Journal for the History of Analytical Philosophy
Volume 12, Number 6 (2024)

Bertrand Russell's Doxastic Sentimentalism (and Neutral Monism)

Ryan Hickerson
hickersr@wou.edu
Western Oregon University

ABSTRACT

This paper reinterprets doxastic sentimentalism and neutral monism, as these doctrines appear in Bertrand Russell's "On Propositions" (1919) and The Analysis of Mind (1921). It argues that Russell's theory of belief, in this particular period, posited at least seven distinct types of feeling, but only one type of entity. The paper's principal thesis is that Russell treated believing as feelings, but it also draws the conclusions that monism and sentimentalism are logically independent of one another, and that sentimentalism and (at least one type of) behaviorism are inconsistent, qua theories of belief.

I dislike the heart as an inspirer of beliefs; I much prefer the spleen. (Russell 1922, 645)

1. Introduction

Doxastic Sentimentalism is the doctrine that believing is, or is closely associated with, feeling. The historical source for the doctrine is David Hume (1711-1776), who treated the feeling in question as (or as closely associated with) *force & vivacity*. It is safe to say this has not been one of Hume's most popular theses. Nevertheless, it attracted no less a luminary than Bertrand Russell (1872-1970). In his *Analysis of Mind* (1921), Russell treated believing as the feeling of *assent*. When retrospective it is the feeling of *memory*, and when prospective it amounts to *expectation*.[1] It appears that Russell adopted this sentimentalism at exactly the moment he was coming to terms with *neutral monism*, namely the doctrine that mental and physical entities each derive from a third kind of neutral

[1] See Section 5. My references to "sentimentalism", throughout, are to *doxastic sentimentalism* rather than moral sentimentalism.

"stuff", itself neither mental nor physical.[2] The task of the present essay, first and foremost, is a recovery of Russell's unique type of *doxastic sentimentalism*. But secondarily its task is an historical accounting of the relationship between Russell's lesser-known sentimentalism and his better-remembered monism.

I shall try to be careful about the relation either doctrine has to Russell's understanding of psychology, namely his understandings of the more-traditional introspectionist psychology and the then-budding behaviorism. Russell came to both *doxastic sentimentalism* and *neutral monism* under the influence of William James (1842-1910). And of the approximately forty-eight works Russell read over the four and half months he was jailed in the Brixton Prison, May through September of 1918, only a half dozen were devoted to topics unrelated to psychology. The result of his jail-time study (and his on-going reaction to James) were his essay "On Propositions" (1919) and his book *The Analysis of Mind* (1921).[3] Both are now widely recognized as possessing a more naturalistic and psychologically-informed character than his earlier breakthroughs in logic and epistemology. In these two works Russell produced what is now recognizably a "philosophy of mind". But despite their innovations, their debts to the intertwined histories of psychology and philosophy are still recognizable, so much so that Thomas Baldwin is able to introduce *AMi* in the following way: "For this Humean theory shows that on this issue Russell had not altogether freed himself from the seductive illusions of introspective psychology" (*AMi*, xii).

The "Humean theory" Baldwin refers to is Russell's *doxastic sentimentalism*. And whether or not it succumbs to the "seductive illusions" of introspectionism, I will provide an interpretation below that reads Russellian beliefs, of the period, as irreducible to behavioral dispositions. In order to do so I will first need to give an account of what his beliefs were supposed to be, i.e., one or another of a small set of feelings, and even more fundamentally "spacetime events", themselves

[2] Readers should be warned that even this definition is controversial. While it derives from Russell (1921), there is considerable controversy over whether it equally applies to Russell (1919), and there seems to be no formulation that does not beg an important question. For further details see Section 4, and the updated entry for "neutral monism" by Stubenberg and Wishon (2023). Another important discussion is Wishon (2021).

[3] "On Propositions: What They Are and How They Mean" is hereafter cited in *CP8*. And *The Analysis of Mind* is hereafter abbreviated *AMi*.

neither mental nor physical until joined together in relations with other spacetime events. This makes for an oddly rebarbative combination of topics: feelings and spacetime events. Perhaps this is a reason so few have yet to venture comment on the relation of Russell's *doxastic sentimentalism* to his *neutral monism*.[4] Or perhaps it is merely for fear of the feelings. While a critical literature has grown up around Russell's *neutral monism*, Russell's shifting attitudes about believing have, rightly or wrongly, been overshadowed by his more famous accomplishments.

Undaunted, here is the trajectory for my own comments. While Russell was attracted to both a Jamesian/pragmatist theory of believing, and a Watsonian/behaviorist one, his commitment to belief's form as propositional meant that he could wholly adopt neither.[5] In this period, Russell posited *images* as the contents for the simplest forms of believing, and treated those *images* as propositions. Because the *images* in question were also understood by him (problematically, in ways discussed below) as "exclusively mental," he appears to have remained at odds with *neutral monism*, by his own definition. Nevertheless, and despite disagreement over this key point in the literature, I will argue that Russell was indeed a monist as early as 1919, of an appropriately limited sort: Russell endorsed the reducibility of the mental, but he was a *local monist* rather than a *global monist*.[6] And, as I will show, it was already somewhat earlier, by 1913, that he had adopted James' doctrine that believing is at least associated with a feeling. Russell's rightful sentimental inheritance (from Hume) took a transatlantic detour through the American pragmatist. But I will argue that it then underwent significant expansion in the Brixton Prison. By 1919 Russell was not merely grudgingly affiliating some beliefs with a feeling, as he had in 1913, but positively identifying belief with one or another of a small set of feelings. Despite Russell's avowal of monism and sentimentalism, over roughly the same period, neither doctrine strictly requires the other. So, notwithstanding their interesting intertwinement in Russell's particular case, my secondary conclusion must be that *doxastic sentimentalism* and *neutral monism* need not be paired. My primary interpretive contention will be that *assent*,

[4] An exception is Robert Tully, who writes that the introduction of feelings, "appears to contaminate, rather than merely complicate, the thesis of neutral monism which Russell had adopted" (Tully 1988, 209–24, especially 218–19).

[5] John B. Watson (1878-1958).

