

URIAH KRIEGEL

MOORE'S PARADOX AND THE STRUCTURE OF CONSCIOUS BELIEF

ABSTRACT. Propositions such as <It is raining, but I do not believe that it is raining> are paradoxical, in that even though they can be true, they cannot be truly asserted or believed. This is Moore's paradox. Sydney Shoemaker has recently argued that the paradox arises from a constitutive relation that holds between first- and second-order beliefs. This paper explores this approach to the paradox. Although Shoemaker's own account of the paradox is rejected, a different account along similar lines is endorsed. At the core of the endorsed account is the claim that conscious beliefs are always partly about themselves; it will be shown to follow from this that conscious beliefs in Moorean propositions are self-contradictory.

1. INTRODUCTION: MOOREAN PROPOSITIONS, MOOREAN ASSERTIONS, AND MOOREAN BELIEFS

Moore's paradox pertains to propositions such as:

- (1) It is raining and I do not believe that it is raining.

Call such propositions – propositions of the form < p & I do not believe that p >, or as I will occasionally write it, < p & \sim IB p > – *Moorean Propositions*.^{1,2} Moorean propositions are paradoxical, in that they (i) cannot be truly asserted, even though they (ii) can be true.³ This is in contrast to propositions of the form:

- (2) It is raining and it is not raining.

These are *not* paradoxical, because they *cannot* be true. Asserting (2) is absurd, then, but not paradoxical, since there is nothing puzzling about its absurdity: the fact that (2) cannot be truly asserted is easily explained in terms of the fact that it cannot be true.

The trouble with Moorean propositions, then, is that although they can be true, assertions of them cannot.⁴ Call assertions of Moorean propositions *Moorean assertions*. Another problem Moorean propositions present is that they cannot be truly *believed*. Call believings of Moorean propositions *Moorean beliefs*.⁵ In Section 2, I will argue that Moorean assertions are absurd *because* they express



Moorean beliefs – more specifically, *conscious* Moorean beliefs. This suggests that a key to Moore's paradox may be the consciousness involved in Moorean beliefs. The account of Moore's paradox I end up endorsing (in Section 4) appeals crucially to the structure of conscious belief. I will reach this account through an examination, in Section 3, of Shoemaker's recent discussion of the paradox.

2. PSYCHOLOGISM ABOUT MOORE'S PARADOX

Moore's paradox, as originally discussed by Moore and Wittgenstein, concerned Moorean *assertions* exclusively. More recently, philosophers have taken interest in the paradoxical nature of Moorean *beliefs* as well. It has become a widely accepted constraint on the adequacy of an explanation of Moore's paradox that it account for the absurdity not only of Moorean assertions, but also of Moorean beliefs (see Heal 1994; Williams 1998 and de Almeida 2001).

Heal (1994) takes the extra step of requiring a *unified* explanation of the absurdity of Moorean assertions and beliefs. This is probably misguided. Moorean beliefs are not absurd in quite the same way Moorean assertions are. In fact, as Williams (1994; 1996) notes, they are *more* absurd than Moorean assertions.⁶ The added absurdity of Moorean beliefs can be explained by considering the proposition:

(3) It is raining and I am not asserting that it is raining.

Like (1), proposition (3) can be true, yet the *asserting* of it cannot. However, unlike the assertion of (1), assertion of (3) is directly *self-falsifying*, in that its tokening constitutes its own falsity-maker, if you will. (More accurately, a tokening of it would constitute one of its possible falsity-makers; the absence of rain would be another possible falsity-maker.⁷) By contrast, assertion of (1) is *not* self-falsifying in this way. (1) would be falsified either by the non-occurrence of rain or by the occurrence of a belief that it is raining, and neither is constituted by the asserting of (1).⁸ The absurdity of asserting (3) is thus more straightforward (hence more easily explained) than the absurdity of asserting (1).

The pathology involved in believing $\langle p$ and I do not believe that $p \rangle$ is logically the same as the pathology involved in asserting $\langle p$ and I do not assert that $p \rangle$. So the absurdity involved in *believing* (1) does *not* parallel the absurdity involved in *asserting* (1), but rather the absurdity involved in asserting (3). And since the absurdity involved in asserting (3) is different from the absurdity involved in asserting (1), the absurdity involved in *believing* (1) is different from the ab-

surdity involved in *asserting* (1). If so, these two absurdities call for different explanations. The argument, again, is this:

1. The absurdity of asserting (1) is unlike the absurdity of asserting (3);
2. The absurdity of asserting (3) is like the absurdity of believing (1); therefore,
3. The absurdity of asserting (1) is unlike the absurdity of believing (1).

And moreover, the absurdity of believing (1) is, like the absurdity of asserting (3), more straightforward than the absurdity of asserting (1). This argument suggests that, while it is correct to require an explanation of the absurdity of Moorean beliefs from a full account of Moore's paradox, it is mistaken to require that this be *the same* explanation as the explanation of the absurdity of Moorean assertions. The absurdity is not the same, so the explanation should not be the same. That is, it is misguided to require (as Heal does) a *unified* account of Moorean assertions and beliefs.

Moreover, the disanalogy between the absurdities of Moorean assertions and Moorean beliefs, and particularly the fact that the latter is more straightforward than the former, suggests a two-step strategy in approaching Moore's paradox. The first step is to explain the absurdity of Moorean assertions in terms of the absurdity of Moorean beliefs; the second is to explain independently the absurdity of Moorean beliefs. On this approach, Moore's paradox is, in its core, a paradox about belief. It is Moorean beliefs that are fundamentally and inherently paradoxical, and Moorean assertions simply inherit their paradoxicality (if you please) from the Moorean beliefs they are used to express. Let us call this approach *psychologism about Moore's paradox*.

Psychologistic explanations of Moore's paradox have recently been offered by a number of philosophers – e.g., Shoemaker (1988; 1995), Williams (1994; 1996 and 1998). But what is missing from their writings is a clear argument for it. In the remainder of this section, I outline such an argument. I argue that Moorean assertions are not universally paradoxical, but are so only when they express Moorean beliefs. More accurately, my claim is that *if* there are Moorean assertions that do not express corresponding beliefs, they are not absurd.⁹ That is, Moorean assertions are absurd only to the extent that they express Moorean beliefs. This would suggest (though not entail) that Moorean assertions are absurd *in virtue* of expressing Moorean beliefs.

Normally, our assertions express corresponding beliefs of ours. In asserting that such-and-such is the case, we *give voice* to our belief that such-and-such is the case.¹⁰ When one says that it is raining, one's assertion normally expresses a belief of one's to the effect that it is raining. But this is not always so: we can say things jokingly, or as actors in a play, or in some other way that is inexpressive of our beliefs. And it is often noted that Moorean assertions are problematic only to the extent that they are sincere, serious, etc. Thus, if one says jokingly "it is raining but I do not believe it is raining," one's assertion is nowise puzzling: it is just a joke, after all.^{11,12}

Inexpressive assertions of Moorean propositions (i.e., Moorean assertions that do not express Moorean beliefs) are thus unparadoxical.¹³ If the absurdity of Moorean assertions depends on their expressing Moorean beliefs, it is reasonable to conclude that Moorean assertions are absurd, when they are, *because* they express Moorean beliefs.¹⁴

It may be objected that inexpressive assertions are not genuine assertions. To say something jokingly, or insincerely, amounts perhaps to *pretending* to assert something, but not to actually asserting it (see Rosenthal 1995a, p. 323). On this view (which I find highly plausible), there is an *internal relation* between genuinely asserting something and expressing a corresponding belief.

