

JOHN N. WILLIAMS

WITTGENSTEINIAN ACCOUNTS OF MOOREAN ABSURDITY

(Received 24 May 1996)

1. INTRODUCTION

Over fifty years ago, G. E. Moore pointed out that to say,

(A) I went to the pictures last Tuesday but I don't believe that I did
(1942, p. 543)

or

(B) I believe that he has gone out. But he has not
(1944, p. 204)

would be “absurd” (1942, p. 543; 1944, p. 204).

Wittgenstein's letters to Moore show that he was intensely interested in this discovery of a class of possibly true yet absurd assertions.

Wittgenstein thought that the absurdity is important because it is “something *similar* to a contradiction, though it isn't one” (1974, p. 177). What is the explanation of the absurdity of saying or believing something about myself that might be true? Wittgenstein thought that although the explanation will say “something about the *logic* of assertion” it will also show that “logic isn't as simple as logicians think it is”. So although the explanation should

appeal to a contradiction, the contradiction is not to be unearthed by logic when narrowly conceived as the study of syntactic relations between propositions. Rather it will be unearthed by the application of a broader conception of logic to the act of assertion.

What is Wittgenstein's explanation of the absurdity? The popular answer among his followers (Linville and Ring 1991, Goldstein 1993 and recently Heal 1994) is that avowals of belief just are (or entail) assertions. I argue that if Wittgenstein held this "avowal-as-assertion" account then this is bad news for Wittgensteinians. The account faces serious difficulties, notably that it cannot explain the absurdity of assertions typified by (A). A similar failure afflicts the "performative" account of Moorean belief (that normally, one can't believe that one has beliefs without having them) which Heal attributes to Wittgenstein.

A satisfactory alternative account of Moorean assertion is that genuine assertions (and avowals of belief) entail *expressions* of belief. I recently sketched this account elsewhere (Williams 1994, 1996). Wittgenstein can plausibly be seen as anticipating this position in *Remarks on the Philosophy of Psychology*. A two-fold extension of this alternative account correlates with developments in his later philosophy. Cohering with this is an alternative account of Moorean belief. This is that Moorean propositions, if true, cannot be believed or cannot be believed without holding contradictory beliefs. This half of the account is consistent with Wittgenstein's transition from the early to the later philosophy.

2. CONDITIONS OF ANY ACCOUNT OF MOOREAN ABSURDITY

What conditions must be met by a successful account of Moorean absurdity? The first, to be defended shortly in §4, is that the account must identify a contradiction but without identifying it with the Moorean proposition itself. When confronted with Moorean utterances most people express a belief that a contradiction is involved. But aren't Moore's propositions possibly true? Since I am not omniscient, an (A)-type sentence might name a truth that I fail to believe. Since I am fallible, a (B)-type sentence might name a falsehood that I do believe.

Heal recognizes that a satisfactory solution “must be of adequate generality to explain the oddness of both *thought* and assertion” (1994, p. 6, my italics). But not just any thought. As Wittgenstein points out (1974, p. 177 and 1980b §280), there is no absurdity in *supposing* (or at the level of speech, verbally conjecturing) that Moore's propositions are true. Indeed Wittgenstein repeatedly stresses the paradoxical nature of this difference between assertion and supposition (1953, p. 190, 1980a, §§478 and 490). So thoughts of Moore's propositions are not absurd unless they constitute thinking that the propositions are true. Then the absurdity remains even if the proposition is not asserted. So, the absurdity of these *beliefs* needs to be explained as well. Ideally, this condition should cohere with the first. Explaining the connection between the occurrence of the contradiction-like phenomenon in assertion and its occurrence in belief would provide the fit between the two sorts of explanations.

But a *third* condition is equally important. Moore's examples typify two different *forms* of Moorean proposition (Williams 1979). Moore himself probably failed to see the difference. His account of one form fails to explain

the other. This fault is shared by many early explanations of the absurdity (Williams 1982b).

The difference hides under two layers of disguise in Moore's examples. The obvious initial representation of (B)¹ is

(b') I believe that p but not-p.

Linville and Ring stop here (1991, p. 295). Heal (1994, p. 6) and Gombay (1988, p. 192) go a step further. They achieve canonical reference to belief by commutation of (A), yielding

(a) I don't believe that p but p.

But a second layer of disguise remains. As Heal points out (1994, p. 7), my response, "I don't think so" is often intended to be taken non-literally as, "I think not". On the other hand, "where 'I do not believe that p' is taken as a self-ascription of ignorance as to whether p, then we do have a genuinely different structure" (1994, p. 7). This structural difference becomes visible if we now maximize the canonical form of Moore's examples by making the second conjunct of (b') identical to that of (a), yielding

(b) I believe that not-p but p.

The structural difference between Moore's propositions is now conspicuous as the difference between the internal and external negation of the belief-operator. This apparently slight difference turns out to be crucial.

The authenticity of the difference can be demonstrated. If believing that not-p entailed a lack of believe that p then contradictory beliefs would be impossible (unwelcome for reasons given in Williams 1982a and Sorensen 1988, p. 27). For if I believe that p and believe that not-p then (by applying the disputed principle to my second belief) I believe and don't believe that p. And the converse entailment prohibits agnosticism. Since if I neither believe that p nor believe that not-p then (by applying it to my first belief) I believe and don't believe that not-p.

