notwithstanding the reliability of the judgment.

It is in this dialectical context that Brown offers her own reply to the achievement problem. The main selling-point of her proposal is that, unlike the traditional reply, her proposal successfully meets the challenge posed by the discriminability argument. Her proposal is to endorse a relevant alternatives conception of knowledge, on which (roughly) one's true belief that *p* amounts to knowledge that *p* iff one's belief that *p* is based on grounds that are inconsistent with any relevant alternatives (= relevant scenarios in which $\sim p$ holds). If we endorse such a view, Brown notes, we can respond to the discriminability argument straightforwardly: we need only note that, in the self-knowledge cases under discussion, the alternative knowledge-undermining scenarios are normally *not relevant* (2004, pp. 136–37). Since knowledge requires discriminability only with regard to relevant alternatives, the result (Brown reasons) is that, even assuming AI,

normally a thinker who expresses an FPPT judgment by way of a sentence such as (W) can be credited with *a priori* knowledge of her thought.

3. *Objections to Brown's proposal*

I have three objections to the claim that Brown's proposal is to be preferred to the traditional (Burge-inspired) reply to the achievement problem.

3.1

The first is that Brown's proposal is not well-motivated. As noted, her reconstruction of the traditional reply to the achievement problem regards that reply as depending on the sufficiency (for knowledge) of cross-world reliability. Although such a reconstruction is arguably in the spirit of some replies to the achievement problem,⁶ is not in the spirit of Burge's own reply. On the contrary, the thrust Burge's (1988a) reply to the achievement problem is two-fold. First,⁷ FPPT judgements enjoy a sort of epistemic support that is both *truth-guaranteeing* and *introspectively accessible* in the following sense: a subject who possessed all of the relevant (semantic and epistemic) concepts would be in a position to appreciate from the armchair the truth-guaranteeing nature of FPPT judgments. And second, subjects who make such judgements enjoy an *epistemic entitlement* to do so.⁸ Given its appeal to the notion of an epistemic entitlement, this two-part answer to the achievement problem is not happily regarded as depending on the sufficiency (for knowledge) of mere cross-world reliability. Consequently, even if the discriminability argument provides a reason to prefer Brown's to *certain* other proposed answers to the achievement problem, it does not provide a reason to prefer hers to the traditional (Burgean) answer.

This first objection depends on the soundness of Burge's reply to the achievement problem: if Burge's reply is not sound, then Brown need not worry about it, and so can rest her case having shown that hers is to be preferred to those replies (of which there are several) that depend on reliabilist conceptions of knowledge. Let us then see whether Brown's argument has the resources to establish that Burge's response is not sound. She can do so if, and it would appear only if, she can reject any one of the following claims: (i) that FPPT judgments are self-verifying; (ii) that the self-verifying nature of FPPT judgements is introspectively accessible in the sense noted above; or (iii) that judgements with these two properties amount to knowledge. (I disregard for the moment Burge's notion of an entitlement; it will arise when we consider how (iii) might be challenged, as one way to do so would be to challenge the claim that thinkers are

entitled to make judgements possessing the first two properties.) I want to argue that Brown cannot successfully challenge any of (i)–(iii).

How might Brown think to challenge (i), the claim that FPPT judgments are self-verifying?⁹ She can deny that FPPT judgements are self-verifying by showing one such judgement to be false. To do this, what would have to be shown is that a situation is possible in which a subject makes an FPPT judgement, yet where in fact she is not thinking the thought she is self-ascribing. Attempts to make out such a view appear doomed to self-defeat. Suppose (for *reductio*) that *S*'s FPPT judgment amounts to a self-ascription of the thought that *p*, under conditions in which *S* is *not* thinking that *p*. Since *S* is not thinking that *p*, it is false that *S* judges that she is thinking that *p*. This reflects what appears to be a truism: judging that one is thinking that *p* involves thinking that *p* – as this first-order thinking is what *fixes the subject-matter* of the self-ascriptive judgment regarding one's thought. But if it is false that *S* judges that she is thinking that *p*, then our original supposition is false: it is not the case that *S*'s FPPT judgment amounts to a self-ascription of the thought that *p*, under conditions in which *S* is not thinking that *p*. In this way, the attempt to undermine the self-verifying nature of FPPT judgements is self-defeating.