[6] Discussion of this difference is in Section 4.

memory, and *expectation*, *conviction*, *doubt*, *non-belief*, and *disbelief*, were all treated by Russell as feelings.

2. The *Theory of Knowledge* Manuscript (1913)

For present purposes, and with significant reservation, I will identify Russell's theories of judgment as equivalent to his theories of belief, and vice versa. Russell did not use the terms 'belief' and 'judgement' interchangeably, preferring the former when referring to dated mental episodes (as might be studied by psychologists), and preferring the latter when emphasizing either that the contents of those episodes are propositional in form, or that they stand in logical relations to (the contents of) other beliefs/judgments. Nevertheless, in 1913 Russell saw fit to write: "When I speak of 'belief', I mean the same kind of fact as is usually called 'judgment' " (*CP*7, 136). And so, at least for present purposes (albeit with reservation), shall I.

That said, it might be useful to begin when Russell had not yet endorsed *neutral monism*. The following quotation comes from his *Theory of Knowledge* manuscript (1913), famously (and at least partially) abandoned due to criticism by Wittgenstein.[7] This paragraph was the conclusion of the published "Chap. II: Neutral Monism":

> For these reasons—some of which, it must be confessed, assume the results of future discussions—I conclude that neutral monism, though largely right in its polemic against previous theories, cannot be regarded as able to deal with all the facts, and must be replaced by a theory in which the difference between what is experienced and what is not experienced by a given subject at a given moment is made simpler and more prominent than it can be in a theory which wholly denies the existence of specifically mental entities. (*CP*7, 32)

So, clearly, Russell was not a *neutral monist* in 1913. However, at least some of the five "reasons" against the doctrine that he had mustered in Chap. II were later recanted.[8] And at the same time, i.e., the second

[7] For a more detailed account of the manuscript and its place in Russell's oeuvre see Pears (1989, 169–82). Another influential reconstruction has been Landini (2011, particularly 251–72). For discussion, particularly in connection to Russell's theory of truth, see Baldwin (2018, 125–49, especially 145–48).

[8] See *CP8* (196) where he labeled at least some of them "not valid". He didn't specify which, unfortunately, but he almost certainly meant one of the first four, as opposed to

decade of the 20th Century, his thinking about believing/judging was undergoing significant revision. Scholars have counted at least four distinct doxastic theories over the period: (1) a dyadic relation theory (–1907), (2) the so-called "multiple-relation theory," (1910–1913), (3) a provisionally adopted "two verbs" theory (1918), and (4) the theory that Russell ultimately settled upon in "On Propositions" (1919) and the *Analysis of Mind* (1921), i.e., the theory positing *images* as the contents of our most basic beliefs. I will delve into only the fourth of these theories below, in Section 5. Presently, i.e., in 1913, Russell was still advancing his "multiple-relation theory" of belief/judgment. For example, Debra believing that Karl is the father of Ryan was treated by Russell as a (mental) subject Debra standing in a (believing) relation to: the (extra-mental) object Karl, the (ordered) "father relation" holding between Karl and Ryan, and the (extra-mental) object Ryan. The belief itself was construed as the set of relations between Debra and the constituents of the fact, and like the fact it admitted of both objects and their relations, i.e., both particulars and universals, as constituents.[9] When Russell was arguing against *neutral monism*, i.e., 1913, he was arguing for a fundamentally dualist theory of believing, namely, one that treated beliefs as relations between mental subjects and extra-mental particulars and universals.

This point is significant for my own purposes because any treatment of believing as a relation will be inconsistent with the strongest forms of *doxastic sentimentalism*, which instead claim that the intrinsic nature of believing is feeling rather than relation. This preliminary observation is hasty (because the relations could always be accompanied by feelings), but readers may see by it where I am heading. For those who go looking, *doxastic sentimentalism* does not pan out of the first six (i.e., published in

the last regarding "emphatic particulars." On my reading Russell retracted arguments mentioned here as arguments against *neutral monism*, but did not disavow them as arguments against a Jamesian theory of belief.

[9]There is a sizable literature on Russell's multi-relation theory of judgment. The most influential accounts are probably the two cited above at the beginning of footnote 7. For a more recent reappraisal of the multiple-relation theory, see Hochberg (2000, 3–40). For an important reframing, including the iconoclastic claim that Russell did not ever abandon the multiple-relation theory of judgment, see MacBride (2013, 206–41). In a previous version of this paper I mischaracterized Russell's multi-relation theory; I am extremely grateful to a reviewer for this journal, who has tried patiently to set me straight about it, multiple times.

The Monist) chapters of the 1913 manuscript, but we do eventually hit pay dirt in one of the later (unpublished) chapters, particularly the one directly devoted to "Belief, Disbelief, and Doubt." In Part II, Chapter IV of the *Theory of Knowledge* Russell discussed what (following James) he called an "emotion" of belief. Acknowledging that beliefs in some circumstances produce the emotion of *conviction* in us, in 1913 Russell declared this emotion could not itself constitute believing (or judgment) because (a) it can be identical in cases where beliefs themselves differ, and (b) because it is not always present when belief is present. Therefore, we should interpret the "emotion of conviction" as one that merely *accompanies* beliefs, in some circumstances. The passage staking these claims is worth quoting at length:

> As regards the "emotion" of belief, the only thing that needs to be understood here is that, however real and important it may be as a psychical fact, it does not concern epistemology, and must be noticed only to avoid the confusions which might result from its unobserved intrusion, like an undesirable alien whose photograph is furnished to the authorities at the frontier. There is an emotion of conviction, capable of many degrees, arising with judgments that hold our unwavering assent, or with perceptions that put an end to a doubt. But this emotion, though it often accompanies judgments, does not by any means constitute them; in fact, it may be exactly the same in the case of two different judgments. A person of a patriotic disposition will feel exactly the same emotion of conviction in entertaining the belief that his country is the best in the world, as in entertaining the belief that his school or club is the best in the country. The emotion is not a relation to the objects of the belief, but a fresh mental fact, caused, perhaps, by the belief, but quite distinct from it. And it would seem that its intensity is not really proportional to our certainty, but to the energy with which we repel doubt. No one feels much of this emotion in contemplating the facts in the multiplication-table, because doubt of them is not conceived to be possible. But religious and political beliefs, just because they are denied or doubted by so many, rouse the utmost fervor of conviction. Such beliefs, however, are not those which a philosopher should take as his model. (*CP*7, 141)

So, Russell here requires we distinguish at least two types of *doxastic sentimentalism*.