So Heal is right to acknowledge (1994, p. 6) that “there are really two paradoxes”. As Gombay observes (1988, pp. 191-192), we can't give them the same label (whether Wittgenstein's label, “Moore's paradox” (1968 p. 190) or Linville and Ring's label, “MS” (1991, p. 295)) and then discuss the labeled phenomena (as Linville and Ring do) as if there were no difference. Admittedly, Heal does not explicitly claim that a satisfactory explanation of the one paradox must satisfactorily explain the other as well. But she does expose (1994, p. 7) a serious weakness in the suggestion that an avowal of belief that p presents itself as conclusive evidence that p. This suggestion explains the absurdity of asserting (b) but fails for (a). And she observes (1994, p. 11) that Baldwin's account of Moorean belief (that I cannot consciously believe (a) or (b) if I'm rational) suffers from the reverse lopsidedness. It explains the absurdity of believing (a) but fails for (b). These criticisms commit her to the third condition that any satisfactory explanation of Moorean assertion or belief must account for (a) *and* (b). Her own account however (1994, §§IV-VI) deals only with (b).

More importantly, the third condition is a methodological virtue. Given that rival explanations are on a par in other respects, the more general the explanation the better. But a single explanation of the four phenomena of

(a)-type assertion, (a)-type belief, (b)-type assertion and (b)-type belief, may be too much to hope for. An explanation of Moorean assertion may be unable to provide the explanation of Moorean belief. Indeed Moorean belief seems more absurd than Moorean assertion. The next best choice is therefore some pair of explanations which ideally should cohere together.

A fourth condition is suggested by Wittgenstein's observation (1980b, §290) that "under unusual circumstances [the] sentence ['It's raining but I don't believe it'] could be given a clear sense". Note here that what ceases to be absurd is the use or utterance of a *sentence* rather than an *assertion*. In (1980a, §§485-487) Wittgenstein gives two consecutive examples of non-absurd uses of (a)-type sentences followed by an example of a non-absurd use of a (b)-type sentence. An ideal account of Moorean absurdity must explain why these unusual circumstances remove the contradiction-like phenomenon that is usually present.

So in summary, any satisfactory account of Moorean absurdity must

- (i) identify a contradiction, or something contradiction-like, but not with the Moorean proposition itself
- (ii) make this identification for assertion and belief
- (iii) be equally plausible for (a) and (b)

and

- (iv) explain the role of circumstances in which absurdity disappears.

3. THE AVOWAL-AS-ASSERTION ACCOUNT

Wittgenstein credits Moore for having “said something about the *logic* of assertion” (1974, p. 177). What is it about *asserting*, as opposed to supposing or hypothesizing (a) or (b), that creates the absurdity? In (1980a, §478) Wittgenstein says:

Moore’s paradox may be expressed like *this*: “I believe p” says roughly the same as “+ p”; but “Suppose I believe that p...” does not say roughly the same as “Suppose p”.

Linville and Ring read this as the strong claim that an avowal of belief that p, “is a form of the assertion that p” (1991, p. 296), namely

(1b') My assertion that I believe that p just is an assertion that p.

The absurdity of asserting any proposition of the form of (b) can now be explained (Linville and Ring 1991, p. 296) as follows. If I assert that (I believe that not-p but p) then I assert that I believe that not-p, which is, by (1b'), just to say that *I assert that not-p*. But since I assert (b) *I assert that p*. So on (1b') the assertion of (b) “*consists of two contradictory assertions*” (Linville and Ring 1991, p. 296, my italics).

A weaker version of the avowal-as-assertion account avoids commitment to (1b') but accepts its entailment, namely

(1b) If I assert that I believe that p then I assert that p.

Heal seems to adopt this weaker position. In contrast with Linville and Ring, she does not explicitly identify the contradiction with the Moorean assertion itself but more cautiously claims that “p but not p” is what “[i]n effect has been said” (1994, pp.20-21, my italics). So on (1b) the assertion of (b) need not consist of two contradictory assertions but must at least entail them.

4. A DEFENCE OF THE FIRST CONDITION

Condition (i) prohibits identifying the contradiction with Moore’s propositions themselves. This prohibition is inconsistent with the strong avowal-as-assertion account. Because Linville and Ring accept (1b’) they must hold that my assertion of (b) is a self-contradiction and so cannot be true (1991, pp. 295-96). To insist that (b) is not self-contradictory because it might be true therefore begs the question. But there is good evidence that Wittgenstein endorses condition (i) and that therefore the strong avowal-as-assertion account cannot be consistently ascribed to him. Wittgenstein says that the absurdity is “something similar to a contradiction, though *it isn’t one*” (1974, p. 177, my italics).

Linville and Ring attempt to rebut Wittgenstein’s point that I can suppose that (a) or (b) are true without supposing a self-contradiction by dismissing appeals to imagination as tests of logical possibility (1991, p. 299). But although the dismissal is correct, it misses the point. The appeal to supposition is not an appeal to mental picturing, but merely to the fact that to assert or believe a subjunctive counterfactual condition such as

If I were to believe falsely that it's raining then I would be surprised to discover that the streets are wet

is in no way absurd. And as Sorensen points out, my belief that

If I eat this mushroom and it is, contrary to my belief, poisonous, then I will die

has the practical consequence that I am unlikely to eat the mushroom unless I have great confidence in my belief that it is not poisonous. But conditionals with self-contradictory antecedents could have no such practical interest. And as Sorensen adds, the negations of Moore's propositions are not tautologies (1988, p. 15).

5. DIFFICULTIES IN THE ACCOUNT

One point to note is that the account requires an additional conjunction principle for assertion, namely:

(Conj^{Ass}) If I assert that (p and q) then I assert that p and I assert that q.

Since this principle is highly plausible we must now look to (1b') and (1b) in search of any flaw.