Nor is the present defense of (i) undermined by appeal to discriminability considerations. To see this we need only note the work done by discriminability considerations in the context of epistemic appraisals. Take the familiar case in which a subject *S* sees what she regards as zebras in the local zoo. Suppose that, on the basis of sight, *S* cannot discriminate zebras from cleverly painted mules. And let *H* be the hypothesis that the animals in question are in fact cleverly painted mules. If *H* is relevant to *S*'s circumstances, then, given a discrimination condition on knowledge, *S* does not know (by sight) that those animals are zebras. It is noteworthy here that the discriminability condition would establish this conclusion *whether or not* the animals *S* sees are, in fact, zebras. The point of the discriminability condition, then, is not to address the *truth-value* of *S*'s judgement that the animals *S* sees are zebras; it is rather to address the *epistemic goodness* (as it were) of the grounds of *S*'s judgement. Since this is the way that discriminability considerations generally figure in epistemic appraisals, it would seem that there is no direct route, from the claim that a subject *S* fails the discriminability condition despite making an FPPT judgement, to the conclusion that the FPPT judgement is *false*. So if Brown hopes to reject (i) by showing that FPPT judgements can be false, she must make her case on the basis of more than the mere appeal to discriminability considerations.

Can Brown argue that (i) is false by appeal to world-switching cases? It might seem so. After all, I noted above that in some such cases, the subject's ignorance of her switching regimen will put her in a position in which she will fail to be able to distinguish between situations in which

she is self-ascriptively thinking a WATER-thought, and those in which she is not (because she is self-ascriptively thinking e.g. a TWATER-thought). If Brown could show that there is at least one such case in which the subject would make a *false* judgement regarding her own occurrent thought, Brown would have shown what needs to be shown. But can this be done? I submit that it cannot, precisely for the reason Burge, 1988a gave (and which Brown, 2004 appears to have accepted): in making FPPT judgements, *one thinks the thought in the very act of self-ascribing it*. Nothing about the world-switching scenario described touches this point. But, still, it might be wondered, how can there be a situation in which the subject's judgement regarding her own thought is guaranteed to be true, yet where the subject fails to be able to discriminate the actual case (in which she is thinking a WATER-thought) from counterfactual cases (in which she is not thinking a WATER-thought)? Doesn't the guaranteed truth of her FPPT judgement undermine her discriminatory inability? In fact, as Brown herself acknowledges throughout her discussion (pp. 43–73), such cases are indeed possible: even perfect reliability does not entail discriminability. As she herself notes, “the slow switch case itself shows us that the reliable ability to form true beliefs about which of two types is instantiated is not sufficient for the ability to distinguish those two types” (p. 48).

Brown's idea on this score is as follows. Given their self-verifying nature, FPPT judgements are invariably true. Consequently, given any situation in which one finds oneself, one is always in a position to judge which, of the various possible content-types one might be thinking, one is actually thinking: one can merely form the relevant FTTP judgment. However, this does not establish that one can discriminate situations in which one is thinking a first-order thought of that content, from counterfactual situations in which (because of one's alternative embedding history) one would have been thinking some “twin” content. Brown effectively argues that it will not do for the subject simply to repeat the FPPT judgement itself; for while it has already been established that such a judgement would be true, even so, the slow-switched subject who makes such a judgment would lack *other* abilities indicative of the presence of the discriminatory ability. So for example such a subject would lack “the abilities to notice change, to make reliable judgements of sameness and difference, and to act differentially” (Brown, 2004, p. 49). Brown's case for this is compelling, and so we can conclude that even perfect reliability is compatible with the lack of relevant discriminatory abilities. However this just shows that (i) is consistent with the lack of discriminatory abilities; so one cannot appeal to such a lack in order to argue that (i) is false. (Again, this is no criticism of Brown, who does not argue in this way.)