Whether or not this is so, however, the expressiveness of utterances of Moorean sentences is still a necessary condition for their absurdity.¹⁵ Thus, if Moorean assertions necessarily express Moorean beliefs, then the question now becomes, Why are utterances of Moorean sentences absurd only when they constitute assertions? And the answer seems to be that the absurdity of utterances of Moorean sentences depends on their expressing Moorean beliefs, and only assertions express Moorean beliefs. So it is still the expressing of a Moorean belief that confers absurdity on linguistic utterances of Moorean sentences.

This line of thought suggests that if we can explain the absurdity of Moorean beliefs, we will have obtained a satisfactory explanation of the absurdity of Moorean assertions: namely, that they are absurd because they express absurd beliefs.

This line also brings into the picture another important element: consciousness. This is because only *conscious* beliefs can be expressed. Unconscious beliefs – tacit beliefs, repressed beliefs, etc. – can be *betrayed* or *reported*, but not *expressed*. It is impossible to give voice to a belief one is unaware of having.¹⁶ This fact – that only conscious beliefs can be expressed – is itself interesting and deserving of ex-

planation; Rosenthal (1990b) offers a promising one. But in any event it surely *is* a fact.¹⁷ So Moorean assertions are absurd only when they express *conscious* Moorean beliefs. To repeat, the reasoning is this: since Moorean assertions are absurd only when they express Moorean beliefs, and Moorean beliefs are expressed only when they are conscious, it follows that Moorean assertions are absurd only when they express conscious Moorean beliefs.

Given these two facts – that (i) only Moorean utterances expressive of Moorean beliefs are paradoxical, and (ii) only conscious Moorean beliefs can be expressed through Moorean utterances – it is plausible that Moore's paradox will succumb to an explanation in terms of the structure of conscious belief. Perhaps the structure of conscious belief is such that Moorean propositions cannot be consciously believed without explicit logical impropriety. Shoemaker's (1988, 1995) recent treatment of Moore's paradox considers this line of thought. In the next section, we turn to an examination of his argument.

3. SHOEMAKER ON MOORE'S PARADOX

Shoemaker starts his discussion of Moore's paradox with upholding psychologism (1995, p. 213; italics original):

What really needs to be explained is why someone cannot *believe* that it is raining and that she doesn't believe that it is, despite the fact that the conjuncts of the belief can both be true. If one can show that such beliefs are impossible, or at least logically defective, and if we come up with an explanation of this, then an explanation of why one cannot *assert* a Moore-paradoxical sentence will come along for free...

Shoemaker also thinks that a key to explaining why Moorean beliefs are logically defective is the following fact (Ibid.):

[C]onsider the proposition... expressed by the sentence "It is raining and I don't believe that it is raining, and that this is so (viz., that it is raining and I don't believe that it is) is something I believe." That is self-contradictory. So it is a feature of the contents of Moore paradoxical sentences that if they can be believed at all, the subject of such a belief could not believe that she had it without believing a contradiction.

That is, a second-order belief to the effect that one holds a Moorean belief involves contradiction. So it may be precisely this sort of contradiction that accounts for the logical impropriety of Moorean beliefs. This fact points to two possible explanations of the paradox, according to Shoemaker.

The first explanation appeals to the Higher-Order Thought theory of consciousness, as developed mainly by Rosenthal (1986; 1990a).¹⁸ According to the Higher-order thought (HOT) theory, a mental state *M* of a subject *x* is conscious when, and only when, *x* has a suitable thought *about M*. What makes a higher-order thought 'suitable' will not concern us here, except for the following restriction: the thought must represent not only the occurrence of *M*, but also that one is, oneself, *in M*. Thus it must represent not only *M*, but also oneself. Indeed, it must represent oneself *as oneself*. A higher-order thought representing that one's mother's nieceless brother's only nephew is in *M* would not ensure *M*'s being conscious (e.g., in case one is unaware that one is one's mother's nieceless brother's only nephew). So, on the assumption that the occurrence of a thought implies the occurrence of a belief,¹⁹ HOT theory is committed to something like the following principle:

- (A) If *x* believes that *p* *consciously*, then *x* believes that *p* and *x* believes that she herself believes that *p*.

Which we may formalize as follows:

$$(A^*) xB^c[p] \rightarrow xB[p] \& xB[IBp]$$

Where the contents of beliefs are in brackets and " $B^c[p]$ " stands for consciously believing that *p*. It is important to formulate the second conjunct of the consequent as $xB[IBp]$ (i.e., *x* believes <I believe that *p*>) and not as $xB[xBp]$ (i.e., *x* believes that *x* believes that *p*) because, as we just saw, it is crucial that the second-order thought represent oneself specifically *as oneself*.²⁰

Principle (A) is appealed to in the first explanation entertained by Shoemaker.²¹ This explanation is similar to the one advocated by Baldwin (1990: 230):²²

... a rational thinker will not consciously hold a Moorean belief. For to hold a belief consciously is both to hold the belief and be aware, and thus believe, that one holds it; and no rational thinker will believe... that he both believes and fails to believe the same thing (which is required by a conscious belief that *p* and that one does not believe that *p*)...

But unlike Baldwin's, the explanation under consideration does not restrict the scope of (A) to *rational* agents. Instead, it maintains that it is impossible to believe a Moorean proposition such as (1) without logical impropriety, because if (A) is true, then *x* consciously believing (1) entails that:

- (4) (a) x believes that it is raining and that she does not believe that it is raining and (b) x believes that she herself believes that it is raining and that she does not believe that it is raining.

And more generally, x consciously believing a Moorean proposition (i.e., $x\mathbf{B}^c[p \ \& \ \sim \ \mathbf{I}Bp]$) entails:

$$(5) \quad x\mathbf{B}[p \ \& \ \sim \ \mathbf{I}Bp] \ \& \ x\mathbf{B}[\mathbf{I}Bp \ \& \ \sim \ \mathbf{I}Bp]$$

Under the assumptions (i) that believing a conjunction entails believing the conjuncts (i.e., $x\mathbf{B}[p \ \& \ q] \rightarrow x\mathbf{B}[p] \ \& \ x\mathbf{B}[q]$) and (ii) that believing that one believes a conjunction entails believing that one believes the conjuncts (i.e., $x\mathbf{B}[\mathbf{I}Bp \ \& \ q] \rightarrow x\mathbf{B}[\mathbf{I}Bp \ \& \ \mathbf{I}Bq]$), (5) entails:

$$(6) \quad x\mathbf{B}[p] \ \& \ x\mathbf{B}[\sim \ \mathbf{I}Bp] \ \& \ x\mathbf{B}[\mathbf{I}Bp] \ \& \ x\mathbf{B}[\mathbf{I}B \sim \ \mathbf{I}Bp]$$

And (6) reports contradictory beliefs (in its second and third conjuncts).²³ A conscious Moorean belief thus implicates its subject in contradictory beliefs.