First consider the stronger version. One reason against it has been given already. The possible truth of Moore's propositions is inconsistent with (1b'). Now inspect (1b') alone. To turn it around, must my assertion that p be an avowal of belief that p? Suppose that you arrest me on suspicion of

perpetrating Friday's robbery. Wanting to divert the blame to my accomplice I look you straight in the eye and say, "Charlie did it, I tell you". But Charlie and I have had a busy week and I have confused Friday's robbery, which really *was* perpetrated by Charlie, with Thursday's, which was perpetrated by myself. You are sufficiently experienced in deception to know that I am lying but you also have Charlie's fingerprints, so you know that I am insincerely telling the truth. The strong intuition is that you should judge that what I have said is true. Now suppose that the situation is exactly the same, except that now I say instead, "I think Charlie did it". Shouldn't you judge that I have said something false? Intuitions may now be ambiguous, which itself casts doubt upon (1b').

Linville and Ring are committed to judging my avowal in this second case to be true, in order to protect (1b'). This is why they endorse (1991, p. 305) Lakoff's claim that "in statements it is the propositional content, not the entire sentence, that will be true or false" (1975, p. 257). But now suppose that I bend under interrogation and make the false confession, "I did it". Shouldn't we judge this false, but judge the avowal "I think Charlie did it" (said as my memory clears) to be true? If (as I argue §7) the truth-conditions of avowals of belief that p are the sincerity-conditions of assertions that p, this is the answer we should give. Lakoff's general claim seems much less plausible for avowals of hope, suspicion and fear. Suppose I say "I hope you will get the promotion". You know that I am lying, because you know that I have just made an unsuccessful attempt to stop you getting it. Shouldn't you judge me to have said something false?

I now turn to (1b). For on either version of the avowal-as-assertion account, (1b) is what delivers the contradiction-like phenomenon. So if (1b) is false or fails to deliver the goods then the account fails.

There is an apparent counterexample to (1b). You ask me whether the post office is open and I make the reluctant reply,

“I wouldn’t like to *say* (don’t quote me), but I *think* so”.

Here I make an avowal of a belief that p which is qualified by a statement of reluctance to go as far as *claiming* that p. This statement serves as a disclaimer of responsibility. If you now waste a trip to the post office only to find it shut, you cannot reasonably blame me for having a false assertion, especially since I went to the trouble of clearly stating my reluctance to make one. Such disclaimers seem to prevent the avowal from being assertions of what is avowedly believed. If so then (1b) is false.

There is no absurdity in the reluctant reply. But if (1b) were true then it would be a way of remarking

“I wouldn’t like to *say* (don’t quote me), but it’s open”.

Arguably this exhibits a performative absurdity. Here I tell you that I am not about to make an assertion. But this is then contradicted by my making one. Three defences of the avowal-as-assertion account seem to surface in discussions of the counterexample.

The first claims that with the right intonation my statement, “... but I *think* that it’s open” should be taken as an assertion that I believe that it is more (perhaps only *just* more) likely than not that it’s open. So I cannot be blamed if the post office is shut because I did not make the false assertion that it was open, but only the (possibly true) assertion that it was likely to be so. So my reluctant reply should really be taken as

“I wouldn’t like to *say* that it’s open (don’t quote me), but it’s likely to be so”

which is not absurd.

But this commits us to saying that with the same intonation Moorean assertions must be taken as, “I don’t believe it’s likely, yet it’s true” and “I believe it unlikely yet it’s true”. Arguably the absurdity which we were supposed to explain has disappeared. Is it absurd to report that an outsider has won a horse race against the accepted odds? And if there is a pragmatic absurdity is it still of the original Moorean variety?

A second defence is to insist that I *did* make an assertion, albeit a self-confessedly reluctant one. My confession of reluctance disclaims responsibility for the consequences of my assertion (as avowal) that the post office is open. This is precisely why I cannot be blamed if the assertion turns out to be false. The price of this defence is that it entails rejecting an attractive view that Wittgenstein appears to adopt. In the later philosophy Wittgenstein frequently uses the verb “*behaupten*” for “assert”, one which he uses only twice (4.21, 6.2322) in the *Tractatus*. This term can be translated as “declaration of certainty”. On this view, my assertions are a bit like promises (cf. Linville and Ring 1991, p. 303) insofar as I give you some kind of guarantee that you can rely on them. This guarantee is explicitly annulled in the reluctant reply by my disclaimer of responsibility. So, on Wittgenstein’s view I could not have made an assertion.

One attraction of this view is that it is consistent with the intuition that claims to know, assertions, convictions, beliefs and avowals of belief lie on a diminishing scale of commitment to the truth. Thus to adapt Wittgenstein’s

example (1980b, §420), if A says “I believe it’s raining” and B says “I believe it’s not” the intuition that each contradicts the other is less substantial than if they had made flat contradictory assertions about rain. But (1b) renders this difference inexplicable.

The third defence of (1b) is to concede that it fails in exceptional counterexamples like the reluctant reply because it is normally, but not necessarily, true. Heal appears to make this concession in saying that “ a person learning the language is... trained to say ‘I believe that p’ *sometimes* as a substitute for the plain assertion ‘p’” (1994, p. 20, my italics). This non-universality of (1b) is consistent with Wittgenstein’s claim that the meanings of “I believe that p” and “p” are *roughly similar* but not identical. But now suppose that I add “but it’s not” to my reluctant reply thus expanding it to

“I wouldn’t like to *say* (don’t quote me), but I *think* that it’s open, but it’s not”.