It seems, then, that (i) is secure: FTTP judgements are self-verifying. Might Brown then think to challenge the soundness of Burge's reply to

the achievement problem by challenging (ii), the claim that the self-verifying nature of FPPT judgements is *introspectively accessible*, in the sense that a subject who possessed all of the relevant (semantic and epistemic) concepts would be in a position to appreciate these features from the armchair? There is an immediate difficulty. Above I presented a defense of the self-verifying nature of FPPT judgements; my claim was that attempts to *deny* the self-verifying nature of FPPT judgements will be self-defeating. But this self-defeatingness was demonstrated by appeal to considerations available from the armchair. For the argument itself involved nothing beyond some trivial points about the conditions under which one would count as thinking that *p*, and the conditions under which one would count as self-ascribing the thought that *p*. Or so it would seem; those who would seek to deny (ii) have the burden of showing otherwise, and I for one don't see how this can be done.

Let us turn, finally, to (iii). What are the prospects of challenging the soundness of Burge's reply to the achievement problem, by challenging the thesis that judgements regarding which both (i) and (ii) hold amount to *knowledge*? It is with respect to (iii), I suspect, that Brown will mount an objection. The best prospect for such an objection would be by appeal to the discriminability condition on knowledge. However, I want to argue that such an appeal is *dialectically unstable*: given any judgement of which (i) and (ii) hold, the application of the discriminability condition on knowledge will have very unhappy results. I bring this out in my second objection to Brown's reply, which follows. But in advance it is worth noting that if I am correct in this, then the defender of Brown's reply cannot plausibly deny any of (i)–(iii), and so cannot plausibly maintain that Burge's reply to the achievement problem is unsound. And this, in turn, would vindicate my objection that Brown's appeal to discriminability considerations does not give us a reason to prefer her proposal to Burge's own.

3.2

Whether or not it is in the service of her attempt to deny (iii), Brown's application of the discriminability condition on knowledge to FPPT judgements is dialectically unstable. To see this, consider the case of Sam. Sam, a philosophy major with a particular interest in the epistemic consequences of AI, has read Burge (1988a) very carefully. (Needless to say, she has all of the semantic and epistemic concepts needed to follow Burge's reasoning.) Now on a particular occasion, Sam forms an FPPT judgement which she expresses with an assertoric utterance of (W). Since she has read Burge, she has a reflective appreciation for the self-verifying nature of the judgement she just made (she can repeat the reasoning for (i) to herself, without leaving the armchair). Since she has reflective

grounds for thinking that her FPPT judgement is self-verifying, she has reflective grounds for thinking that the judgement is true. And indeed it is true (as guaranteed by the very reasoning she herself articulates). What is more, Sam is also in a position to know from the armchair that her judgement regards her present thought: in uttering (W), she is giving expression to knowledge to the following effect: with that very thought, she herself is presently thinking that water is wet. In sum: when Sam makes a judgement of the form '(With this very thought) I am presently thinking that *p*', she is making a judgment that is *guaranteed* to be true (= Sam is indeed presently thinking that *p*), where Sam herself *appreciates the basis* of this guarantee, and appreciates as well what this guarantee is a guarantee *of* (namely, the truth of her thought-self-ascription). Given what appears to be Sam's excellent epistemic situation *vis-à-vis* her judgement that (with that very thought) she herself is presently thinking that *p*, it would seem that any reasonable theory of knowledge ought to grant that this judgement amounts to knowledge.

Consider then the effect in this dialectical context of an insistence on a discrimination condition on knowledge. Suppose Sam has been the victim of slow-switching regimen, such that she cannot discriminate between the case in which she is self-ascriptively thinking that *p*, and the case in which she is not (e.g. because she is thinking some "twin" content). And suppose that she forms an FTTP judgment by way of an assertoric utterance of (W). By the lights of the discrimination condition, her FTTP judgment does not amount to knowledge – and this, despite the facts that (a) this judgment is *guaranteed* to be true, (b) Sam herself *appreciates the basis* of this guarantee, and (c) Sam appreciates as well what this guarantee is a guarantee *of* (namely, the truth of her thought-self-ascription). But this outcome is absurd: better we should use (a)–(c), together with our intuition (supported by (a)–(c)) that she has judged *knowledgeably*, as the basis for rejecting the doctrine that *discrimination is required for knowledge* in such cases.