> *strong doxastic sentimentalism*: For all x, if x is believed, then the manner of x's conception is a particular feeling.

> *weakest doxastic sentimentalism*: For all x, if x is believed, then the manner of x's conception is sometimes accompanied by a feeling (but sometimes not).[10]

Of particular importance is that *strong doxastic sentimentalism* posits an identity. Believing simply is a feeling, of a particular sort. The most famous proponent of this strong thesis was of course David Hume, who at one time treated the feeling in question as the very same *force & vivacity* of the once-present *impression*.[11] So it is no surprise that in Russell's chapter mentioning *conviction* we also find quotation from Hume's *Treatise*. More interestingly, the passages from Hume are quickly followed by quotation from James' *Principles of Psychology*, wherein James proclaims that, "*In its inner nature belief, or the sense of reality, is a sort of feeling more allied to the emotions than to anything else*" (James 1981, 913).

Weakest doxastic sentimentalism, by contrast, does not posit an identity. Unlike the stronger versions of the doctrine, it is not a claim about the intrinsic nature of believing. Instead it posits a mere association between believing and some feeling (or feelings). Nevertheless, it might still be classified as a kind of sentimentalism, because it still posits some feeling (or other) as at least affiliated with (in at least some cases) believing. To put the point in our contemporary idiom: *Weakest doxastic sentimentalism* still treats believing as having a "cognitive phenomenology", albeit without the stronger thesis that believing is a feeling. I have found it appropriate to label this sort of sentimentalism "weakest," precisely because Russell did not even posit the presumed "association" as a strictly causal one, let alone a logical necessity or sufficiency, at least not in 1913.

On the basis of the passage quoted above (including the unpublished chapter from which it was drawn) I attribute only this weakest form of sentimentalism to Russell in 1913. Unlike beliefs themselves, which were supposed to be relations to objects (and their relations), the "emotion" of *conviction* was supposed by Russell to be a "fresh mental fact". It is caused in us, "perhaps" by the belief relations, but is "quite distinct

[10] This "manner of conception" language is borrowed from Hume.

[11] When Hume began writing his first *Treatise* the feeling of believing was conceived as strictly identical to the *force & vivacity* of the erstwhile *impression*. But, as David Owen has observed (2000, 172–74), by the time Hume was writing the Appendix to that same *Treatise* he had already softened that claim to assert merely that the feeling of believing has an effect similar to *force & vivacity*. For further discussion see Hickerson (2020, 53–57).

from it". While it is a feeling that comes in many degrees, says Russell in 1913, it is not to be confused for certainty. Of particular importance in this latter regard is the possibility of doubt. According to Russell, some certainties command no *conviction*, precisely because there can be no conception of doubt. For example, no one feels *conviction* when contemplating the multiplication tables. Still in that case, certainty is manifest.

It is also worth noting that, while Russell committed himself to this weakest form of sentimentalism in 1913, he also clearly wanted nothing to do with it. He says that the feeling of *conviction* should be scrutinized only to avoid "confusions which might result from its unobserved intrusion, like an undesirable alien whose photograph is furnished to the authorities at the frontier." Such emotions may be interesting to the psychologists, insofar as they frequently accompany religious or political beliefs. But those of us in the business of epistemology (rather than "convictions") need only be on the lookout for them, i.e., on guard against any feelings' infiltration into our proper business. Such sentiments are not the sort "which a philosopher should take as his model" (*CP*7, 141). Compare this to Russell's attitude toward *neutral monism* at the same moment. In the earlier (published) chapter of the manuscript he wrote that *neutral monism* is "preferable to dualism if it can possibly be made to account for the facts", and that it is "largely right in its polemic against previous theories", and that it "has performed an important service to philosophy" (*CP*7, 21, 32, 31, respectively).

If the reader wished to stop now, drawing my conclusions, I would cordially invite this early denouement. My theses are proved on the basis of the two passages quoted above. I conclude that Bertrand Russell was at least the weakest sort of *doxastic sentimentalist* in 1913, even if begrudgingly so. At the same moment he was no sort of *neutral monist*, despite his enthusiasms for that doctrine, because he was still, at that time, presenting arguments against *neutral monism*. Therefore, the two doctrines appear unrelated, at least in Russell's mind.

3. Belief Behaviors in "The Philosophy of Logical Atomism" (1918)

But there is considerably more to the story, for those who would follow it further. Let's skip ahead several years. The years in question, from 1914–1918, were tumultuous for Russell. In addition to delivering the Lowell Lectures at Harvard, the Great War had begun. And Russell had begun a new romantic relationship, with Lady Constance Malleson (Colette O'Niell). The Lowell Lectures were immediately published as *Our Knowledge of the External World* (1914), and Russell became highly involved in anti-war activities in the No-Conscription Fellowship. Because of his stance against the war Russell was dismissed from Trinity College and denied a passport for purposes of returning to Harvard to teach. Just prior to his imprisonment, after publishing a particularly inflammatory two-and-a half-page article in *The Tribunal*, Russell gave his famous lectures on logical atomism at Gordon Square.[12]

Amidst the political and personal tumult, he was still quite clearly thinking about believing:

> I come back now to the theory of behaviourism which I spoke of a moment ago. Suppose, e.g. that you are said to believe that there is a train at 10.25. This means, we are told, that you start for the station at a certain time. When you reach the station you see it is 10.24 and you run. That behavior constitutes your belief that there is a train at that time. If you catch your train by running, your belief was true. If the train went at 10.23, you miss it, and your belief was false. That is the sort of thing that they would say constitutes belief. There is not a single state of mind which consists in contemplating this eternal verity, that the train starts at 10.25. They would apply that even to the most abstract things. I do not myself feel that that view of things is tenable. It is a difficult one to refute because it goes very deep and one has the feeling that perhaps, if one thought it out long enough and became sufficiently aware of all its implications, one might find after all that it was a feasible view; but yet I do not *feel* it feasible. It hangs together, of course, with the theory of neutral monism, with the theory that the material constituting the mental is the same as the material constituting the physical, just like the

[12] *The Tribunal* was a publication of the No-Conscription Fellowship. These biographical details, and the article for which Russell believed he was imprisoned, are printed in Russell (1968, 3–39, 103–6).