Moorean absurdity has not appeared. But if (1b) is falsified by the reluctant reply then it cannot be the explanation of the absurdity. Against this it might be claimed that the absurdity here is not Moorean. This would be consistent with the defence’s commitment to holding that the disappearance of Moorean absurdity coincides with the falsehood of (1b). The expanded reply is a conjunction of three elements. These are a statement of reluctance to assert that p, an avowal of belief that p and an assertion that non-p. If the absurdity of the expanded reply is not Moorean then there must be some pairing of these three elements which is absurd in a way different from the (b)-type Moorean absurdity of pairing its last two. The original reluctant reply, which conjoins the statement of reluctance with the avowal, is not

absurd. So the supposed non-Moorean absurdity must lie in the conjunction of the statement of reluctance with the assertion that non-p, namely

“I wouldn’t like to *say* (don’t quote me) but it’s not open”.

But consistently with the reluctant reply this can be plausibly taken as

“I’m not going to say (I don’t want to say) that the post office is open, it’s not”

which is not absurd either.

6. THE FAILURE OF THE ACCOUNT

If these difficulties can be resolved then the absurdity of asserting (b) is satisfactorily explained by the avowal-as-assertion account as contradictory assertions which are entailed by the assertion of (b), although (b) is not itself a contradiction. So far so good. But how is the absurdity of a (a)-type assertions to be explained?

Obviously, (1b), which is a principle about asserting a belief, cannot explain the absurdity of asserting (a), which is not an assertion of believe, but rather an assertion of *lack* of belief. Goldstein recently (1993, pp. 94-95) attempts to remedy this by extending (1b) to

(G) If I assert that I don’t believe that p then I deny that p or express my refusal to accept that p.

Since denying that p is just asserting that not-p, this is ambiguous between

(1a) If I assert that I don't believe that p then I assert that not-p

and

(G') If I assert that I don't believe that p then I express my refusal to accept that p.

Forget (G') for the moment. (1a) explains the absurdity of (a)-type assertion thus. If I assert that (I don't believe that p but p), then by (Conj^{Ass}) I assert that I don't believe that p, and so, by (1a), *I assert that not-p*. But by (Conj^{Ass}) again, *I assert that p*.

But now the account demonstrably fails. If (1a) were true then an agnostic who asserted, "I neither believe that God exists nor believe that He doesn't" would be both asserting and denying the existence of God. Clearly he isn't so (1a) is false. As a principle about self-ascription of ignorance, (1a) is entirely implausible. My admission of ignorance of your innocence is not an accusation of guilt. Moreover, as Wittgenstein points out (1980b, §420), if A says "I believe it's raining" and B says "I don't believe so", they are not contradicting each other. Wittgenstein's point is especially salient if B's remark is a self-ascription of ignorance. But A and B would contradict each other if both (1b) and (1a) were true.

7. AN ALTERNATIVE WITTGENSTEINIAN ACCOUNT OF MOOREAN ASSERTION

Wittgenstein's clearest commitment to (1b) is in (1980a, §490):

The paradox is this: the *supposition* may be expressed as follows: "Suppose *this* went on inside me and *that* outside" – but the *assertion* [*Behauptung*] that *this* is going on *inside* me asserts [*behauptet*] this is going on outside me. As suppositions the two propositions about the inside and the outside are quite independent, but not as assertions.

The central idea here is that of independence between what I say about the inside (my beliefs) and what I say about the outside (the outside world). The formulation of the dependence is crucial in explaining Moorean absurdity. So Heal is quite right to point out that a serious defect of functionalism is that it makes belief into an inner state which is non-casually "independent of how things are outside" (1994, p. 18). Once the independence is severed there can be no case in which (a) or (b) are absurdly asserted.

Wittgenstein can be read as endorsing two such dependencies which in turn satisfactorily explain Moorean assertion. An integral part of both dependencies is the idea of sincerity, or believing what one says. In (1980a, §472) Wittgenstein says:

I want to say first of all with the assertion "it's going to rain" one expresses [*drückt aus*] belief in that just as one expresses the wish to have wine with the words "Wine over here!" One might also put it like this: "I believe p" means roughly the same as "p".²

So the inside depends upon the outside in that:

(2) If I assert that p then I express a belief that p.

In other words, my avowal of belief that p is an *Äusserung* not in the sense of a statement or assertion (*Behauptung*) that p but in the sense of an expression or voicing of the belief (*einen Glauben ausdrücken*). How is the belief expressed? In (1953, p. 190), Wittgenstein explains the rough equation of meaning between “I believe that p” and “p” as the fact that they are used in similar ways. In other words, as Heal point out (1994, p. 20), a person is taught to make avowals of belief that p in the same way that he is taught to make assertions that p. So unless I’m self-deceived, I don’t need to look to my behaviour (see Wittgenstein 1980a, §501) nor inside myself to Cartesian or Tractarian mental objects (see Wittgenstein 1980a §488) in order to make a sincere avowal of belief that p. All I need to do is to look at the world and ask myself whether p. In (1980a, §477) Wittgenstein in turn identifies the rough equation of meaning or use with the fact that “We react in roughly the same way when anyone says the first and when he says the second”. How we react depends on what we are looking for. A priest looking for converts reacts to a presumedly sincere assertion of the existence of God in the same way as he would react to a presumedly true avowal of belief in God. The existence of God is an issue which the priest regards as settled. What he hopes to discover is the belief. By contrast, your reaction to a presumedly true avowal of belief that the trains are still running might be different from your reaction to a presumedly true assertion that they are. What interest you are not my beliefs but catching the train. But sincerity is a far less reliable indicator of truth-telling than truth-telling is of sincerity.