I anticipate that Brown might object that the foregoing argument begs the question against the discriminability requirement on knowledge. After all, given such a requirement, Sam does *not* count as knowing that (with that very thought) she herself is thinking that *p*. In reaction to the charge of begging the question, it is worth pointing out that, however we ultimately resolve the issue of whether Sam's FPPT judgement amounts to knowledge, Sam's epistemic situation (as it bears on the epistemic standing of the FPPT judgement she has just made) is a strong one indeed. This is seen in the fact that, as argued above, conditions (a)–(c) hold. Suppose that Brown grants (a)–(c); must she then grant that, for these reasons, Sam's FPPT judgement itself amounts to Sam's knowledge that (with that very thought) she herself is thinking that *p*? One way Brown might try to resist this would be to deflate the epistemic significance of

(a)–(c), e.g. by suggesting that at best these points put Sam in a position to acquire a piece of *metacognitive* or *metalinguistic* knowledge. This would be the case if, for example, the *de dicto* content of Sam's knowledge is something that she herself might express by uttering, "The judgement I am now making, when I judge that (with this very thought) I am presently thinking that water is wet, is true". Such a metacognitive construal of Sam's knowledge might give Brown a way to accept (a)–(c) without treating Sam as knowing that (with that very thought) she herself is thinking that water is wet.

But the metacognitive construal is unmotivated in the extreme. We see this when we see *what* piece of knowledge a subject is in a position to have when she has a reflective appreciation of the reasoning of Burge (1988a). It is true that, having digested such reasoning, a subject is in a position to have a general piece of knowledge, to the effect that all FPPT judgements are true. (This general piece of knowledge can then be applied to particular cases, yielding metacognitive knowledge of the sort just described.) But more than this, a subject who has digested the reasoning of Burge (1988a) is *also* in a position to have the particular pieces of knowledge expressed by the particular FPPT judgements themselves. This follows on the simple assumption that *in making an FPPT judgment one thinks the thought in the very act of self-ascribing it*. For insofar as one is thinking the thought in the very act of self-ascribing it, one's judgement will amount to a correct judgement of the very thought one is then thinking; and this, together with the guaranteed truth of the judgment and the subject's reflective appreciation of the basis for the judgement's truth, renders *that very judgement* knowledge. I do not see how this primitively compelling line of argument can be resisted. In particular, it would simply be a bald assertion to resist this conclusion by arguing that what is known in such a case is something metacognitive or metalinguistic. And, since the metacognitive proposal would appear to be the only way for Brown to resist the conclusion in question – the conclusion, namely, that Sam knows something to the effect that (with that very thought) she herself is currently thinking that *p*, despite the fact that Sam is not in a position to discriminate this case from other cases in which she would not be currently thinking that *p* – I conclude that, in the case at hand, Sam has the relevant piece of self-knowledge despite failing the discrimination condition.

This last result paves the way for making a point of more general epistemological interest, which is simply that discriminability (from relevant alternatives) is not a necessary condition on knowledge. It is worth presenting the case for this from a point of view that is not so directly linked to the particularities of Brown's position. Most everyone will concede that, in addition to true belief, knowledge requires some sort of positive (truth-indicative) epistemic standing.¹⁰ However, it is controversial what

sort of “epistemicizing” feature(s) such a standing supervenes on.¹¹ One suggestion is provided by the relevant alternative conception of knowledge:

- (RA) For all true beliefs that p , such a belief is knowledge iff it is based on grounds that are incompatible with all relevant alternative scenarios (in which the belief would be false).