Note that the two assumptions appealed to here are highly plausible. The first (i.e., $x\mathbf{B}[p \ \& \ q] \rightarrow x\mathbf{B}[p] \ \& \ x\mathbf{B}[q]$) is commonly appealed to – explicitly or implicitly – in discussions of Moore's paradox (see, e.g., Williams 1994). It is hardly deniable, since any evidence for the proposition $\langle p \ \& \ q \rangle$ that may lead a person to believe that $p \ \& \ q$ is already evidence for $\langle p \rangle$ (and for $\langle q \rangle$) and would thus lead a person to believe that p (and that q).²⁴ The second assumption (i.e., $x\mathbf{B}[\mathbf{I}Bp \ \& \ q] \rightarrow x\mathbf{B}[\mathbf{I}Bp \ \& \ \mathbf{I}Bq]$) appears to me to be even more plausible: it is difficult to see what could possibly lead a person to believe that she believes $\langle p \ \& \ q \rangle$ without leading her to believe that she believes $\langle p \rangle$ (and that she believes $\langle q \rangle$).

As we saw above, against the background of these two assumptions and principle (A), consciously believing a Moorean proposition entails having contradictory beliefs. This result provides the proponent of the HOT theory of consciousness with what appears to be a promising account of Moore's paradox: conscious Moorean beliefs are absurd because having them necessarily involves having contradictory beliefs; and Moorean assertions are absurd because they express beliefs the having of which necessarily involves having contradictory beliefs.

In Section 4, I will argue that the HOT account of Moore's paradox is lacking. But it certainly appears to go in the right direction, especially if one hopes to explain the paradox in terms of a constitutive relation between first-order and second-order beliefs. Despite this apparently satisfactory result, Shoemaker prefers a different

explanation of Moore's paradox. According to this other explanation (1995: 214; italics mine):

... believing something *commits* one to believing that one believes it, in the sense that in some kind of circumstances, yet to be specified, if one believes something, and *considers* whether one does, one must, on pain of irrationality, believe that one believes it.

Shoemaker's explanation does not appeal to consciousness. Instead, the principle it appeals to is:

- (B) If x believes that p , then x is committed to the belief that she herself believes that p .

That is

$$(B^*)xB[p] \rightarrow xCB[IBp]$$

Where ' $xCB[IBp]$ ' is to be read " x is committed to believing the proposition $\langle I$ believe that $p \rangle$." Thus Shoemaker does not require that x believe that she believes that p , but only that she be *committed* to the belief that she believes that p . On the other hand, he imposes this requirement on *all* beliefs, not only on conscious ones.

Whether principle (B) can successfully explain Moorean absurdities depends, of course, on what is meant by "commitment to a belief." What does it mean for someone to be *committed* to believing a proposition, as opposed to simply believing it? In the passage just quoted, Shoemaker offers the following explication of "commitment to a belief": x is committed to a belief that p iff, if (in circumstances C , yet to be specified) x considers whether she believes that p , then she will come to believe that she believes that p .²⁵ To be committed to believing something, then, is to believe it if one considers it, where the consideration in question is of a doxastic kind. Against the background of this explication, (B) cashes out as:

- (C) If x believes that p , then if x doxastically considers (in circumstances C) whether she herself believes that p , then x believes that she herself believes that p .

Or more formally:

$$(C^*)xB[p] \rightarrow ((xD[IBp] \& C) \rightarrow xB[IBp])$$

Where ' $xD[IBp] \& C$ ' is to be read " x doxastically considers whether the proposition $\langle I$ believe that $p \rangle$ is true & circumstances C obtain."

Principle (C), however, does not appear to yield a satisfactory explanation of the absurdity of Moorean beliefs. If it is true, then x believing (1) entails that:

- (7) (a) x believes that it is raining and that she does not believe that it is raining, and (b) if x considers (in circumstances C) whether she believes that it is raining and that she does not believe that it is raining, then x believes that she believes that it is raining and that she does not believe that it is raining.

And more generally, x believing a Moorean proposition entails:

$$(8) \quad xB[p \ \& \ \sim IBp] \ \& \ ((xD[IBp \ \& \ \sim IBp] \ \& \ C) \rightarrow xB[IBp \ \& \ \sim IBp])$$

The question is what (8) implies in case x *does not* consider whether she believes <It is raining and I do not believe that it is raining> (or if she does but circumstances C fail to obtain). It is certainly not irrational to fail to consider whether one believes some proposition,²⁶ so as far as principle (C) is concerned, a perfectly rational agent could have a Moorean belief without considering whether she does. In that case, she will not *actually* have the belief that she believes the Moorean proposition, and therefore will not actually have the contradictory beliefs principle (A) imputes on her.

Yet a Moorean belief is absurd even when the agent does not consider whether she has it. Shoemaker must therefore claim that this absurdity is accounted for by the mere *possibility*, or *potentiality*, of the subject having contradictory beliefs – not by her actually *having* contradictory beliefs. On this explanation, a Moorean believer is guilty of irresponsibly putting herself in a position to easily fall into contradictory beliefs. For the only gap between having her current beliefs and having contradictory beliefs is the consideration of whether she has her current beliefs.

It is difficult to see, however, how the mere potentiality of having contradictory beliefs accounts for absurdity. And it is even more difficult to see what would make such an explanation preferable to one that exposes *actual* contradictory beliefs, which is what the explanation by appeal to (A) offers us. So, although I will later (in Section 4) argue that the explanation by appeal to principle (A) is not as satisfactory as it initially appears, that explanation nonetheless seems more promising than the explanation by appeal to principle (C), the one Shoemaker actually embraces.

It might be suggested that the explanation by appeal to (C) is preferable to the explanation by appeal to (A) in that (C) is not restricted to *conscious* Moorean beliefs as (A) is. But as we saw above, only conscious beliefs can be expressed in assertions, and therefore an explanation of the absurdity of conscious Moorean beliefs is sufficient

for the purposes of deriving the absurdity of Moorean assertion (in accordance with the psychologicistic program).²⁷

4. MOORE'S PARADOX AND CONSCIOUS BELIEF

Both accounts considered in the previous section appeal to a certain constitutive connection between first-order and second-order beliefs. There are two main differences between them. First, the HOT account appeals to an occurrent second-order belief whereas Shoemaker's account appeals only to a dispositional second-order belief. Second, the HOT account construes the second-order belief as numerically distinct from the first-order belief, whereas Shoemaker's account construes the second-order belief as somehow inherent in the first-order belief. It may be thought that Shoemaker can allow himself to construe the second-order belief as embedded in the first-order one precisely because it is only a dispositional belief. In this section, however, I explore an account of Moore's paradox we can obtain if we construe a conscious belief as one that is at once an occurrent first-order belief and an occurrent second-order belief. This results in the following principle:

- (D) If x believes that p *consciously*, then x believes that p and that she herself believes that p .

That is

$$(D^*) xB^c[p] \rightarrow xB[p \ \& \ IBp]$$

On this view, conscious beliefs always include an awareness of themselves. This is different, then, from principle (A), according to which every conscious belief is accompanied by a *separate* belief about it. (According to (D), $xB^c[p] \rightarrow xB[p \ \& \ IBp]$, whereas according to (A), $xB^c[p] \rightarrow xB[p] \ \& \ xB[IBp]$; the latter involves two separate belief states, whereas the former only one.) (D) yields, however, a solution to Moore's paradox that is quite similar to that yielded by (A). For if it is true, then x consciously believing (1) entails:

- (8) x believes (a) that it is raining and that she does not believe that it is raining and (b) that she herself believes that it is raining and that she does not believe that it is raining.