But whether we are looking for the belief or the truth, we would *not* react in the same way if we knew that the avowal of belief was made truly (or falsely) on one occasion but that the assertion that p was made *insincerely* (or sincerely) on another. The priest would react differently. And if we are interested in whether the trains are running then knowing that the speaker is insincere should make a difference to our confidence that they are running.

The connection of meaning then between “I believe that p” and “p” is that my *practice* of truly using the first is normally the same as my practice of *sincerely* using the second. There can be abnormal cases in which I insincerely (untruthfully) tell the truth. For example, I might pretend to let you know something which in fact is just a lucky guess. Or my attempt to deceive you with a lie may be defeated by the fact that my belief in the falsehood of my assertion is mistaken. But unless you recognize such rare cases for what they are, your justification for thinking me a truth-teller includes the justification for thinking me sincere.

So in asserting that p, I normally *express* a belief that p in the sense that I afford you the *prima facie* justification for thinking me sincere, by affording you the *prima facie* justification for thinking me a truth-teller.

What justifies this ascription of truth? Perhaps not Moore’s claim (1942, pp. 542-543) that lying is *vastly* exceptional, nor my strengthened version of it (Williams 1994, p. 164) that insincerity is *necessarily* exceptional. Both claims invite challenge by skeptical brains-in-vats or evil demon scenarios³. Nonetheless, the rational thing to do is to take appearances at face-value, unless keen observation indicates otherwise. So what entitles you to take me as a truth-teller and as sincere is the experimentally undefeated presumption that falsehood and insincerity *are*

exceptional. When asking for directions in a strange town, it is better bet to err on the side of gullibility than paranoia, even if (for once) the paranoid prejudice is correct and all the natives *are* out to get you.

The absurdity of asserting either of Moore's propositions can now be explained: If I assert either then by (Conj^{Ass}) I assert that p. So by (2) *I express a belief that p*. But by (Conj^{Ass}) again, *I assert a lack of belief that p* in (a) or *I assert a belief that not-p* in (b). So what I conjointly assert and express is a contradiction in (a) and contradictory beliefs in (b). As one might expect, the different conceptual structures of (a) and (b) result in a difference in the contradiction-like phenomena.

The only plausible reading of (G') is one which collapses into (2). (G') explains the absurdity of asserting (a) as follows. If I assert that (I don't believe that p but p) then I assert that I don't believe that p. So by (G') I express my refusal to accept that p. Goldstein comments that such a speaker "is, or purports to be, simultaneously putting a proposition forwards and taking it back" (1993, p. 95). But how is the metaphor of putting and taking to be read? Surely "accept" is used here as a synonym of "believe". So in asserting a proposition, I put it forward in the sense that I purport to believe it. On reading or "purport" is "pretend". But while this fits lies, it doesn't fit other types of assertion. A liar pretends to believe or know his assertion in order to establish the authority needed to make his dupe believe it too. But if I sincerely let you know something then I don't *pretend* to believe it because I *do* believe it. The only plausible reading of "purport" is the non-factive sense of "express", which yields (2).

8. THE FIRST EXTENSION OF THE ALTERNATIVE ACCOUNT

The alternative account locates the contradiction-like phenomena on a hybrid level of assertion-plus-expression. But wouldn't it be better to locate it on a single-level? The answer to this is that the account can be extended to locate the contradiction-like phenomenon purely at the level of expression of belief as well. This extended account is again a consequence of the key idea that your justification for thinking me a truth-teller includes the justification for thinking me sincere.

Wittgenstein's view of the outside is that the dependence is two-way. The outside depends upon the inside as well. As Wittgenstein remarks in (1953, p. 191) my assertion "It is so", "throws light" on my state of mind in the same way as, "I believe it is so". So if you are justified in thinking that I tell the truth when I avow a belief that p then you are justified in thinking that I really do believe that p. This does not commit you to thinking that I have asserted that p but only to thinking that any assertion I might make would be sincere. This can be formulated as

(3b) If I assert that I believe that p then I express a belief that p.

Similarly, if you are justified in thinking that I tell the truth when I avow a *lack* of belief that p then you are justified in thinking that I really don't believe that p. This can be formulated as

(3a) If I assert that I don't believe that p then I express a lack of believe that p.

In (1974, p. 177) Wittgenstein comments:

If someone asks me “Is there a fire in the next room?” and he answers “I believe there is” I can’t say: “Don’t be irrelevant. I asked you about the fire, not about your state of mind!”

The avowal of belief is relevant because in making it, I say something about my inner self (in the double sense of avowing and therefore expressing my belief) *and* say something about the outside world in the sense that I say that I have a view of how the world is (as including a fire in the next room) although I might not be prepared to *claim* that this is how it is.

This shows that Linville and Ring’s argument for the avowal-as-assertion account is unsound. In effect the argument is this: an avowal of belief is either about the avower or is about the outside world and since it is not about the avower it is therefore an assertion about the outside world (1991, p. 296). Because the avowal is about both, the second premise is false.

Goldstein’s (G’) commits him to (3a) as well as (2). (G’) says that in asserting that I don’t believe that *p*. I express my refusal to accept (believe) that *p* in the sense that I purport to take it back. For the same reasons given in the last section, purporting to not believe it can only be plausibly taken as expressing the lack of belief in it.

The absurdity of asserting either of Moore’s propositions can now be explained as follows. If I assert either then by (Conj^{Ass}) I assert that *p* and so by (2) *I express a belief that p*. But by (Conj^{Ass}) again, I assert that I don’t believe that *p* in (a) or I assert that I believe that not-*p* in (b). So by (3a) *I express a lack of belief that p* or by (3b) *I express a belief that not-p*. The absurdities are thus explained yet distinguished as expressing a contradiction

in (a) as opposed to expressing contradictory beliefs in (b). The connection between Moorean belief is also simply explained as the fact that a Moorean assertor expresses Moorean beliefs.