I am arguing that the case of self-knowledge – at least the sort of self-knowledge manifested in FTTP judgements – is a counterexample to RA. The counterexample is brought out by focusing on cases in which the subject herself is familiar with the Burgean reasoning leading up to the conclusion that her FPPT judgement is self-verifying, and where she presents an explicit Burge-style defense of her claim to know that (with that very thought) she herself is thinking that p .¹² In such a case, her FTTP judgement is based on grounds that *guarantee its truth*, and the subject herself is *reflectively aware* of this fact. It should be uncontroversial that in such a case, the subject counts as knowing something to the effect that (with that very thought) she herself is thinking that p . That is: whatever the proper account of the epistemicizing feature turns out be, this will be a case in which such a feature is present. The result is that, if Brown is correct in her contention that a slow-switched subject may fail to be able to discriminate the actual situation (in which she is self-ascriptively thinking that p) from one in which she is not self-ascriptively thinking that p ,¹³ discrimination is not a necessary condition on the possession of the sort of self-knowledge in question. The falsity of RA follows trivially: discrimination (from relevant alternatives) is not a necessary condition on knowledge as such.^{14,15}

3.3

So far, I have presented two objections to the claim that Brown’s alternative response to the achievement problem is to be preferred to the traditional, Burge-inspired response: first, her response is not adequately motivated *vis-à-vis* Burge’s response; and second, her response depends on what, in the context of FPPT judgements, is a dialectically precarious appeal to the discriminability condition on knowledge. I turn now, briefly, to my third objection: even if otherwise motivated, Brown’s reply to the achievement problem is ultimately too concessive. Brown’s position concedes that in exotic circumstances, where the knowledge-undermining scenarios *are* relevant, the subject fails to know her own occurrent thoughts (if AI is true). Since the scenarios in question are exotic, Brown regards this concession as negligible (2004, pp. 136–37). But even granting that the knowledge-undermining scenarios are exotic, and hence are rarely relevant, even so Brown’s concession is not negligible. The core

worry behind the achievement problem is not that AI makes failures of self-knowledge *too frequent* an occurrence; rather, it is that AI makes self-knowledge hostage to “external” changes in the first place. A key intuition on this score is that changes in the “external” environment should *never* get in the way of a thinker’s achieving non-empirical knowledge of her own thoughts. (In the Cartesian demon scenario, the intuition is that, whatever other evils he is capable of, the demon *cannot* so arrange things as to ensure that I fail to achieve non-empirical knowledge of one of my own thoughts.) Yet it is precisely this key intuition that Brown’s position surrenders. To those who want to honor such an intuition – and surely this includes all incompatibilists and many would-be compatibilists¹⁶ – Brown’s concession will seem to be *too* concessive. It is here, of course, that we see one other virtue of the traditional case for compatibilism, as developed above: if it can be made to work, it would preserve the intuition in question, for the simple reason that Burge-style reasoning about the self-verifying nature of the relevant judgements applies, and is available for being thought, in all contexts.

4. *Lessons*

I have just presented three reasons for preferring the traditional reply, over Brown’s proposed alternative reply, to the achievement problem. But I focus on Brown’s argument because doing so highlights what I consider to be some important characteristics of the sort of self-knowledge expressed in FPPT judgements. Here I outline two such characteristics, as these have emerged in my discussion of Brown’s view. First, the epistemic status of FTTP judgements is best cast, not in terms of brute-reliability, but rather in terms of the two features of such judgements as are highlighted by (my reconstruction of) Burge’s own reply to the achievement problem: their *self-verifying* nature, and the (in principle) *introspective availability* of this self-verifying nature to any subject possessing the relevant semantic and epistemic concepts. Second, and more importantly, when it comes to the knowledge a subject has of her own occurrent thoughts, she can *know* that she is occurrently thinking that *p* even as she fails to be able to discriminate the actual case from a counterfactual case in which she is not occurrently thinking that *p* (e.g. because she is thinking some other “twin” content). To say this just is to say that we ought not to impose a discriminability condition on such self-knowledge. If so, then it would be wrong to think that discriminability (from relevant alternatives) is a condition on knowledge as such.

Department of Philosophy
University of Kentucky

NOTES

¹ Thanks to Jessica Brown for discussion; and to an anonymous referee, for helpful comments on an earlier version.

² Although nothing I say here will tell against treating discriminability as a necessary condition on knowledge in *certain domains* – perceptual knowledge, say.

³ For some recent papers on the topic, see the selections in five recent volumes: Ludlow and Martin (1997), Wright, Smith, and Macdonald (1998), Nuccetelli (2003), Frapolli and Romero (2003), and Schantz (2004).