> Post Office directory which gives you people arranged geographically and alphabetically. (*CP8*, 194–95)

None of this should be read as denial of the absolutely deep and vital connection between believing and acting. Action, after all, requires belief. But amidst even love and war and politics, psychology marches on, and the behaviorists' question presses: is the success condition for a belief, i.e., its being true, the same as the goal of an action? Or even more fundamentally: are believing and acting simply the same? In 1918 Russell clearly considered the fate of *neutral monism* to "hang together" with behaviorism. But about behaviorism, of at least one kind, he was clearly still harboring some doubts.

Not just in religion and politics, but in philosophy too, we should beware "-isms". And many different ideas have circulated under the generic label "behaviorism", so I will confine my attention below to only a particular species, and only to some of Russell's views about it.

> *reductive behaviorism about belief*: all believing is reducible to either an action or a disposition to an action.

Like Russell in various places, I will refer to this doctrine, i.e., *reductive behaviorism about belief*, when I use the term "behaviorism" below. I will leave open the interpretive question whether or not James espoused *reductive behaviorism about belief*; I will also leave open how it may or may not be connected to other doctrines espoused by pragmatists of various eras (not to be confused for it being a "pragmatic doctrine" itself.) I will also leave aside questions about its relationship with various other ways of construing "behaviorism", and/or the wide variety of understandings of behavioristic methodology in psychology, those circulating in Russell's time or still today. What will be important for my present purpose is merely whether Russell himself became more behaviorist in the specified sense, i.e., whether Russell came to accept *reductive behaviorism about belief* at the moment he adopted *neutral monism*.

There are *prima facie* reasons to think that he might have. As quoted above, Russell claimed in 1918 that this exact sort of behaviorism (about belief) "hangs together" with *neutral monism*. A page prior he had already claimed that this sort of behaviorism (about belief) "hangs together" with James' pragmatism (*CP8*, 193). Assuming that he eventually adopted *neutral monism* (further discussion of this is in Section 4), consider that the "hangs together" relation might be transitive. That

would suggest that anyone who came to embrace *neutral monism* might hold open their arms to behaviorism, and perhaps thereby Jamesian pragmatism. Additionally, the equivocal language Russell continued to use (through 1918) in expressing his attractions to *neutral monism* is very much mirrored in his language about "behaviorism". Also, scholars have argued that while Russell was initially critical of pragmatism and behaviorism, he warmed to both over this period.[13] So, at least *prima facie*, we might think Russell himself eventually adopted *reductive behaviorism about belief*.

I will argue below that he did not. Russell realized that a complete solution to the problem of false belief, not to mention a satisfactory account of propositions, would not come from logical analyses alone, but must also incorporate psychology. This is because the study of truth-taking, whether it be believing or judging, lies on the border of logic and psychology. Policing that border may remain necessary, but we must also look to the psychological frontier for inspiration and solutions. And while Russell suggested a close connection between *neutral monism* and behaviorism, when he claimed in the "Logical Atomism" lectures of the Winter of 1917–18 that behaviorism "hangs together" with *neutral monism*, he could not have meant that the latter logically requires the former. However tight that "hangs together" relation may or may not have been, it cannot have been so tight as a logical requirement. Because *reductive behaviorism about belief* is inconsistent with the sentimentalist doxastic theory that Russell himself would pair with *neutral monism* after 1918.

[13] Cheryl Misak, in particular, has made this argument: "...by 1921 he will have accepted so much of the behaviourist position that he was thought of as a pragmatist" (Misak 2016, 139), and: "But while it is well known that Russell became friendly to these aspects of pragmatism, other aspects of his emerging pragmatism are an under-discussed part of the history of analytic philosophy. That more general shift was toward a behaviourist position, culminating in the 1921 *The Analysis of Mind*, which drew on pragmatist resources" (Misak 2016, 140). Or even more directly: "He [Russell] appeals, that is to the idea, first held by Peirce and taken further by Dewey, that belief is a disposition or a habit of action" (Misak 2019, 68). Compare Landini, who associates the two doctrines but does not stake this interpretive claim directly: "...it is not at all surprising to find that Russell turned to neutral monism and behaviorism for help" (Landini 2011, 272).

4. Interpreting *Russell's Neutral Monism*

Along with William James and Ernst Mach (1838-1916), Russell is remembered as one of three "classical" *neutral monists*. It is often said that Russell named the doctrine.[14] And Russell reports adopting this doctrine in late 1918 or early 1919. There is clear evidence that he thought himself a monist, immediately upon publication of the *Analysis of Mind*.[15] And there is also clear evidence that he retrospectively understood himself to have adopted *neutral monism* at this moment, and never to have subsequently relinquished it.[16] So, nobody doubts that Russell thought of himself as a monist by 1921. But, vexingly, questions remain about whether Russell was entitled to the doctrine, on the basis of what he wrote in *The Analysis of Mind*. Vexingly, this extremely simple interpretive question—was Russell a *neutral monist* in 1921?—has no consensus answer amongst Russell-scholars. Despite his being remembered as one of the three classical monists, and despite Russell's recounting of his own philosophical development, amongst those who know him best (other than Russell himself) there remains this wide disagreement about when, or even if, Russell adopted the doctrine.[17]

[14]Though he did not. See Wishon (2021, 138–39). I have recently learned that this name was already in use by James Ward (1843-1925), several decades prior: "The old materialism has been repudiated and an agnostic or neutral monism—nihilism some would call it—has come into vogue in its stead" (Ward 1906, viii). On this point I have again benefited from another of the reviewers for this journal.