9. THE SECOND EXTENSION OF THE ALTERNATIVE ACCOUNT

In §7 we saw that by making an assertion I afford a hearer the *prima facie* justification to think that I am *telling the truth* and the *prima facie* justification to think that I am *sincere*. Combining these two justifications constitutes the justification to believe *me*. Both justifications are needed. If you know that I am sincere but mistaken in what I assert then you won't believe me. And if you know that I'm making a lucky guess, inadvertently telling the truth in failed attempt to deceive you or merely parroting information again you won't believe me, although you will believe what I say. To believe me is to believe that I am sincerely (truthfully) telling the truth. So when I assert something, I give you *prima facie* justification to believe me. If the assertion is Moorean then the justification is worthless because in fact it is a justification for *not* believing me in the case of (a) or for not believing me on the charitable assumption that I am minimally rational in the case of (b).

To see this, suppose that you *do* believe me. Since you think me sincere in asserting the second conjunct, *you believe that I believe that p*. And since you believe what I say in the first conjunct, *you believe that I don't believe that p* in (a) and *you believe that I believe that not-p* in (b). If you are to believe me, you must have contradictory beliefs in the first case (conclusive grounds for disbelieving me, which will therefore force any

rational person to withhold belief) and you must think *I* have them in the second (conclusive grounds for believing me irrational). So you cannot believe me unless you are irrational or judge me to be so.

The spirit of this account is arguably late-Wittgensteinian, in being a treatment of the whole person rather than just of relations between propositions. The treatment clarifies what goes right in communication by diagnosing what goes wrong. Central to the diagnosis is the idea of expression which is in turn located in a social context. Cohering with this is the provision of a psychological explanation of our puzzlement when confronted with real-life Moorean speakers. When I say (a) or (b), I express doxastic contradiction in me or you. Detecting this, you look inside yourself and fail to find it there and by the charitable principle of avoiding ascriptions of irrationality (see Sorenson 1988, p. 132) you hesitate to locate it in me. So you look at what I've asserted, but you can't detect the contradiction there either.

10. THE "PERFORMATIVE" ACCOUNT OF MOOREAN BELIEF

As might be expected from her treatment of Moorean assertion, Heal concentrates exclusively on (b). Avowals of belief don't constitute believing in the way that marriage vows normally constitute marrying. Nonetheless Heal insists that second-order beliefs have "a sort of performative character" (2004, p. 21) in that "when I come to thought that I believe that p then I do, in virtue of that very thought, believe that p" (1994, p. 22), i.e.:

(4b) If I believe that I believe that p then I believe that p.

The absurdity of believing any proposition of the form of (b) can now be explained as follows. If I believe that (I believe that not-p but p) then I believe that not-p. So by (4b) *I believe that not-p*. But since I believe (b), *I believe that p*.

11. THE FAILURE OF THE PERFORMATIVE ACCOUNTS

Just as the avowal-as-assertion account requires the principle that assertion distributes over conjunction, so the performative account requires its parallel for belief:

(Conj^{Bel}) If I believe that (p and q) then I believe that p and I believe that q.

Again this seems undeniable. But the absurdity of believing (a), which is about a *lack* of belief, cannot be explained by Heal's (4b). If I believe that (I don't believe that p but p) then by (Conj^{Bel}) I believe that I don't believe that p but I do believe that p. But (4b) cannot yield contradictory beliefs from this. We could try maximizing coherence with the original explanation by appealing to the analogue of (4b):

(4a) If I believe that I don't believe that p then I don't believe that p.

But now the possibility of absurd (a)-type beliefs refuses this parallel principle. If I believe that (I don't believe that p but p) then by (Conj^{Bel}) I

believe that I don't believe that p and *I believe that p*. But by (4a) *I don't believe that p*.

Flat contradiction is likewise entailed by applying both (4b) and (4a) to a more sophisticated form of (b), in which "I believe that q" is substituted for "p", yielding

(b*) I believe that I don't believe that q but I believe that q.

If I believe (b*) then by (Conj^{Bel}), I believe that I believe that I don't believe that q and I believe that q. Applying (4b) to strip off the first layer of belief-operators entails that I believe that I don't believe that q and *I believe that q*. But applying (4a) to this first belief entails that *I don't believe that q*.

We must therefore either deny the possibility of believing (a) or (b*) or reject one of the two principles which entail the flat contradiction. On the first option the explanation is not that Moorean belief entails contradictory beliefs but rather Hintikka's claim (1962, p. 67) that Moorean belief is impossible. But the performative account explains the absurdity of (b)-type beliefs *once held* in terms of contradictory beliefs (which are possible). So since (Conj^{Bel}) is unassailable, (4a) is false whenever it is absurd to believe (a) and either (4a) or (4b) is false whenever it is absurd to believe (b*).

Admittedly, Heal could reject (4a) but retain (4b) without formal contradiction. They are logically independent. But they do seem to stand or fall together, depending upon the corrigibility of judgments about my own mental states. If my conviction that I believe that God does not exist is performative then why isn't my conviction that I don't believe that God exists?

One temptation to accept (4b) might be an assimilation of “belief” with “thought”. I can’t think about a thought of eels without thinking about eels. But it doesn’t follow that I can’t have a belief about a belief without having that belief. For although believing is a kind of thinking, most of us have long believed that eels don’t eat glass without having thoughts of glass-eating eels.