⁴ The other way in which some have sought to argue against the compatibility of AI and PA is to raise what Brown calls the “consequence problem”, regarding whether the assumption of AI and PA has unacceptably strong epistemic consequences. This will not be at issue here. (See Goldberg, 2003a and 2003c for a discussion.)

⁵ A version of what is arguably the same sort of reply can be found in Heil, 1988.

⁶ See e.g. Falvey and Owens, 1994 and Gibbons, 1996.

⁷ Admittedly, Burge himself does not quite put his argument this way; but in Goldberg, 2005 I argued that such a formulation captures the spirit of Burge’s position.

⁸ For the notion of an entitlement, see Burge, 1998a and 1998b, p. 241.

⁹ I should note that in her (2004), Brown takes (i) for granted. (Thus, summarizing her argument, she writes, “as we have seen, a subject’s *cogito* thoughts cannot be mistaken” (p. 48).) But since the combination of (i)–(iii) pose a problem for her view, I am examining whether she ought to have challenged (i).

¹⁰ There are two traditional ways to deny this: first, by holding that knowledge just is true belief; second, by holding that knowledge is not further analyzable, hence not analyzable into true belief plus some additional epistemicizing feature. For the former see e.g. Goldman, 2002 (where Goldman speaks of “weak knowledge”); and for the latter, see e.g. Williamson, 2000.

¹¹ I borrow the term ‘epistemicizing’ from Plantinga, 1993.

¹² I note in passing that such reasoning may not be reflectively accessible to many, or perhaps most, subjects: either they will lack the relevant semantic or epistemic concepts, or they will not be sufficiently sophisticated to be able to produce the sort of Burge-style defense of their own FPPT judgements, etc. However, this point does not affect the present argument, which merely depends on the existence of a subject for whom such reasoning *is* reflectively accessible. See e.g. Burge, 1998a and 199b, in addition to Goldberg, 2005, for a discussion of how the present point can be developed into a general account of knowledge via FPPT judgement.

¹³ I have not argued for this claim here, or defended Brown’s own argument for it (Brown, 2004, pp. 45–59). But it is widely accepted by others (see e.g. Burge, 1988a, 1988b, 1998a, 1998b; Boghossian, 1989, 1992, and 1994; Owens, 1986; Falvey and Owens, 1994; and Brown, 2000); and in any case I have argued for it elsewhere (see Goldberg, 2000, 2003a, and 2003b.)

¹⁴ Of course, it is one thing to say that the appeal to discriminability condition on knowledge is dialectically inappropriate in the context of FPPT judgements; it is another to provide a positive account of what makes those judgements *knowledge* even in cases in which the subject is not reflectively aware of the Burge-style reasoning. For his part, Burge appeals to the notion of an epistemic entitlement at this point in the dialectic. Such an account is well worth pursuing; but it is one that I do not pursue here (but see Burge, 1998a and 1998b, as well as Goldberg, 2005, for an extended discussion of these matters).

¹⁵ Arguably, the discriminability condition fails even in self-knowledge cases having nothing to do with anti-individualism. Consider the self-knowledge that I express when I assert (1) and (2):

- (1) I am now entertaining the thought that whoever believes that all bachelors are bachelors believes that all bachelors are unmarried men.
- (2) I am now entertaining the thought that whoever believes that all bachelors are bachelors believes that all bachelors are bachelors.

I can understand both, both of these can be true of me, and I recognize this; but I need not know if the thought I know myself to be entertaining in (1) is the same or different as the thought I am entertaining in (2). (I thank an anonymous referee from this journal for this very case; see also Burge, 1978 and Owens, 1986 for detailed discussions of such cases.)

¹⁶ As Brown herself notes, the exceptions to this are Warfield, 1998 and Sawyer, 1998, both of whom offer compatibilist arguments that in effect surrender this intuition.