And more generally, x consciously believing a Moorean proposition entails:

$$(9) \quad x\mathbf{B}[p \ \& \ \sim \ \mathbf{IB}p \ \& \ \mathbf{IB}p \ \& \ \sim \ \mathbf{IB}p]$$

Which, on our assumption that believing that one believes a conjunction entails believing that one believes the conjuncts (i.e., $x\mathbf{B}[\mathbf{IB}p \ \& \ q] \rightarrow x\mathbf{B}[\mathbf{IB}p \ \& \ \mathbf{IB}q]$), entails:

$$(10) \quad x\mathbf{B}[p \ \& \ \sim \ \mathbf{IB}p \ \& \ \mathbf{IB}p \ \& \ \mathbf{IB}\sim \mathbf{IB}p]$$

The belief reported in (10) is *self-contradictory*, since the second and third conjuncts of its content (i.e., $\sim \mathbf{IB}p$ and $\mathbf{IB}p$) contradict each other.

This result provides the proponent of (D) with an explanation for Moore's paradox that is in some respects similar and in others dissimilar to the one proposed by the proponent of (A). According to the emerging explanation, conscious Moorean beliefs are absurd because they are self-contradictory, and Moorean assertions are absurd because they express self-contradictory beliefs.²⁸

This is different from the explanation provided by the proponent of (A), in that it imputes on the Moorean believer a single self-contradictory belief, rather than two distinct beliefs that contradict each other. This difference seems to constitute an advantage for the (D)-explanation over the (A)-explanation, since believing a self-contradictory proposition (i.e., $\mathbf{B}[p \ \& \ \sim p]$) is *worse* (more absurd) than believing contradictory propositions (i.e., $\mathbf{B}[p] \ \& \ \mathbf{B}[\sim p]$). In fact, as I will now argue, holding contradictory beliefs is not necessarily absurd, whereas holding a single self-contradictory belief is.

To see the gap between believing a self-contradiction and believing contradictories, consider the assumption needed to derive the former from the latter. This is the assumption that holding a conjunction of beliefs entails believing the conjunction (i.e., $x\mathbf{B}[p] \ \& \ x\mathbf{B}[q] \rightarrow x\mathbf{B}[p \ \& \ q]$). If this assumption was true, then believing contradictories would entail believing a self-contradiction. But despite the fact that its converse is highly plausible, and is appealed to by several of the accounts discussed above, the present assumption is false, as demonstrated by Williams (1981). First of all, if p and q are highly complex propositions, a child may be able to understand them separately but unable to wrap her mind around their conjunction, and consequently she will believe that p and that q but not that $p \ \& \ q$, since it is impossible to believe a proposition one does not understand. Second, x might come across independent evidence both in support of $\langle p \rangle$ and in support of $\langle \sim p \rangle$, and therefore believe both *separately*, but only a madman would believe them *conjointly*, that is, *as* $\langle p \ \& \ \sim p \rangle$.²⁹ Propositions $\langle p \rangle$ and $\langle \sim p \rangle$ cannot be both true,

of course, and so whatever evidence there is in favor of the falsehood among them, there must be greater *counter-evidence* against it. But x may be simply unaware of that counter-evidence.³⁰

Plausibly, believing a self-contradiction is psychologically impossible. But even if it is not, it is certainly acutely pathological. By contrast, believing contradictories is to some degree common, especially when one of the contradictories is believed *unconsciously*. Thus, upon being confronted by her therapist, x may consciously believe that her mother did not have an affair when x was a toddler, despite having a repressed belief to the effect that she did. The proposition consciously believed and the proposition repressively believed contradict each other, but there is nothing *absurd* – only infelicitous – about x believing both. Similarly, one may honestly come to believe, on the basis of solid evidence, that women are not inferior to men, while still harboring (say, due to early-childhood indoctrination) a deep-seated belief to the effect that they are. Such epistemic predicaments are not all that uncommon. And they show that one can readily believe contradictories when one of the beliefs is unconscious. Believing contradictories is a failing, to be sure, but there need not be anything *absurd* about it.

This last fact constitutes a crucial advantage of the explanation of Moorean absurdity by appeal to (D) over the explanation by appeal to (A). Recall that according to the (A)-explanation, the reason a conscious Moorean belief is absurd is that whenever x has it, both $\langle \sim IBp \rangle$ and $\langle IBp \rangle$ are believed by x . For consciously believing $\langle p \ \& \ \sim IBp \rangle$ requires believing $\langle IBp \ \& \ \sim IBp \rangle$. The contradiction here is between the second conjunct of the first-order belief (i.e., $\sim IBp$) and the believing of the first conjunct in the second-order belief (i.e., IBp). Now, according to the literature on HOT theories of consciousness, the second-order belief must be normally *unconscious* (see especially Rosenthal 1986 and 1990). For if it was normally conscious, the account of the consciousness of the first-order belief by appeal to a second-order belief about it would lead to infinite regress.³¹ But if the second-order belief is unconscious, then the contradiction imputed on the Moorean believer by the (A)-explanation is relatively untroubling, since it is a contradiction between a conscious content and an unconscious content: $\sim IBp$ is believed consciously, whereas IBp is believed unconsciously.³² The logical impropriety involved in Moorean beliefs is much more dramatic, however, than that involved in unconsciously believing something that contradicts what one consciously believes, which as we saw above, is not particularly absurd.

On the (D)-explanation, by contrast, both IBp and $\sim IBp$ are *consciously* believed. This is because these two propositions form part of the single self-contradictory content of a conscious Moorean belief.

It seems, then, that the mere contradiction between two beliefs may be insufficient to explain the absurdity of Moorean beliefs after all, but a single self-contradictory belief is sufficient. This is a crucial advantage of the (D)-explanation over the (A)-explanation. Given these significant advantages, the (D)-explanation appears to be preferable to the (A)-explanation.

It might be objected that the (D)-explanation of Moorean absurdity may be a little too strong. Conscious Moorean beliefs are absurd, all right, but not as absurd as self-contradictory beliefs. In any event, the absurdity of the former is surely *different from* the absurdity of the latter. Otherwise, there would be no difference between propositions of the logical form of (1) and propositions of the logical form of (2). Moorean propositions would just be self-contradictory propositions.

My rejoinder to this objection is twofold. First of all, it is crucial to note that even on the (D)-explanation of Moorean absurdity, there is a difference between propositions (1) and (2). The difference is that belief in (1) is self-contradictory only when conscious, whereas belief in (2) is self-contradictory whether conscious or unconscious. So the difference between the absurdity of Moorean beliefs and the absurdity of beliefs in self-contradictory propositions is preserved in the (D)-explanation. The objector might insist that, even so, it is false to compare the absurdity of a conscious Moorean belief to the absurdity of a self-contradictory belief. The latter is *more absurd* than the former, and therefore the absurdity of conscious Moorean beliefs must be construed as subtler or more complicated than it is in the (D)-explanation. Here I must bite the bullet, and concede that although it appears that the absurdities of conscious Moorean beliefs and self-contradictory beliefs are different in kind, it is a consequence of the (D)-explanation that in reality they are the same. To my mind, the intuition that the absurdity of conscious Moorean belief is different from that of self-contradictory belief is not as sharp and powerful as to make this conclusion unacceptable. Rather, this is one of those intuitions one would be well justified in discarding given the theoretical benefits of doing so.