Heal recognizes that (4b) might be occasionally defeated by tiredness, stress or drugs (Heal 1994, p. 22). A further possible falsification of both principles is self-deception. My sincere professions or open-mindedness about the status of women or of lack of prejudice against them, may both be mistaken. In such exceptional cases my observation of my outward behaviour might help me to discover what my inner beliefs really are and so remove the self-deception.

But like Moorean assertion, Moorean belief is *always* absurd. Therefore the circumstances in which the performative principles are false are those in which the absurdity of Moorean belief persists. Hence the principles cannot be the explanation of the persistent absurdity. For example, assume that I’m convinced that I have a *false belief* in the equality of women, as a case of (b), or that I’m convinced that I *lack the correct belief* that women are unequal, as a case of (a). Both convictions are certainly absurd. Now suppose that both are the product of self-deception so that both are mistaken and (4a) and (4b) are both false. Far from expunging the absurdity of my convictions, this supposition seems to heighten it.

To advert to Heal’s consideration, the fact that I have been deprived of sleep for a week, have been arrested for murder and am high on LSD, might be the cause of my absurd Moorean conviction (why I’m mad). But it

might also be the cause the failure of (4a) and (4b) in which case these principles won't explain the nature of the absurdity (the madness) itself.

12. AN ALTERNATIVE ACCOUNT OF MOOREAN BELIEF

The account I propose appeals to the fact that Moore's propositions are "blindspots". Here I use Sorensen's term but not for his reasons. Sorensen argues that (a) and (b) are belief-blindspots for me in that I cannot believe them. Here "cannot" is to be specified by background constraints which can range from

Merely the rules of logic,... to a more substantial set of constraints such as the laws of physics, the principles of psychology, or immunity from structural criticisms (1988, pp. 52-53).

Sorensen is methodologically committed to avoiding logics of belief (1988, pp. 19-22). This is why he restricts the use of doxastic principles to a hypothetical test of structural flaws in beliefs, the results of which are to be compared to their constraining *desiderata* (1988, p. 43).

My account of the blindspots is simpler. Take (a). If I believe that (I don't believe that p but p) then by (Conj^{Bel}) I believe that p. But then (a) is false because its first conjunct is false. It is therefore a mistake to think that because the content of a belief could be true, it would be possible to "believe-it-correctly". For (a) is a possible truth and I can believe it but it cannot be true *if* I believe it. By contrast, I *can* correctly believe (b). If I

believe that (I believe that not- p but p) then by (Conj^{Bel}) I believe that p . Nonetheless (b) might be true if I hold contradictory beliefs.

Alternatively put, my *believing* either (a) or (b) entails (by applying (Conj^{Bel}) to its second conjunct) that *I believe that p* . But the *correctness* of what I believe entails that *I don't believe that p* in (a) or that *I believe that not- p* in (b). So my believing-correctly entails a contradiction in the case of (a) and contradictory beliefs in the case of (b). Yet (a) and (b) themselves might be true of me and might be believed by me.

This account of the blindspots is more economical than Sorensen's. I have two constraints on the believability of Moore's propositions. In the case of (a) it is truth, the weakest one of all. In the case of (b) it is minimal rationality or the absence of contradictory beliefs. This economy is enabled by my minimal commitment to (Conj^{Bel}) as opposed to a full-blown logic of belief, of which I am equally suspicious. I see no methodological objection to using a plausibly true principle which is about what it explains.

Conditions (i) and (ii) have now been met for belief, as they were for assertion, and so (ii) is met as well. We also have the desirable fit between the explanations of Moorean belief and of Moorean assertion. The contradiction-like phenomena entailed by Moorean correctly-believing is identical to that expressed by Moorean assertion or ascribed by those who believe Moorean assertors.

We can also explain why Moorean assertion seems less absurd than Moorean belief. Moorean assertors express Moorean beliefs which are conclusive signs of irrationality. But Moorean assertors can express Moorean beliefs which they don't have. It is a lesser sin to licence the criticism of irrationality than to be guilty of it.

Wittgenstein never explicitly mentions Moorean *belief*. What criteria should we adopt for ascribing an account of it to him? The charitable answer is that we should prefer those accounts which are maximally satisfactory and which maximally cohere with the rest of Wittgenstein's thought. Obviously Heal appeals to the first criterion. And she appeals to the second in claiming that (4b) parallels Wittgenstein's views on first-person pain statements (1994, p. 20, fn.9). My own account of Moorean belief satisfies this second criterion at least as well because it coheres with Wittgenstein's views in the *Tractatus* (1922, 5.63 ff.).

Just as (a) could not report a truth about an omniscient being so (b) could not report a truth about an infallible one. But for human beings, who cannot contemplate the world *sub specie aeterni*, Moorean belief involves an interesting modal shift. Of course I can have non-Moorean beliefs that I am non-omniscient or fallible. These are beliefs that there is some substitution-instance of "p" for which (a) or (b) are true for me. They are rational beliefs precisely because there very probably *are* and thus *can* be such instances of ignorance or error. But if (a) reports a specific instance of such ignorance then I *cannot* believe it correctly. Given my ignorance, I cannot believe correctly that this instance of ignorance occurs. Similarly, I am necessarily ignorant of my specific errors reported by (b)-type propositions if I am minimally rational. The reason that there are truths about my specific ignorance (or errors) which I cannot believe is that as an *epistemological* subject they cannot appear in my conception (or rational conception) of the world (cf. Baldwin, 1990, p. 231). Truth (or rationality) constrains us from seeing our own individual failures of vision. This account does not commit us to a view of the self as a metaphysical limit nor to solipsism, both of which might be seen as unwelcome Tractarian consequences. So

Wittgenstein might have consistently carried the account over into the later philosophy.