REFERENCES

- Boghossian, P. (1989). "Content and Self-Knowledge," *Philosophical Topics* 17.
- Boghossian, P. (1992). "Externalism and Inference," *Philosophical Issues* 2.
- Boghossian, P. (1994). "The Transparency of Mental Content," in J. Tomberlin (ed.) *Philosophical Perspectives 8: Language and Logic*. Ridgeview: Atascadero.
- Brown, J. (2000). "Critical Reasoning, Understanding and Self-Knowledge," *Philosophy and Phenomenological Research* 61, pp. 659–676.
- Brown, J. (2004). *Anti-Individualism and Knowledge*. Cambridge, MA: MIT Press.
- Burge, T. (1986). "Intellectual Norms and the Foundations of Mind," *Journal of Philosophy* 83, pp. 697–720.
- Burge, T. (1978). "Belief and Synonymy," *Journal of Philosophy* 75, pp. 119–38.
- Burge, T. (1988a). "Individualism and Self-Knowledge," *Journal of Philosophy* 85, pp. 649–63.
- Burge, T. (1988b). "Wherein is Language Social?" in *Reflections on Chomsky*. Oxford: Basil Blackwell.
- Burge, T. (1998a). "Our Entitlement to Self-Knowledge," in Martin and Ludlow (eds.) *Externalism and Self-Knowledge*.
- Burge, T. (1998b). "Reason and the First-Person," in C. Wright *et al.* (eds.) *Knowing Our Own Minds*.
- Falvey, K. and Owens, J. (1994). "Externalism, Self-Knowledge, and Skepticism," *Philosophical Review* 103, pp. 107–137.
- Frapolli, M. J. and Romero, E. (eds.) (2003). *Meaning, Basic Self-Knowledge, and Mind*. Palo Alto: CSLI Publications.
- Goldberg, S. (2000). "Externalism and Self-Knowledge of Content: A New Incompatibilist Strategy," *Philosophical Studies* 100, pp. 51–78.
- Goldberg, S. (2003a). "Anti-Individualism, Conceptual Omniscience, and Skepticism," *Philosophical Studies* 116, pp. 53–78.
- Goldberg, S. (2003b). "What Do You Know When You Know Your Own Thoughts?" in S. Nuccetelli (ed.) *New Essays on Semantic Externalism and Self-Knowledge*. Cambridge, MA: MIT Press, pp. 241–56.
- Goldberg, S. (2003c). "On Our Alleged *A Priori* Knowledge That Water Exists," *Analysis* 63, pp. 38–41.
- Goldberg, S. (2005). "The Dialectical Context of Boghossian's Memory Argument," *Canadian Journal of Philosophy* 35, pp. 135–48.
- Goldman, A. (2002). "What is Social Epistemology? A Smorgasboard of Projects," in A. Goldman, *Pathways to Knowledge: Public and Private*. Oxford: Oxford University Press.
- Gibbons, J. (1996). "Externalism and Knowledge of Content," *Philosophical Review* 105.

- Heil, J. (1988). "Privileged Access," *Mind* 97, pp. 238–51.
- Martin, N. and Ludlow, P. (eds.) (1998). *Externalism and Self-Knowledge*. Palo Alto: CSLI Publications.
- Nuccetelli, S. (ed.) (2003). *New Essays on Semantic Externalism and Self-Knowledge*. Cambridge, MA: MIT Press.
- Owens, J. (1986). "Synonymy and the Non-Individualistic Model of the Mental," *Synthese* 66, 361–382.
- Plantinga, A. (1993). *Warrant: The Current Debate*. Oxford: Oxford University Press.
- Sawyer, S. (1998). "Privileged Access to the World," *Australasian Journal of Philosophy* 76, pp. 523–533.
- Sawyer, S. (1999). "An Externalist Account of Introspective Self-Knowledge," *Pacific Philosophical Quarterly* 80, 358–78.
- Schantz, R. (ed.) (2004). *The Externalist Challenge*. Berlin: Walter de Gruyter.
- Warfield, T. (1997). "Externalism, Self-Knowledge, and the Irrelevance of Slow Switching," *Analysis* 57, 282–284.
- Warfield, T. (1998). "A Priori Knowledge of the World: Knowing the World by Knowing our Minds," *Philosophical Studies* 92, pp. 127–47.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.
- Wright, C., Smith, B. C. and Macdonald, C. (eds.) (1998). *Knowing Our Own Minds*. Oxford: Oxford University Press.