[15]See his letter to the editor of *The Japan Weekly Chronicle*, dated 20 April 1922, responding to a review of *AMi* that had misinterpreted him on exactly this point (*CP9*, 31–32).

[16]See Russell (1959, 134–39). See also Stubenberg (2018).

[17]Some examples: Banks (2014) contrasts James's and Mach's "pure neutral monism" with Russell's "lingering dualism between psychology and physics", calling Russell's position in *AMi*, "weird psychophysical dualism. . .Mentalistic, but at least 'psychologically realistic,' image propositions thus resided in an unhappy halfway house between Russell's old realistic theory and the more thorough neutral monism and causal theory of knowledge and error of Mach and James. . .In effect, Russell had become a dualist and not a neutral monist after all (as Tully also holds: 1993, p. 34)" (Banks 2014, 115, 130, 131, respectively). Following Banks' suggestion, but instead consulting the more-recent Tully (2003), we find much the same attribution of dualism: "Perhaps, then, it is preferable to see Russell's metaphysical doctrine during this transitional period as one of revisionary dualism rather than unalloyed monism. At most, in *The Analysis of Mind*, he has formed a federation with the Neutral Monists, not a union" (Tully 2003, 255). Compare those dualist readings with John G. Slater (1986), editor of numerous volumes of Russell's *Collected Papers*, who says that "By 1921, in *The Analysis of Mind*, Russell had accepted James's rejection of

Russell's simplest articulations and earliest endorsements of *neutral monism* claim that (1) some entities, e.g., *sensations*, are both mental and physical.[18] Alternatively formulated this is the claim that (2) some "neutral" entities (e.g., *sensations*) are neither mental nor physical.[19] Following Stubenberg (2018), I'll call the first of these his "both/and formulation" and the second his "neither/nor formulation". Russell's failure to distinguish the two formulations was not mere carelessness on his part. It is because one of the most distinctive features of the Russellian version of the doctrine is his claim that (3) the mental and the physical are *constructions*, in the sense that they are arrangements/groupings of particulars. The claim here is based upon Russell's earlier work with *logical construction*.[20] In just the way that a mathematical point in space, rather than be posited as an entity, may instead be treated as a class of all objects containing that point, Russell proposes to treat physical objects (e.g., a table) as all *aspects* of that table perceived from various points

consciousness, but he was still not prepared to adopt neutral monism in its entirely" (*CP8*, xxii). Wishon (2015), similarly claims that in 1919 Russell's commitment to *neutral monism* was only "partial and provisional," and that "he is not yet ready to conclude that there is in fact only one kind of fundamental (neutral) stuff in the world." Nevertheless, Wishon disagrees with Slater in claiming that Russell "embraces a comprehensive version of Neutral Monism by the time of his 1921 AMi" (Wishon 2015, 101, 102, respectively). Compare both to Pincock (2019), who says that Russell rejected *neutral monism* in 1919, and embraced it in only a "qualified" sense in 1921, but then fully embraced it in 1925, arriving at the position when he eventually published the *Analysis of Matter* (1927). (Pincock 2019, 325–30). Bostock (2012), on the other hand, says that "Russell's version of neutral monism was never properly 'neutral' or 'monistic' ". He argues that Russell was not a *neutral monist* in 1919, or 1921, or 1927, or really ever (Bostock 2012, 190). And all of the above should be contrasted with the remembrance of Russell as a "classical proponent", i.e., the account originally given by Russell himself, and now nicely re-articulated by Stubenberg (2018), namely that Russell in fact adopted *neutral monism* in 1919 and did not relinquish it subsequently. See Stubenberg (2015, 69–75). Ultimately, I cannot make common cause with everyone listed above, particularly because the theories of belief presented in "On Propositions" (1919) and the *Analysis of Mind* (1921) are substantively the same. I will prosecute my case for this latter claim in subsequent notes, citing passages from both works whenever possible (which is often, as many have an almost identical twin in the other). On the *neutral monism* question, I conclude closest to Stubenberg (2015) or Koç Maclean (2014) and the memory of Russell as one of the "classical" *neutral monists*, from at least 1919 onward.

[18] Three examples are "On Propositions" (*CP8*, 287), see (*AMi*, 120–21); "On Propositions" (*CP8*, 289), see (*AMi*, 25–26), and "On Propositions" (*CP8*, 295).

[19] See *AMi* (25); see "On Propositions" (*CP8*, 289).

[20] For Russell's notion of *construction* see Linsky (1999, 110–38). Another nice account of *construction* is given in Bostock (2012).

of view, and to treat mentality (e.g., a subject experiencing a table) as all *aspects* perceived by a subject. The particulars themselves, i.e., the "aspects" in both cases, were conceived by Russell as neither mental nor physical. Once arranged/grouped they serve (with their fellows) as equivalent to some mental or physical entity, but they do so only in virtue of their arrangement/grouping in a *construction*.

(4) The neutral atoms for such *constructions*, i.e., the particulars themselves, are simply "spacetime events".[21] Here we must fight a severe temptation to interpret anything labeled a "spacetime event" as something physical.[22] Contrariwise, we are easily misled by the fact that Russell called these same spacetime events "sensations", into misinterpreting him as meaning something mental.[23] Arrangements/groupings are broadly conceptual, in the sense that a type (e.g., a property) is identifiable on their basis. But Russell claimed also that (5) such complexes are only generated, and are ultimately only discoverable, by causal relations. Uncontroversially (I hope), physical laws govern causal relations amongst spacetime events. Similarly, according to Russell, psychological laws (Russell's paradigmatic case was the Behaviorists' "Law of Association") were supposed to govern causal relations amongst (possibly those very same) spacetime events. These were not only thought by him as different laws, but different kinds of law, in much the way the Post Office may categorize people either alphabetically (by their names) or geographically (by their street addresses). The task of the physical sciences, and of the psychological sciences, is then accounting for each's

[21] I should again warn readers that, while this is a safe interpretation of *AMi*, it is controversial as applied to "On Propositions" because Russell did not use the language of "spacetime events" in that earlier essay. There is on-going scholarly debate about whether the "particulars" Russell referred to in 1919 are identical to the "spacetime events" proposed as the neutral entities of 1921, and some of that debate is directly related to the broader dispute about the proper interpretation of Russell's monism in 1921, noted above. I am here interpreting Russell as merely choosing new language in the *Analysis of Mind* for the very same "particulars" he had already adopted as the neutral entities in 1919.