Another, more fundamental objection is that principle (D) is itself completely implausible. Now, it is impossible to detail here the full case for (D); I have attempted to do so elsewhere (Kriegel (2003a; 2003b and 2005)).³⁴ But let us briefly note a couple

of considerations in favor of (D) – other than the fact that it appears to buy us a satisfactory explanation of Moorean absurdity!

It is widely accepted that a belief **B** of a subject *x* is conscious only if *x* is aware of **B**.³⁵ For *x* to be aware of **B**, *x* must be in a mental state **A** that *makes x* aware of **B**. We then face a dilemma concerning **A**'s relation to **B** (i.e., the awareness relation to the belief): either **A** is numerically the same as **B** (as suggested by (D)), or it is numerically different (as suggested by (A)). There are certain advantages to assuming that it is numerically the same. The main advantage is that if **A** is numerically different from **B**, then it is possible for **A** to misrepresent not only the properties of **B**, but also the very existence of **B**, whereas if **A** is the same as **B**, such that they are one and the same state, then while that state may misrepresent its own properties, it cannot misrepresent its own existence.³⁶ The possibility of **A** misrepresenting the existence of **B** is disconcerting, because it would mean that *x* is not in any conscious state, but is under the impression that she is – which sounds absurd.³⁷

Another advantage concerns the epistemology of consciousness. How do we know that conscious beliefs are beliefs we are aware of having? We know this by immediate acquaintance with the awareness in question. That is, we have first-person knowledge of being aware of our conscious beliefs. But as we noted above, according to the HOT theory of consciousness, and therefore (A), the awareness of our conscious beliefs is normally an unconscious state. We cannot have first-person knowledge of unconscious states, however, and so the HOT theory entails that we do not normally have first-person knowledge of our awareness of our conscious beliefs – in contradiction with the manifest facts. This problem does not apply to (D), because according to (D) the second-order belief is itself conscious, since it is one and the same as the first-order belief (which is conscious *ex hypothesi*).³⁸

Principle (D) is therefore highly plausible. It derives its plausibility from two facts: (i) the fact that conscious beliefs are beliefs we are aware of having; and (ii) the fact that there are serious disadvantages to construing the awareness of our conscious beliefs as a numerically different mental act or state. If we accept these two facts, then it seems that our conscious beliefs must constitute an awareness of themselves.

5. CONCLUSION: A NEW SOLUTION TO MOORE'S PARADOX

If conscious beliefs are partly about their own occurrence, then the reason conscious Moorean beliefs are absurd is that they are self-

contradictory beliefs. And if psychologism about Moore's paradox is correct, then Moorean assertions are absurd because they express self-contradictory beliefs. This is the solution to Moore's paradox that arises from the foregoing discussion.

To summarize the argument of this paper, we considered three possible accounts of, or solutions to, Moore's paradox. All three share a psychologistic approach according to which Moorean assertions inherit their absurdity from the Moorean beliefs they express. Furthermore, all three appeal crucially to second-order beliefs allegedly involved in certain kinds of first-order believing.

According to the first account we considered, the occurrence of a conscious belief that p entails the occurrence of a (normally unconscious) second-order belief that one believes that p ; and therefore the occurrence of a conscious Moorean belief entails the occurrence of contradictory beliefs (hence its absurdity). According to the second account, the occurrence (in the right circumstances) of a belief that p entails that if the subject considers whether she believes that p then she comes to believe that she herself believes that p ; and therefore the occurrence (in those circumstances) of a Moorean belief entails that if the subject considered whether she herself believes the Moorean proposition she would fall into contradiction (hence its absurdity). According to the third account, the occurrence (in any circumstances) of a conscious belief that p entails the occurrence of a belief that p and that one believes that p ; and therefore the occurrence of a conscious Moorean belief entails the occurrence of a conscious self-contradictory belief—hence its absurdity.

In Section 2, I argued for the psychologistic approach shared by all three accounts. In Section 3, I argued that the first account is superior to the second, because it exposes actual contradictory beliefs in a Moorean believer, rather than the mere possibility of contradictory beliefs. In Section 4, I argued that the third account is superior to the first, because it attributes to the Moorean believer a single self-contradictory belief, rather than two separate contradictory beliefs. The solution to Moore's paradox I am recommending here, then, appeals to the combination of psychologism about the paradox and an account of conscious beliefs as beliefs that are partly about their own occurrence.

One interesting advantage of all three solutions discussed in this paper over more common solutions in the literature on Moore's paradox is that they cast Moorean beliefs as *semantically*, rather than merely *pragmatically*, defective.³⁹ In particular, if the solution I suggest is correct, then conscious Moorean beliefs are downright self-

contradictory, meaning that there is something absurd about their *contents*, not only about the *acts* of believing those contents. This sort of solution is preferable to the more common pragmatic solutions, in that it accounts better for the fact that Moorean assertions and beliefs appear to involve a certain *logical* impropriety, rather than merely some sort of behavioral transgression. It is not only the subject's acts of asserting and believing that are absurd, but also, and primarily, *what* she asserts, or believes, that is absurd.⁴⁰

NOTES

- ¹ See Moore 1942: 543. The term "Moore's paradox" derives from Wittgenstein (1953). In this paper, I call "Moorean" only propositions of the form that bothered Moore. I thus do not follow the practice of some recent writers who expand the predicate "Moorean" to cover other, importantly similar propositions. Thus, Sorensen (1988) claims that many other propositions (e.g., "The atheism of my mother's nieceless brother's only nephew angers God") are Moorean; and de Almeida (2001) claims that "It is raining, but I am not justified in believing that it is raining" is Moorean. Some of these are indeed importantly similar to (1), but also significantly dissimilar. Whether one wants to call them Moorean is partly a verbal matter. In this paper, I call only propositions of (1)'s form Moorean. Indeed, except in endnote 28, I will not even consider propositions of the form $\langle p$ and I believe that not- $p \rangle$ (i.e., $p \ \& \ IB \sim p$), which have been commonly treated as Moorean but are nonetheless different in form from (1), as Williams (1979) pointed out.
- ² Some caution is needed here, however. On the most natural conception(s) of propositions, there are no I-propositions. There are only "I"-sentences, and perhaps also I-statements. The *proposition* expressed by different persons uttering "I am hungry," or "I am tall," is not the same proposition. To bypass this problem, we must therefore either (i) embrace an unusual conception of proposition, one that allows for I-propositions, or (ii) formulate the point not in terms of propositions, but instead in terms of sentences or statements, or rather in terms of sentence types or statement types. We can also do something that is neutral as between these two: we can stipulate that in the present paper, by "proposition" we will mean a sentence type or statement type; and remain silent on whether this *stipulated sense* captures the *common notion* of proposition. Quine (1960) and others have defined propositions as sets of sentences, so for Quineans "proposition" in our technical, stipulated sense means the same as "proposition" in the common sense, or perhaps in the only scientifically and/or ontologically acceptable sense. For others, it means something different. But in any case we can work with this sense of "proposition" for the purposes of the present paper without committing to any particular conception of propositions.
- ³ The paradox is differently formulated by different writers. For many, the problem is that Moorean propositions can never be *rationally* asserted, even though they involve no contradiction and may well be true. But this is a limiting construal of the paradox. There is something much more *dramatically* wrong about asserting Moorean propositions.

