

© Conor Smith Gaffney, 2011. Please do not share without attribution or reproduce.

Part II, Chapter 4: Moore's Paradox

...“Moore's paradox” is not a name that Moore himself came up with. It first appeared in a 1945 discussion at the Moral Sciences Club in which Wittgenstein referred to “Professor Moore's paradox” as being contained in the assertion “*p*, but I do not believe that *p*.”¹ What Wittgenstein called “Moore's paradox” was actually only a specific interpretation of a more general class of paradoxes that Moore and others had noted beginning in at least 1903. In his first book, *Principia Ethica*, Moore wrote:

It is often pointed out that I cannot at any given moment distinguish what is true from what I think so: and this is true. But though I cannot distinguish *what* is true from *what* I think so, I always can distinguish what I mean by saying *that* it is true from what I mean by saying *that* I think so.²

Later, in a 1942 “Reply to my Critics,” Moore made reference to the specific *paradoxical* nature of propositions that capture the distinction noted above: “If I say that I went to the pictures last Tuesday, I *imply by saying* so that I believe or know I did, but I do not *say* that I believe or know this, nor does *what* I say...*imply* (in the sense of ‘entail’) that I do believe or know it.”³ Yet, the paradox as formulated then still did not have the significance it would soon gain through Wittgenstein's interpretation. Indeed, in a 1940 paper delivered to the Moral Sciences Club, J.L. Austin made use of the “absurd” proposition, “The cat is on the mat, and I do not believe it,” to show that “asserting *p* implies ‘I (who assert *p*) believe *p*’.”⁴ The minutes of this meeting note merely that “Mr. Austin has not explained clearly what the view he was criticizing exactly amounted to, nor has he clearly explained the connexion between this view and other philosophical

¹ Minutes of the Moral Sciences Club, October 25, 1945.

² G.E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903) 132. See also Thomas Baldwin, *G.E. Moore* (London: Routledge, 1990) 226-7.

³ G.E. Moore, “A reply to my critics,” in *The Philosophy of G.E. Moore*, ed. P.A. Schlipp (Evanston: Northwestern, 1942) 541. See pp. 542-3 of the same volume for Moore's explicit characterization of the above phenomenon as a paradox.

⁴ J.L. Austin, “The Meaning of a Word,” in *Philosophical Papers* (Oxford: Clarendon, 1961) 63-4.

questions and errors which, according to him, result from it.”⁵ It wasn’t until 1944 that this phenomenon became known as “Moore’s paradox,” and then only through Wittgenstein’s understanding of the paradox as showing something about the *logic* of assertion, and about logic more generally.

On October 26, 1944, Moore delivered a paper to the Moral Sciences Club again titled “Certainty.” There are no minutes recorded for that meeting to indicate what the paper concerned. However, in a letter to Moore of October of that year, Wittgenstein enthusiastically thanked Moore for a paper he delivered to the club. Wittgenstein wrote that Moore had made a genuine discovery in logic and urged him to publish his paper:

I should like to tell you how glad I am that you read us a paper yesterday. It seems to me that the most important point was the “absurdity” of the assertion “There is a fire in the room and I don’t believe there is.” To call this, as I think you did, “an absurdity for *psychological* reasons” seems to me wrong, or *highly* misleading...Pointing out that “absurdity” which is in fact something *similar* to a contradiction, though it isn’t one, is so important that I *hope you’ll publish* your paper. By the way, don’t be shocked at my saying it’s something “similar” to a contradiction. This means roughly: it plays a similar role in logic. You have said something about the *logic* of assertion. Viz.: It makes sense to say “Let’s suppose: p is the case and I don’t believe that p is the case”, whereas it makes *no* sense to assert “p is the case and I don’t believe that p is the case”. This *assertion* has to be ruled out and *is* ruled out by “common sense”, just as a contradiction is. And this just shows that logic isn’t as simple as logicians think it is. In particular: that contradiction isn’t the *unique* thing people think it is. It isn’t the *only* logically inadmissible form and it is, under certain circumstances, admissible. And to show this seems to me the chief merit of your paper. In a word it seems to me that you’ve made a *discovery*, and that you should publish it.⁶

Unfortunately, a manuscript of this paper is not to be found among Moore’s papers.⁷ However, there is a manuscript available in which Moore discussed his paradox, and

⁵ Minutes of the Moral Sciences Club, February 23, 1940.

⁶ Wittgenstein, *Cambridge Letters*, 315-7.