[22] See Landini (2011) who considers whether the particulars are B1: "The neutral stuff are physical events in space-time. . ." or instead B2: "The neutral stuff are phenomenal qualitative states. . ." (2011, 291). Ultimately, Landini interprets Russell as claiming that they are B1. See Landini (2011, 297), who interprets them as physical, when Russell clearly claimed that they are both B1 and B2, and neither B1 nor B2.

[23] Stubenberg dryly quips: "Few seem to be persuaded by Russell's case for the neutrality of the events at the basis of NM, and there is a long tradition of classifying Russell's events as mental" (2015, 83).

unique causal arrangement/grouping of the event particulars according to its unique kind of law. But the most basic "stuff" each investigates was supposed to be the same in both cases, i.e., "spacetime events", grouped either as physical effects of physical causes (in the one case), or mental effects of mental causes (in the other).[24]

Russell holds open the question whether (6) a future science may one day investigate all spacetime events under the same kind of law, i.e., whether psychological laws will eventually be treated as merely special cases of physical laws.[25] So Russell was, indeed, a kind of dualist in the period from 1919-1921. But this was not a dualism of a particularly "weird" sort. (At least no weirder than what typically goes about in philosophy.) Russell was a *nomological dualist*, insofar as he claimed (7) there are, according to the current state of the art, two fundamentally different kinds of law. I must emphasize that this last claim was made provisionally, as he believed it "highly probable" that psychological laws would eventually be reduced to special cases of physical laws. But far from regarding this (i.e., our present day) *nomological dualism* as an impediment to the development of psychology, Russell thought the discovery of distinctively psychological laws was a vital part of the science's development (just like Boyle's Law was necessary for the development of the Kinetic Theory of Gasses.)

[24]Still, it is not immediately obvious what Russell meant by "spacetime events" in the supposedly neutral sense. Clearly, they were not meant to be points, but regions/durations. Somewhat less clearly, I think they were not meant to be regions of spacetime themselves, but occupants of regions, i.e., events. Leaving their grouping to the conventions of one or another science seems fair enough, but that creates a knock-on problem for claiming "the same" event is variously so grouped. "The same" according to which's criterion of individuation? I am grateful to Ivan Welty for drawing this problem to my attention, and wish that I could give a better account on Russell's behalf.

[25]He imagines this a remote, if "highly probable" eventuality (*AMi*, 36–37). Importantly, the subject here is an eventual *inter-theoretic reduction*, not to be confused for the unique type of *ontological reduction* (via *construction*) under discussion thusfar. In much the way that Bernoulli needed Boyle's Law, before deriving it from Newtonian mechanics, the derivation of psychological laws from physical ones will only be accomplished (presuming it will be) by future scientists standing on our contemporaries' shoulders. In all cases, and this means even by the lights of their future science, the physical and mental should be treated as "logical fictions," rather than ontologically fundamental. This is because they are capable of *construction* using only materials provided by neutral spacetime events and (what remain today, but perhaps what merely remain today) the two distinct kinds of causal relation.

This broadest sketch of *Russell's neutral monism* can be reinforced with a bit of standard philosophical technical terminology: we say that *intrinsically* the particular spacetime events are neither mental nor physical.[26] Spacetime events themselves, i.e., *intrinsically*, are neither mental nor physical, even when arranged in a physical or mental *construction*. This is because they have the character of being physical or mental merely *extrinsically*, i.e., only insofar as they are causally related to other particular spacetime events. Nevertheless, any particular event could (in principle, if perhaps only by the lights of a science of the remote future) be part of a *construction* of the one type or other. This is enough to allow us to understand why Russell was sanguine rather than sloppy when using the both/and and neither/nor formulations interchangeably. On Russell's account, the two modes of expression are strictly equivalent. The former means possibly "both physical and mental," depending on the causal relations that some given spacetime event has entered into with others, and the latter means "neither physical nor mental" *intrinsically*.[27]

Crucially, for Russell, (8) not all particulars serve atomically in such *constructions*. Presumably there are many remote regions of spacetime containing events that will never be useful as building blocks for anyone's mental life. Correlatively (and more controversially) Russell thought that there are some spacetime events that are not capturable in the nomological net of the physical laws, i.e., not currently capturable given our current level of scientific development. These latter particulars (i.e., spacetime events) Russell labeled "images". *Images*, unlike *sensations*, currently figure only in the psychological laws, rather than in both the psychological laws and the physical laws. I will provide a further attempt at interpreting *images* in Section 5.

But presently, the eight claims numbered above sketch *Russell's neutral monism* in its broadest outline, according to me. It is particularly claim (3), with its emphasis on *construction*, that gives the sketch its distinctively Russellian character, so I would hazard to call these eight claims "Russellian Monism," if that name hadn't already been stolen by

[26] For further discussion of the intrinsic/extrinsic distinction, see Humberstone (1996, 205–67). My usage here is the first of the three Humberstone discusses.

[27] A similar point, albeit with slightly different emphasis, is already made by Stubenberg (2015, 73).

philosophers of mind to mean something different.[28] I'll instead label them "Russell's neutral monism," with the proviso that what I have intended to sketch here is the metaphysical doctrine espoused by the historical Bertrand Russell, from 1919 and following. Summarizing:

> *Russell's neutral monism*: the mental and physical are not ontologically fundamental, but can instead be treated as *constructions* using particular spacetime events that are themselves neither *intrinsically* mental nor *intrinsically* physical.

There remains, however, a basic equivocation in the doctrine just labeled, one that is related to claim (8), above. To proceed, I think it necessary to also distinguish two quite different interpretations of *Russell's neutral monism*, versions that I will now dub "global monism" and "local monism", and only one of which was espoused by Russell himself:

> *global monism*: all particular spacetime events, which are themselves neither *intrinsically* mental nor *intrinsically* physical, are used in some *constructions* of the mental and some *constructions* of the physical.
>
> *local monism*: some particular spacetime events, which are themselves neither *intrinsically* mental nor *intrinsically* physical, are used in some *constructions* of the mental and some *constructions* of the physical.