- ⁴ This may be for two reasons: either it is because there are assertions of Moorean propositions, but they are all necessarily false; or because there cannot be assertions of Moorean propositions. In the latter case, Moore's Paradox would be more precisely described as follows: Moorean propositions are paradoxical because although they can be true, they cannot be asserted (or they cannot be asserted to be true). In the present paper, however, I will work with the former characterization of Moore's paradox, according to which Moorean propositions are paradoxical because although they can be true, and can be asserted, they cannot be both *at the same time* – that is, they cannot be truly asserted. I work with the former characterization of the paradox because it appears, in my judgement, to be the less radical one. My main goal in this paper, however, is to suggest a solution, or a resolution, of the paradox. This requires having a working characterization of the paradox, but that is not my priority here. I do not think, in any case, that my “working characterization” of the paradox should affect the main thread of argument of this paper. The points I will make at each step of the discussion should be possible to transfer into the more radical characterization.
- ⁵ A point analogous to the one made in endnote 4 should be made here: I characterize the problem in terms of the fact that Moorean propositions cannot be truly believed even though they can be true, and this is compatible both with (i) the notion that Moorean propositions can be believed, but not truly, and with (ii) the notion that Moorean propositions cannot be believed at all. Although in this paper I work with (i), my sense is that similar claims to the ones I will make here can be made against the background of (ii).
- ⁶ Here, and in the remainder of this paper, I join Williams (1994, 1996) and others in assuming that it makes sense to talk of *degrees* of absurdity: that one assertion or belief can be more or less absurd than another. I do not pretend, however, to be in a position to offer an analysis of such talk. For that matter, I do not claim to be able to offer an analysis of the notion of absurdity. Julie Tenenbaum has suggested to me (personal communication) that absurdity may simply amount to something like *dramatic irrationality*. If so, one assertion or belief can be deemed “more absurd” than another when it is more dramatically irrational. But there is certainly a natural understanding of absurdity on which an assertion or belief is either absurd or not, but cannot be absurd to a degree. In this paper I will assume that this is not the case, however.
- ⁷ I would like to thank a referee for this journal for pointing this out to me.
- ⁸ At least the occurrence of the assertion is not *overtly* self-falsifying in the case of asserting (1) as it is in the case of asserting (3). One may argue that a genuine assertion necessarily involves the occurrence of a belief, and thus asserting (1) is self-falsifying after all, but such a claim still requires an *argument*, an argument which would bring out a *covert* sort of self-falsification.
- ⁹ As we shall see later, it is possible to hold that nothing counts as an assertion unless it expresses a belief. If so, I would have to reword the present claim as follows: “if there are Moorean utterances that do *not* express corresponding beliefs, they are not absurd.”
- ¹⁰ How to understand the expression relation between assertions and beliefs is something a full account of Moore's paradox would have to address. Williams (1994, 1999) explains it as follows: in asserting that *p*, one gives one's interlocutor *prima facie* justification to believe that one believes that *p*. In any case, the way I

am using the phrase, it is possible for a person to express a belief she does not really have – e.g., when lying.

- ¹¹ It might be claimed that the utterance itself *is* absurd, even if it does not serve to make any absurd assertion. This would be wrongheaded, however. For there is nothing in the concatenation of symbols constituting the sentence “It is raining and I do not believe that it is raining” to make it absurd in abstraction from the meaning it is used to express. Thus, “Blehp Bluhp Blihp” may mean that it is raining and it is not raining in a language yet to be invented, but that would not make it absurd. Sentences can be absurd only in virtue of the meaning they express.
- ¹² This point should be extended to all forms of “inexpressive” Moorean assertions. Thus, consider Moorean assertions by mindless robots, robots that do not *have* beliefs. Whether mindless robots *can* make assertions is problematic, but if they can, then when a mindless robot asserts a Moorean proposition, its assertion is not absurd in the least. In fact, it may well be true.
- ¹³ If so, the problem with Moorean propositions is not that even though they can be true they cannot be truly asserted, but, more specifically, that even though they can be true they cannot be *truly-and-expressively* asserted (where “expressively” is meant in the sense of “expressive of corresponding beliefs”).
- ¹⁴ It might be objected that “inexpressive assertions” are inexpressive of their propositional content, rather than of their corresponding beliefs. But this would appear to be implausible: an assertion that *p* cannot possibly fail to express the proposition that *p* – or it would not *be* an assertion that *p*.
- ¹⁵ I am using “utterance” here as a term for an act that does not imply the expression of a belief (or any other mental state).
- ¹⁶ I am assuming that conscious beliefs are beliefs one is aware of having (Rosenthal 1986, 1990a, Lycan 2001).
- ¹⁷ It might be objected that a person may believe that *p* unconsciously and say that *p* in an attempt to lie. But in this case, it does not appear that this person says that *p* *because* she believes that *p* in a way that would make the former an *expression* of the latter. When a person lies, she tries to *conceal* her real belief, not to *express* it. A similar objection would be that a person who believes that *p* unconsciously may express this belief in a Freudian slip of the tongue. But again, it seems that the correct thing to say here is that the person in question has *betrayed* her unconscious belief, rather than *expressed* it.
- ¹⁸ See also Mellor (1978), Carruthers (1989, 1996) and Dennett (1969, 1991).
- ¹⁹ I take it that this is a safe assumption to make. On a natural conception of the relation between thought and belief, a thought just is an occurrent belief, or a believing. But the present assumption maintains much less than that. It maintains that some sort of belief that *p* – occurrent or latent – is required for (involved in) the occurrence of a thought that *p*. Since a belief that *p* is constituted by whatever internal state the subject is assessable for truth or falsity in virtue of, and a thought is an internal state in virtue of which the subject is assessable for truth or falsity, we can safely say that the occurrence of a thought involves a belief. There is, of course, a tradition of using the term “thought” to denote *any* mental state, or any cognitive state (including entertaining that *p*, contemplating that *p*, etc.), but this is not how I use the term here. In this I follow Rosenthal’s own practice in his exposition of his Higher-Order Thought theory. By “thought,” Rosenthal means a mental state with an assertoric attitude

(or a mind-to-world direction of fit). Such mental states surely are, or involve, beliefs.