⁷ Brian McGuinness and G.H. von Wright note in their edition of Wittgenstein’s letters that there are several papers written in pencil among the otherwise penned notes for Moore’s 1941 Howison Lecture, “Certainty” (Add. MS 8875 14/11). The penciled pages, they hold, were likely written at a later date and for a less formal presentation. They claim that “everything indicates” that what was penciled in this manuscript was what Moore delivered to the Moral Sciences Club on October 26, 1944, but do not specify anything that does indicate this. So, I am not sure that everything does indicate this. For one, there are penciled notes in the manuscript of “Certainty” (p. 8A) that are notes for “Certainty,” and not for some other lecture which could have been the 1944 presentation to the Moral Sciences Club. Additionally, there is no mention of the example of Moore’s paradox that Wittgenstein referred to in his letter (“There is a fire in the room, but I don’t believe there is”) in the penciled notes in Moore’s manuscript. But McGuinness and von Wright’s point that this penciled insertion is just one more iteration of a paradox Moore employed

which Moore read to the Moral Sciences Club one year after delivering the paper Wittgenstein so enthusiastically reacted to in the letter above. Additionally Wittgenstein discussed Moore's paradox in some depth in the *Philosophical Investigations*. Though the historical record omits the exact manuscript which so excited Wittgenstein, an accurate sense of the form of the paradox which became "Moore's paradox" can be constructed from Moore's and Wittgenstein's writings about the paradox following the seminal presentation in 1944.

The canonical example of a Moore-paradoxical proposition is a 1945 formulation, the exact phrasing of which Moore credits to Wittgenstein: "It is raining, and I don't believe it's raining."⁸ By Wittgenstein's light, Moore's paradox captures an absurdity that is not a logical contradiction. Since belief and the thing which that belief is about are logically independent, it is not contradictory for one to be true and the other to be false. For instance, I could believe, on some grounds, that it is raining in Chicago and it very well could not be raining in Chicago without my *belief* being contradicted or otherwise rendered absurd by this fact. Were I to find out later, when I spoke with my friend in Chicago on the phone, that it was not raining in Chicago, I would not have the feeling that my holding the belief that it was raining earlier was absurd. However, were I to assert that "I believe it is raining in Chicago, and it is not raining in Chicago," there would be something absurd about this thought. Though both propositions are logically independent of each other—their truth or falsity does not depend upon one another—nonetheless, in asserting them I am somehow uttering nonsense.

As Sydney Shoemaker has pointed out, one obvious explanation of how Moore's paradox obtains in the apparent absence of a logical contradiction is to understand it in

throughout his career is correct, and moreover, the relevant point. The discussion of the absurdity, but not contradictoriness, of the statement "Dogs bark, but I don't know they do," in the manuscript of Moore's Howison lecture (p. 18) is an obvious instance of a Moore-paradoxical sentence. But it is also obviously not the example that Wittgenstein heard on October 26, 1944. The exact formulation of the paradox that so enthralled Wittgenstein is likely not on record. But though we do not have the exact point in the history of the paradox at which it became Moore's paradox, we can nonetheless draw distinctions from the record between the earlier versions and the version that was characteristically Moore's paradox so as to describe the difference between the successful and unsuccessful versions.

⁸ G.E. Moore, Add. MS 8875 12/3/10 in Philosophical papers of G.E. Moore, Cambridge University Archives. Reprinted as "Moore's Paradox" in *Selected Writings*: 207-212.

terms of speech act theory.⁹ This interpretation understands an assertion as a verbal expression that *implies* belief in what is being asserted. Thus, uttering the proposition “It is raining” implies that the utterer believes it, and for him to then say that he doesn’t entails a contradiction. However, as Shoemaker noted, this interpretation does not account for every application of the paradox: “There is something paradoxical or logically peculiar about the idea of someone’s believing the propositional content of a Moore-paradoxical sentence, whether or not the person gives linguistic expression to this belief.”¹⁰ Indeed, there is something paradoxical about simultaneously having the belief “that *p*” and the knowledge “that not *p*,” even if one were not to give expression to these two thoughts. Wittgenstein also noticed this, rejecting the explanation that Moore’s paradox merely denotes “a piece of inconsistent behavior.”¹¹ There is a vast literature on the various interpretive ventures that try to explain why Moore-paradoxical sentences are absurd, but it is well beyond the concerns of this dissertation to review it.¹² What is relevant for this argument about the impact of Moore’s paradox on Wittgenstein’s thought is to unpack Wittgenstein’s reaction to Moore’s “discovery,” and examine the presence of Moore’s paradox in his later writing.

According to Wittgenstein, to take Moore’s paradox as concerning exclusively the *assertion* of belief is to miss the point entirely. In discussing Moore’s paradox in the *Philosophical Investigations*, he elliptically remarked, “Don’t take a hesitant assertion as the assertion of hesitancy.”¹³ What this means is that a hesitant assertion is not my plain assertion along with an assertion “I am hesitant,” but rather *a hesitant assertion*. If I hesitate and then assert, “...I think the answer is twenty-four,” I am not presenting two assertions: 1) that I am hesitant or unsure, and 2) that I think the answer is twenty-four. For I very may well be sure that I *think* the answer is twenty-four, while my

⁹ Sydney Shoemaker, “Moore’s Paradox and Self-Knowledge,” *Philosophical Studies* Vol. 77, Nos. 2-3, (1995): 213. See also, Austin’s use of the paradox in footnote 104 above.