With the help of a logical apparatus handed down to us from Russell himself, we can make the quantifications intended in these possible interpretations more conspicuous. When the universe of discourse is spacetime events and the entities constructed out of them, and Cxy means "x is a *construction* out of y," and Mx is "x is mental," and Px is "x is physical, and Sx is"x is a spacetime event," then *global monism* becomes:[29]

[28] They use it to mean a family of monistic doctrines inspired by Russell, none of which are directly attributable to him. For discussion see Alter and Nagasawa (2015, particularly 58–90).

[29] There is a significant problem with this universe of discourse, articulated clearly by one of my anonymous reviewers. I confess that I am not able to overcome it, and can only endeavor to make it more explicit: In order for me to draw the distinction immediately below I must discuss not only events, but also "mental entities" and "physical entities." But it is only the events (i.e., only the "*x*'s") that are members of Russell's fundamental ontology. According to Russell, the mental and physical (i.e., the "*y*'s" and the "*z*'s") are not. So, readers must be wary of this U.D., as it is annoyingly wider than Russell's ontology. I am about to bind variables, the values of which aren't ontologically fundamental! Still, I think this little bit of formality can be useful for seeing the relevant distinction, even if it must be taken *cum grano salis*.

All spacetime events are used in both mental and physical *constructions*.

$\forall x\{Sx \supset [\exists y\ (My\ \&\ Cyx)\ \&\ \exists z\ (Pz\ \&\ Czx)]\}$

. . .and *local monism* is merely:

Some spacetime events are used in both mental and physical *constructions*.

$\exists x \exists y \exists z\ (Sx\ \&\ My\ \&\ Cyx\ \&\ Pz\ \&\ Czx)$

Now, I know quite well that the whole point of being a *neutral monist* is being a monist. The entire reason a monist metaphysics is attractive to philosophers like James and Russell is because it promises to reduce the number of (types of) fundamental entities. Ockham's Razor is sharp for the job. And while the popularity of the shave may be a bit perplexing to those who remain wooly, we can all appreciate the sensibilities of those who think that no right-minded individual ought to walk around merely half-shorn: *Entia non multiplicanda praeter necessitatem*. Russell regarded this as "the supreme methodological maxim in philosophizing," baldly declaring that it, "prescribes James's theory as preferable to dualism if it can possibly be made to account for the facts" (*CP*7, 21). And this was Russell's own, clearly expressed, motive for embracing *neutral monism*, even while he himself remained a dualist, i.e., prior to the end of 1918.[30]

But we must notice that the difference between *global monism* and *local monism* is not a difference in the number of (types of) fundamental entities. Each posits only the one type of fundamental entity, i.e., neutral spacetime events. Both treat all mental entities and all physical entities as *constructions* using only those neutral event tokens. So, the famed Principle of Parsimony remains silent with respect to any difference between them. Their difference is instead that *global monism* stakes, additionally, the much stronger claim that every particular is used in both a mental and a physical *construction*. And that is a strong claim indeed. Some philosophers have believed that the mental is spread evenly into every nook and cranny of the universe. (These philosophers are traditionally called "panpsychists".)[31] Compare their doctrine to the

[30] See *CP8* (195). Compare Russell (1959, 12–13). A distinctive feature of Russell's understanding of the principle is that it does not require actively denying the existence of the "entities" dispensed with, but merely stipulates we show (wherever possible) that we do not need them, via *construction*.

[31] For further discussion of the relationship between *neutral monism* and *panpsychism*, see Stubenberg (2015) or Holman (2008).

more popular, if less frequently noticed, assumption that the physical is spread evenly into every nook and cranny of the universe. (A label for this more popular assumption might be "panphysicalism"?) Clearly, the *global monist* is both a *panpsychist* and a "panphysicalist". But not every *neutral monist* need be. *Local monism* is the more modest thesis that only some particulars are party to both sorts of *construction*. And that claim scrupulously leaves open whether some other particulars are used only in the one sort of *construction*, or only in the other sort of *construction*, or neither.

Now compare both to the following key doctrines, typically thought to be implied by *neutral monism*, and here couched in a way amenable to Russell:

> *the reducibility of the mental*: all mental entities can be treated as *constructions* out of spacetime events
> $\forall x[Mx \supset \exists y(Sy \,\&\, Cxy)]$

> *the reducibility of the physical*: all physical entities can be treated as *constructions* out of spacetime events
> $\forall x[Px \supset \exists y(Sy \,\&\, Cxy)]$

Here is where the Principle of Parsimony pipes up. To my mind, philosophers can't call themselves *neutral monists* unless they are ready to endorse both the *reducibility of the mental* and the *reducibility of the physical*. Russell is often (if not uncontroversially) thought to have been committed to both in *AMi* (1921), but it remains a matter of debate whether he was also committed to both in "On Propositions" (1919). I venture out onto a limb in claiming that he was.[32] The important point presently,

[32] Here yet again I owe thanks to the journal's reviewers, for pressing this point. The *reducibility of the mental* (as defined above, with the universal quantifier) is not explicitly endorsed by Russell in "On Propositions". I interpret his claims there as committing him to this doctrine, but readers are advised to compare Wishon (2015, 91–118) or Wishon (2021, 138–39). One of the reviewers has pointed out that, unlike me, Russell was not shy about claiming "neutral monism" for some elements of mental life but not others, offering the following passage where Russell identified Stout as a "neutral monist" about sensations, but not generally: "There seems in this passage an acceptance, as regards sensation, of the doctrines of neutral monism which Professor Stout would be far from adopting generally" (*CP7*, 21 note 4). If Russell thought Stout could be a monist about sensations but not generally, then perhaps Russell himself adopted such a view? Russell's manuscript from 1918 "On Sensations and Ideas" (*CP8*, 252–55) lends itself in this regard. And readers may also be advised to compare Lecture I of *AMi*, the exact place where Russell introduces *neutral monism* to the book: "I should admit this view as regards sensations. . .But I should

however, is that neither of these doctrines implies *global monism*. Nor do they jointly imply it. My key interpretive application of the distinction made above goes like this: it was open to Russell to claim to be a *neutral monist*, in a sense that embraced both the *reducibility of the mental* and the *reducibility of the physical*, but only in the sense in which he was a *local monist*, i.e., without staking any additional claim that all spacetime events are used in *constructions*. It is my contention (somewhat less controversially) that this was his metaphysics in *The Analysis of Mind* (1921), but also (somewhat more controversially) in "On Propositions" (1919).