- ²⁰ The difference between $x\mathbf{B}[\mathbf{I}Bp]$ and $x\mathbf{B}[xBp]$ is that the former refers to the subject indexically, whereas the latter refers to her descriptively, or at least non-indexically. The difference between these two modes of self-reference is well explored in the literature on *de se* beliefs (see especially Castañeda 1966, 1969 and Perry 1979).
- ²¹ Note that this sort of account of the paradox is *not* offered by Rosenthal himself, who prefers an account in terms of the identity of performance conditions for assertions of “*p*” and of “I believe that *p*” (see Rosenthal 1995a, b).
- ²² Shoemaker’s own articulation of the idea is more convoluted. He writes (1995: 213–4): “One explanation would crucially involve the idea of consciousness. Assertion, on this view, requires that the speaker be conscious of the belief being expressed by the assertion. . . . Let’s assume with D. H. Mellor and David Rosenthal that a belief’s being conscious consists in the subject’s having a higher-order belief, or higher-order thought, to the effect that she has it. The content expressed by a Moore-paradoxical sentence could not be the content of a conscious belief – or, at any rate, could not be the content of a conscious belief without the subject’s having a self-contradictory belief.”
- ²³ By application of the first assumption, $x\mathbf{B}[p \ \& \ q] \rightarrow x\mathbf{B}[p] \ \& \ x\mathbf{B}[q]$, (5) – that is, $x\mathbf{B}[p \ \& \ \sim \mathbf{I}Bp] \ \& \ x\mathbf{B}[\mathbf{I}Bp \ \& \ \sim \mathbf{I}Bp]$ – entails $x\mathbf{B}[p] \ \& \ x\mathbf{B}[\sim \mathbf{I}Bp] \ \& \ x\mathbf{B}[\mathbf{I}Bp \ \& \ \sim \mathbf{I}Bp]$, which, by application of the second assumption, $x\mathbf{B}[\mathbf{I}Bp \ \& \ q] \rightarrow x\mathbf{B}[\mathbf{I}Bp \ \& \ \mathbf{I}Bq]$, entails $x\mathbf{B}[p] \ \& \ x\mathbf{B}[\sim \mathbf{I}Bp] \ \& \ x\mathbf{B}[\mathbf{I}Bp \ \& \ \mathbf{I}B\sim \mathbf{I}Bp]$. With a reapplication of the first assumption, this entails (6), $x\mathbf{B}[p] \ \& \ x\mathbf{B}[\sim \mathbf{I}Bp] \ \& \ x\mathbf{B}[\mathbf{I}Bp] \ \& \ x\mathbf{B}[\mathbf{I}B\sim \mathbf{I}Bp]$.
- ²⁴ Indeed, without getting too deep into tangential epistemological issues, it seems highly plausible that one can only reach a belief in a conjunction through believing the conjuncts.
- ²⁵ The circumstances Shoemaker has in mind are the circumstances in which a belief is “available” to the subject. Now, Shoemaker (1995: 227) admits that the notion of belief availability is far from clear, but we may illustrate this notion as follows. I may have a belief about the identity of Caesar’s murderer that is unavailable to me on a certain weekend: I cannot for the life of me recall it, no matter how hard I try and how confident I am that I *know* who murdered Caesar. On Monday, I could suddenly remember it was Brutus who did it. But during the weekend it was unavailable. According to Shoemaker, during the weekend I did not lose my belief, but rather the belief was unavailable to me. In restricting (C) to the right circumstances, Shoemaker means to exclude unavailable beliefs from the scope of the principle. In any case, this will not play a role in the discussion to follow.
- ²⁶ It is in fact impossible to consider, for every proposition, whether we believe it or not, since there are infinitely many propositions.
- ²⁷ Another advantage Shoemaker claims on behalf of the second explanation of Moore’s paradox over the first one is that the second one provides a better explanation of a related fact, namely, that the assent conditions of $\langle p \rangle$ and $\langle \mathbf{I} \text{ believe that } p \rangle$ are the same (1990: 214–5). He holds that principle (C) explains directly the fact that the assent conditions of $\langle p \rangle$ and $\langle \mathbf{I}Bp \rangle$ are identical, whereas principle (A) explains this fact only if it entails, or is supplemented by, (C). However, even *if* this is so, it surely represents a relatively minor advantage of an explanation that appeals to (C) and not (A) – an advantage which seems to me to be outweighed by the disadvantage discussed in the main text. Moreover, it is

questionable whether (C) truly explains the fact about identical assent conditions – see Larkin (1999) for an argument that it does not.

²⁸ Although I pledged in endnote 1 not to discuss propositions of the form $\langle p \ \& \ IB \sim p \rangle$, let me now say a little about how the present account would explain them. Conjoined with principle (D*), consciously believing such a proposition entails believing $\langle p \ \& \ IB \sim p \ \& \ IB_p \ \& \ IB_{\sim p} \rangle$. Against the background of our assumption that $x\mathbf{B}[IBp \ \& \ q] \rightarrow x\mathbf{B}[IBp \ \& \ IBq]$, this entails believing $\langle p \ \& \ IB \sim p \ \& \ IBp \ \& \ IB_{IB\sim p} \rangle$. The latter is not quite as bad as believing a self-contradiction, but it is very absurd in its own way. The believer here is guilty, not of believing contradictories, but of believing that she believes contradictories. As we shall see momentarily in the main text, believing contradictories is an epistemic failure, but a relatively undramatic one. By contrast, believing one believes contradiction is much more dramatic. When a person becomes aware of a contradiction in her belief system, she immediately revises one of her contradictory beliefs, or at least puts one of them on hold. Consequently, she immediately comes to believe that she *has been* believing contradictories, rather than that she *is* believing contradictories. But in the case of Moorean absurdity, we imagine a person who *stands by* her assertion. So this person is guilty of believing that she believes contradictories without taking back anything she asserts. And that is absurd.

²⁹ Here $\langle \sim p \rangle$ is used as a substitution instance of $\langle q \rangle$.

³⁰ This is impossible in the case of the converse assumption (i.e., $x\mathbf{B}[p \ \& \ q] \rightarrow x\mathbf{B}[p] \ \& \ x\mathbf{B}[q]$), because if the subject is aware of evidence against $\langle p \rangle$ (or $\langle q \rangle$), she will *also* – in fact, *thereby* – be aware of evidence against $\langle p \ \& \ q \rangle$.

³¹ If the belief that p is conscious in virtue of being represented by the belief that IBp and the belief that IBp is itself conscious, yet numerically different from the belief that p , then the belief that IBp must be conscious in virtue of being represented by a belief that $IBIBp$, which is itself conscious yet different from the belief that IBp , and then the belief that $IBIBp$ must be conscious in virtue of being represented by a further belief, and so on *ad infinitum*.

³² Strictly speaking, consciousness is not an attribute of contents, but rather of the beliefs *about* those contents. I use here the phrases “conscious content” and “unconscious content” as shorthand for (respectively) “content consciously believed” and “content unconsciously believed.”

³³ It is also preferable to an (E)-explanation, because (i) unlike (E), it explains the absurdity of *irrational* agents’ Moorean beliefs, and (ii) unlike (E), it does enjoy *prima facie* support, since conscious beliefs are precisely beliefs one is aware of having.

³⁴ It is worthwhile to note that the case has been made by others as well. The most thorough elucidation and defense of (D) is to be found in Brentano’s writings (mainly Brentano 1874). According to Brentano, it is a distinctive property of all conscious mental states that they include within them a “secondary judgment” about themselves. Caston (2002) argues, however, that Brentano inherited this view from Aristotle (see *De Anima* II, 3). It is also almost certain that Kant held something like this (see Brook 1994, especially Ch. 4, and Sturma 1995). More recently, it has been defended by Smith (1986, 1989), Thomasson (2000), Hossack (2002) and Caston (2002). It is impossible to review this literature here. I cite it mainly to suggest that an explanation of Moore’s paradox by appeal to principle

(D) will not go without backing from an important body of philosophical research that points to the truth of (D).