The account also provides the missing explanation of Gombay's correct claim that Moore's propositions report states which are "counterprivate" (1988, p. 193). Anyone except me can (sensibly) think or say that I am in these specific states. This is starkly opposed to the performative account which holds that my vision of my beliefs (unlike that of others) is normally perfect. But in fact I am blind to the relation of my specific beliefs to the world in a way that no-one else is. Given the anti-private tenor of Wittgenstein's later philosophy, the alternative account coheres with it better.

13. THE DISAPPEARANCE OF ABSURDITY

We can now explain the role of circumstances in which the absurdity of saying (a) or (b) disappears, thus meeting condition (iv).

Wittgenstein's first example (1980a, §485) of the disappearance is my exclamation of amazement. "He's coming but I still can't believe it". This is explained by Baldwin's observation that we naturally interpret the second part of the remark as saying, "I find it very difficult to believe it" (1990, p. 229). By the principle of charity, we seek ways of believing the speaker without making self-ascriptions of contradictory beliefs. By the same principle, we can believe that a tabloid reporter has genuine information which is "incredible" only by making this claim as presenting a fact which resists but merits belief.

Wittgenstein's second example is a railway announcer who is convinced (perhaps groundlessly) that the train won't arrive. He announces its arrival and adds, "Personally I don't believe it" (1980a, §§486-487). This corresponds to (a). The third example is a soldier who produces military communiqués but adds that he believes they are incorrect. This corresponds to (b). Both remarks seem absurd unless we suppose that the speaker is merely parroting the assertion of an authority. Then the absurdity vanishes along with our justification to think that the speaker believes what he utters. The parroting is not a genuine assertion but the report of one.

A similarly spurious assertion occurs if I mention someone else's assertion in order to assert its negation. No absurdity arises if I sarcastically intone your claim that the earth is flat and add "I *don't* think!". Obviously I afford you no reason to think I believe the claim I'm denying. The absurdity of saying "it's true but I don't believe it" is similarly expunged by the recognition that the speaker is making a wild guess in a "true-or-false" quiz in which success is merely the right answer (as opposed to the demonstration of knowledge).

In each case, the absurdity is removed by removing the justification to think that the speaker believes what he says. If the exclamation, tabloid exaggeration, announcement, communiqué, denial or guess were genuine assertions then they would be absurd. But none of them are.

Notes

¹ Heal actually uses a different example from either of Moore's namely, "I believe that it is raining but it is not" (1994, p. 5)

² Wittgenstein's view here startlingly resembles Searle's "(1983, p.9):

...in the performance of each illusionary act with a propositional content, we express a certain intentional state with that propositional content, and that intentional state is the sincerity condition of that type of speech act.

Searle goes on to point out that one can express intentional states without being in them.

³Roy Sorensen pointed this out to me.

REFERENCES

- Baldwin, T. (1990): *G. E. Moore*, London: Routledge.
- Goldstein, L. (1993) : 'Inescapable Surprises and Acquirable Intentions', *Analysis* 53,93-99.
- Gombay, A. (1988): 'Some Paradoxes of Counterprivacy', *Philosophy* 63, 191-210.
- Heal, J. (1994): 'Moore's Paradox: A Wittgensteinian Approach', *Mind* 103, 5-24.
- Hintikka, J. (1962): *Knowledge and belief*, New York: Cornell University Press.
- Lakoff, G. (1975): 'Pragmatics in Natural Language', in E. L. Keenan (ed.), *Formal Semantics of Natural Language*, 256-276, Cambridge: Cambridge University Press.
- Linville, K. and Ring, M. (1991): 'Moore's Paradox Revisited', *Synthese* 87, 295-309.
- Malcolm, N. (1984): *Ludwig Wittgenstein: A Memoir*, Oxford: Oxford University Press.
- Moore, G. E. (1942): 'A Reply to My Critics', in P. Schlipp (ed.), *The Philosophy of G. E. Moore*, 535-677, Evanston Tudor.
- Moore G. E. (1944): 'Russell's Theory of Descriptions', in P. Schlipp (ed.) *The Philosophy of Bertrand Russell*, 175-225, Evanston: Tudor.
- Searle, J. R. (1983): *Intentionality: An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.
- Sorensen, R. A. (1988): *Blindspots*, Oxford: Clarendon Press.
- Williams, J. N. (1979): 'Moore's Paradox- One or Two?', *Analysis* 39, 141-142.
- Williams, J. N. (1982a): 'Believing the Self-Contradictory', *American Philosophical Quarterly* 19, 279-285.
- Williams, J. N. (1982b): 'The Absurdities of Moore's Paradoxes', *Theoria* 48, 38-46.
- 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', *Australasian Journal of Philosophy*, 74, 135-149.
- Wittgenstein, L. (1922): *Tractatus Logico-Philosophicus*, London: Routledge and Kegan Paul.
- Wittgenstein, L. (1953): *Philosophical Investigations*, Oxford: Basil Blackwell.
- Wittgenstein, L. (1974): *Letters to Russell, Keynes and Moore*, G. H. von Wright (ed.), Oxford: Basil Blackwell.

Wittgenstein, L. (1980a): *Remarks on the Philosophy of Psychology*, Vol. I, G. E. M. Anscombe and G. H. von Wright (eds.), Chicago: University of Chicago Press.

Wittgenstein, L. (1980b): *Remarks on the Philosophy of Psychology*, Vol. II, G. H. von Wright and H. Hyman (eds.), Chicago: University of Chicago Press.

Department of Philosophy
National University of Singapore
Singapore, 912409