¹⁰ Shoemaker, “Moore’s Paradox,” 213.

¹¹ *Minutes of the Moral Sciences Club*, October 25, 1945.

¹² See, for instance, *Moore’s Paradox: New essays on belief, rationality, and the first person*, ed. Mitchell Green and John N. Williams (Oxford: Clarendon, 2007) for a good overview of the various interpretive themes in philosophical treatments of Moore’s paradox. Moore’s paradox has found employment in the various arenas of epistemology, ethics, justification, and philosophy of mind, among others. See also Kent Linville and Merrill Ring “Moore’s paradox revisited,” *Synthese* Vol. 87 No. 2 (1991): 295-309, for a review of “explanations [of Moore’s paradox] with the exception of Wittgenstein’s” (p. 295).

¹³ Wittgenstein, *Philosophical Investigations*, p. 164.

accompanying assertion of hesitancy seems to suggest that I am unsure of this assertion. We do not say "...I hesitate to assert that I think the answer is twenty-four." The same holds with expressions of belief. Wittgenstein argued that Moore's paradox shows that an expression of belief is not simply an assertion of something that is also believed:

Moore's paradox can be put like this: the expression "I believe that this is the case" is used like the assertion "This is the case"; and yet the *hypothesis* that I believe this is the case is not used like the hypothesis that this is the case. So it *looks* as if the assertion "I believe" were not the assertion of what is supposed in the hypothesis "I believe"!¹⁴

If there is a different "hypothesis" in the expression of a belief, it is not the hypothesis "this is the case," but rather "I believe this" and this latter hypothesis has an entirely different character than hypotheses of the former type (assertions) and hypotheses derived from others' expression of their own beliefs: "My own relation to my words is wholly different from other peoples'."¹⁵ The "asymmetry" between first person and third person expressions of belief, which Wittgenstein noted in his 1945 discussion of Moore's paradox at the Moral Sciences Club, parallels the asymmetry between the absurdity of holding a Moore-paradoxical proposition and a logical contradiction, and additionally highlights the fact that holding and expressing beliefs appear to be governed by rules other than those of logical contradiction and tautology.¹⁶ "How did we ever come to use such an expression as "I believe..."?", Wittgenstein asked, "Did we at some time become aware of a phenomenon of belief? Did we observe ourselves and other people and so discover belief?"¹⁷ The obvious answer Wittgenstein is pointing to is no. But not just no. He was raising the issue of how it is problematic to even begin to think of belief as something analogous to other epistemic positions, like knowledge or doubt.

¹⁴ *Ibid.*, p. 162.

¹⁵ *Ibid.*, p. 163.

¹⁶ Minutes of the Moral Sciences Club, October 25, 1945.

¹⁷ Wittgenstein, *Philosophical Investigations*, p. 162. See also, Jane Heal's "Moore's Paradox: A Wittgensteinian Approach," *Mind* Vol. 103, No. 409 (1994): 5-24, for a more exhaustive examination of the defects of a functionalist theory of belief and Moore's paradox. Heal offers a Wittgensteinian approach to Moore's paradox that critiques this functionalist account of belief. As an alternative, Heal argues that representations of belief have multiple "grammars" (p. 23), one of which is "performative" (p. 21), meaning that the act of representing belief *constitutes* this belief; it is not an exclusively epistemological matter (pp. 21-22). Heal argues that the performative and epistemological grammars of belief clash in Moore-paradoxical sentences which elide the distinction between the two. Heal's reading of Wittgenstein's approach is here very akin to mine.

The argument behind these remarks here is a powerful one about the role of logic in constituting a language-game. It is logic *not* as logicians think of it, as Wittgenstein remarked in his letter to Moore, but a kind of logic that admits more than just contradiction and tautology. Consider the question of how one comes to learn how to hold and express beliefs. Wittgenstein pointed out that we do not learn to believe by observing how other people express beliefs or behave when holding beliefs. If to hold my own beliefs meant simply the same as what it means that someone else holds a belief, then we would be able to apply propositions like “I believe falsely...” to ourselves (because we can apply them to other people), which Moore’s paradox showed to be absurd.¹⁸ Wittgenstein wrote, “If I listened to the words of my mouth, I might say that someone else was speaking out of my mouth.”¹⁹ But this point can be easily misconstrued. Wittgenstein here is critiquing the very notion of a question, “Where did I learn to believe?” since what lies behind this is a paradigm of learning as something typically done through observation and coordination of the expression of one’s understanding with that of other people. Wittgenstein does not give an account of the availability or source of one’s own belief, but rather is concerned with clarifying the *means* with which we can speak and think about our own belief that does not mistakenly assume that there could be a point outside of the language of belief from which we could see how beliefs entered our language. Moore’s paradox shows the paradox of thinking that you are standing at this position.