- ³⁵ See Rosenthal (1986) and Lycan (2001). But see also Dretske (1993) for an argument *against* this.
- ³⁶ This is because it is impossible for a mental state to represent anything (including itself) when it does not exist.
- ³⁷ For a discussion of this advantage of (D) over (A), see Caston (2002). For a compelling presentation of the problem (A) faces here, see Byrne (1997).
- ³⁸ For a discussion of this advantage of (D) over (A), see Kriegel (2003). A similar argument is offered by Goldman (1993).
- ³⁹ Another exception in the literature to this rule is Gillies' (2001) recent solution, which is purely semantic as well. But the rule in the literature has always been, under Wittgenstein's (1953) and Hintikka's (1962) influence, to offer a pragmatic solution.
- ⁴⁰ This paper was presented at Brown University, Georgia State University, McMaster University, University of Arizona, University of California at Santa Cruz, and University of Georgia. For discussion and comments on the paper, I would like to thank the audiences there. In particular, I would like to thank Yuri Balashov, David Chalmers, Juan Comesaña, Randy Cross, Derek Ettinger, Simon Feldman, Allan Hazlett, Julie Tenenbaum, and Joseph Tolliver. A special thank should go to John Williams, for a very helpful exchange of ideas on this topic.

REFERENCES

- Aristotle: 1907, *De Anima*, (trans. R. D. Hicks), Cambridge: Cambridge UP.
- Baldwin, T.: 1990, *G. E. Moore*, Routledge, London and New York.
- Brentano, F.: 1874, in O. Kraus (ed.), *Psychology from Empirical Standpoint*, ed. of English edition L. L. McAlister (trans. A. C. Rancurello, D. B. Terrell, and L. L. McAlister), London: Routledge and Kegan Paul, 1973.
- Brook, A.: 1994, *Kant and the Mind*, Cambridge UP, Cambridge and New York.
- Byrne, D.: 1997, 'Some Like It HOT: Consciousness and Higher Order Thoughts', *Philosophical Studies* **86**, 103–129.
- Carruthers, P.: 1989, 'Brute Experience', *Journal of Philosophy* **85**, 258–269.
- Carruthers, P.: 1996, *Language, Thought, and Consciousness*, Cambridge UP, Cambridge.
- Castañeda, H.-N.: 1966 'He': A Study in the Logic of Self-Consciousness', *Ratio* **8**, 130–157.
- Castañeda, H.-N.: 1969, 'On the Phenomeno-Logic of the I', reprinted in Q. Cassam (ed.), *Self-Knowledge*, Oxford UP, Oxford 1994.
- Caston, V.: 2002, 'Aristotle on Consciousness', *Mind* **111**, 751–815.
- de Almeida, C.: 2001, 'What Moore's Paradox is About', *Philosophy and Phenomenological Research* **62**, 33–58.
- Dennett, D. C.: 1969, *Consciousness and Content*, Routledge, London.
- Dennett, D. C.: 1991, *Consciousness Explained*, MIT Press/Bradford Books, Cambridge MA.
- Dretske, F. I.: 1993, 'Conscious Experience', *Mind* **102**, 263–283.
- Gillies, A. S.: 2001, 'A New Solution to Moore's Paradox', *Philosophical Studies* **105**, 237–250.

- Goldman, A.: 1993, 'Consciousness, Folk Psychology, and Cognitive Science', *Consciousness and Cognition* **2**, 364–383.
- Heal, J.: 1994, 'Moore's Paradox: A Wittgensteinian Approach', in *Mind* **103**, 5–24.
- Hintikka, J.: 1962, *Knowledge and Belief*, Cornell UP, New York.
- Hossack, K.: 2002, 'Self-Knowledge and Consciousness', *Proceedings of the Aristotelian Society* **102**, pp. 163–181.
- Kriegel, U.: 2003a, 'Consciousness as Intransitive Self-Consciousness: Two Views and an Argument', *Canadian Journal of Philosophy* **33**, 103–132.
- Kriegel, U.: 2003b, 'Consciousness, Higher-Order Content, and the Individuation of Vehicles', in *Synthese* **134**, pp. 477–504.
- Kriegel, U.: 'Naturalizing Subjective Character', forthcoming in *Philosophy and Phenomenological Research*.
- Larkin, W. S.: 1999, 'Shoemaker on Moore's Paradox and Self-Knowledge', *Philosophical Studies* **96**, 239–252.
- Lycan, W. G.: 2001, 'A Simple Argument for a Higher-Order Representation Theory of Consciousness', *Analysis* **61**, 3–4.
- Mellor, D. H.: 1978, 'Conscious Belief', *Proceedings of the Aristotelian Society* **78**, 87–101.
- Moore, G. E.: 1942, 'A Reply to My Critics', in P. A. Schlipp (ed.), *The Philosophy of G. E. Moore*, Tudor Publishing Company, New York.
- Perry, J.: 1979, 'The Problem of the Essential Indexical', *Nous* **13**, 3–21.
- Quine, W. V. O.: 1960, *Word and Object*, Harvard UP, Cambridge MA.
- Rosenthal, D. M.: 1986, 'Two Concept of Consciousness', *Philosophical Studies* **94**, 329–359.
- Rosenthal, D. M.: 1990, 'A Theory of Consciousness', ZiF Technical Report 40, Bielefeld, Germany, reprinted in N. J. Block, O. Flanagan, and G. Guzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*, MIT/Bradford, Cambridge MAS: 1997.
- Rosenthal, D. M.: 1995a, 'Moore's Paradox and Consciousness', *Philosophical Perspectives*, **9**: 313–333.
- Rosenthal, D. M.: 1995b, 'Self-Knowledge and Moore's Paradox', *Philosophical Studies* **77**, 195–209.
- Rosenthal, D. M.: 1990b, 'Why Are Verbally Expressed Thoughts Conscious?', Technical Report 32.
- Shoemaker, S.: 1988, 'On Knowing One's Own Mind', *Philosophical Perspectives* **2**, 183–209.
- Shoemaker, S.: 1995, 'Moore's Paradox and Self-Knowledge', *Philosophical Studies* **77**, 211–228.
- Smith, D. W.: 1986, 'The Structure of (Self-)Consciousness', *Topoi* **5**, 149–156.
- Smith, D. W.: 1989, *The Circle of Acquaintance*, Kluwer Academic Publishers, Dordrecht.
- Sturma, D.: 1995, 'Self-Consciousness and the Philosophy of Mind: A Kantian Reconsideration', in *Proceedings of the Eighth International Kant Congress*, vol. 1, Marquette UP, Milwaukee WI.
- Thomasson, A. L.: 2000, 'After Brentano: A One-Level Theory of Consciousness', *European Journal of Philosophy* **8**, 190–209.
- Williams, J. N.: 1979, 'Moore's Paradox – One or Two?', in *Analysis* **39**, 141–142.
- Williams, J. N.: 1981, 'Inconsistency and Contradiction', *Mind* **90**, 600–602.

- Williams, J. N. 1994, 'Moorean Absurdity and the Intentional 'Structure' of Assertion', *Analysis* **54**, 160–166.
- Williams, J. N.: 1996, 'Moorean Absurdities and the Nature of Assertion', in *Australian Journal of Philosophy* **74**, 135–149.
- Williams, J. N.: 1998, 'Wittgensteinian Accounts of Moorean Absurdity', *Philosophical Studies* **92**, 283–306.
- Wittgenstein, L.: 1953, *Philosophical Investigations* (trans. G. E. M. Anscombe), McMillan, New York.

Department of Philosophy
University of Arizona
Tucson, AZ 85721
USA
E-mail: theuriah@yahoo.com

Manuscript Submitted 24 August 2003
Final version received 23 March 2004