Immediately following his chapter on Moore’s paradox in the *Philosophical Investigations* is a chapter discussing the puzzle of multiple images of a single object that seem to come from that same object.²⁰ Wittgenstein gave his famous example of the duck-rabbit:

¹⁸ If there were a verb meaning “to believe falsely,” it would not have any significant first person present indicative,” Wittgenstein, *Philosophical Investigations*, p. 162.

¹⁹ *Ibid.*, p. 163. So many of Wittgenstein’s critical remarks in his later philosophy have this form of “If I were to hold/say/believe X, then I would be committed to Y.” This form captures how Wittgenstein is doing more than criticizing a particular view in these remarks; he is raising the issue of why one would feel tempted toward this view in the first place: “If [you *choose*] to hold X, then you are committed to Y” is different than saying “Holding X commits one to Y.”

²⁰ These two chapters were arranged together purposefully. See Peter Winch’s “The Expression of Belief,” *Proceedings and Addresses of the American Philosophical Association* Vol. 70 No. 2 (1996): 8-9: “The three-page long Section x of *Philosophical Investigations*, Part II is a sustained discussion of Moore’s paradox about the verb “to believe”. Section xi, immediately following, thirty six pages and the longest



Wittgenstein's duck-rabbit, p. 194 of the Philosophical Investigations.

Wittgenstein drew the reader's attention to the apparent paradox of seeing two different objects in the same object. The duck-rabbit looks like a duck from one perspective, and a rabbit from another. Yet it is the same object we are viewing. Thus, a single object, outside of the subject, appears to reach into our thoughts and plant two exclusively different objects which we struggle to reconcile. Like cases of Moore-paradoxical sentences which posit two different and exclusive propositions ("I believe *p*" and "*p* is not so"), the duck-rabbit appears paradoxical in being both a duck and a rabbit. But, Wittgenstein admonished, "our problem is not a causal but a conceptual one."²¹ The problem, according to Wittgenstein, is not that the duck-rabbit *causes* two different pictures, but that we conceptualize the process of our perception of the duck-rabbit with the same conceptual means by which we conceptualize the process of our perceiving either the duck or the rabbit:

The concept of an 'inner picture' is misleading, for this concept uses the 'outer picture' as a model; and yet the uses of the words for these concepts are no more like one another than the uses of 'numeral' and 'number'.²²

Wittgenstein was no more concerned in this chapter or in the one preceding it on Moore's paradox to explain the precise conceptual means by which one materializes either an inner picture as opposed to an outer picture or statements of belief as opposed to statements about the world, but rather he was focused on giving one the means to make the distinction at all: "In giving all these examples I am not aiming at some kind of

section in Part II, is a sustained and fascinating discussion of the whole phenomenon of shifting aspects. The last eight pages of the Section drift back, so unobtrusively that the significance of this movement has often been missed, into a discussion of what is involved in our experience of the mental life of other people, their feelings, emotions and beliefs. This movement must be seen as the culmination of the discussion of shifting aspects, and not as a rather inconsequential afterthought."

²¹ *Ibid.*, p. 174.

²² *Ibid.*, p. 167.

completeness, some classification of psychological concepts. They are only meant to enable the reader to shift for himself when he encounters conceptual difficulties.”²³ What marks these various questions of aspect in the duck-rabbit and his other visual examples from questions of misperceiving, what enables the reader to reject the charge that he does not know what he sees and “shift for himself” out of the dilemma of either duck or rabbit but not both, is that the aspect in question, Wittgenstein wrote, bears “a symbol of the logical.”²⁴ The *recognition* that the apparently shifting aspect of the duck-rabbit is not a trick of the “physiology” of the picture to which our language is inadequate, is the recognition that the shifting aspect of the picture is a presentation of the particular *logical* character of our phenomenon of seeing a duck in a particular tangle of lines giving over to the logical character of our seeing a rabbit in a particular tangle of lines. This is the same kind of recognition that Wittgenstein so enthusiastically wrote to Moore about: the recognition that a Moore-paradoxical sentence is not absurd for “psychological reasons,” but rather absurd because it contravenes the *logical* nature of first person assertions of belief by assimilating these assertions to the logic of statements about the external world. This, according to Wittgenstein’s letter and the arguments in the *Philosophical Investigations*, is why we recognize Moore’s paradox as paradoxical.

²³ *Ibid.*, p. 175.

²⁴ *Ibid.*, p. 179.