As I have now suggested, one explanation of the vexing dissensus over *Russell's neutral monism* is insufficient appreciation of the difference between *global monism* and *local monism*.[33] But, even more directly, I would now like to claim that the reason we fail to agree about Russell is our lingering doubt about whether he really and truly had, by 1919, committed himself to the *reducibility of the mental*. It is easy to see why one may harbor doubt on the point. *Sensations* were claimed by Russell to be both mental and physical. But then, in almost the same breath, he called *images* "purely mental".

> But when we come to consider the stuff of the two sciences, it would seem that there are some particulars which obey only physical laws (namely, unperceived material things), some which obey only psychological laws (namely, images, at least), and some which obey both (namely,

say that images belong only to the mental world. . ." (*AMi*, 25). In each of these places Russell seems amenable to "piecemeal neutral monism", or "impure neutral monism", i.e., to adopting "neutral monism" for *sensations* only, but not also for *images*. Nevertheless, I do not interpret Russell in that way. For further argument against doing so see Koç Maclean (2014, 122–28). My own reason is because what "belongs only to the mental world", i.e., what belongs only to our mental world right now, depends entirely on the current state of our art: the very same spacetime events currently deemed "exclusively mental" (namely, *images*) could well become elements in *constructions* using some future science's better understanding of the physical laws.

[33] Here I must part ways with Bostock (2012), who writes: "Most of the ingredients from which minds are constructed do not also occur in matter, and conversely most of the ingredients from which matter is constructed do not occur in minds. For most of them are unsensed sensibilia, but it is only actual events of sensing that occur in actual minds. So it is only these events that he can claim to be genuinely 'neutral' " (2012, 190). While the premises of this interpretive syllogism are true, its conclusion does not follow. Bostock presumes that in order to be "neutral," a spacetime event must be used in both a mental *construction* and a physical *construction*. But that does not respect the difference between *global monism* and *local monism*.

> sensations). Thus sensations will be both physical and mental while images will be purely mental (*CP8*, 289).[34]

What did he mean by that, exactly? Could it be possible that Russell failed to recognize what so many of us plainly have: that were he to posit irreducibly mental entities it would make his metaphysics non-monistic? Sensitive to the issue, one family of readings has been that Russell only slowly came to his more mature monism. On these gradualist interpretations Russell at some point only "provisionally or partially adopted" neutral monism, and then later came to embrace it "more fully".[35] This sort of account has been given its most recent and able exposition by Wishon.[36] These gradualist interpretations face the challenge of explaining what it would mean to only "partially adopt" a doctrine like neutral monism, however defined. I am tempted to say in response that however many doubts one may harbor, one either believes in some (type of) entity or one doesn't. I hasten to add that I do not take myself to have settled the interpretive question on this point. Russell clearly could have come to *neutral monism* gradually, accepting piecemeal some reductions for some bits of our mental life but not others. But I do wish to offer *local monism* instead, i.e., a kind of monism including the strong (i.e., universal) claim about the *reducibility of the mental*, as a kind of monism that could have been wholly adopted by Russell in 1919, even if the evidence he did so only comes later. I humbly suggest that this was the doctrine Russell came to believe in Brixton, wholeheartedly. My reading of Russell has the virtue of a somewhat greater simplicity than the "gradualist reading," as it does not attribute to him a change in ontology from 1919 to 1921.[37]

Perhaps we have been too stingy in granting *local monism* its due as a type of *neutral monism*. But our interpretive dissensus also demonstrates that we have yet to fully grasp what Russell meant by *images*. (Given the complexities of *AMi,* this latter fault is significantly more excusable in us.) If we could better understand how *images* did not commit Russell to the irreducibility of the mental, then we would be able to better trust his

[34] See also *AMi* (25–26).

[35] This is explicitly argued for by Slater in *CP8* (xxii), and Tully (2003, 332), but can also be traced in many of the other interpreters cited above.

[36] Wishon (2015, 92–118) especially, but also Wishon (2021, 138–39, note 36).

[37] Here, yet again, I have profited enormously from criticisms by the reviewers for this journal.

reporting of his own philosophical development. To be a *neutral monist* Russell must have claimed that neither mental nor physical entities are ontologically fundamental. And while that is a far cry from *global monism*, making this explicit has the additional benefit of focusing us squarely on the particulars (presently) found in only the mental *constructions*. Inevitably, our understanding of *Russell's neutral monism* will hinge upon this question about the nature of *images* and their relationship to *sensations*. Were *images* also meant by Russell to be "neutral," just as the *sensations* were, or were they instead supposed to be irreducibly mental? Whatever else has by now emerged, I hope to have established that the key to settling one interpretive dilemma, regarding whether Russell was or was not a *neutral monist* in "On Propositions" (1919) and *The Analysis of Mind* (1921), is what Russell claimed about *images*.[38]

5. *Image-Propositions* and The Seven Sentiments

So, Russell's proposed relation of *images* to *sensations*, and the nature of *images*, will be the crux of any interpretation of *Russell's neutral monism*. If *images*, like *sensations*, were meant to be neutral spacetime events, then Russell had his metaphysics. But if they were not (as suggested by a plain reading of the passages wherein they were declared "purely mental"), then it appears Russell was not actually a *neutral monist* even in 1921, despite his ambitions to that view. In some places Russell claimed that *images* are "radically distinct" from *sensations* insofar as the former are "not amenable to the laws of physics" (*CP8*, 287).[39] Nevertheless, in other places he claimed that "images and sensations cannot always be distinguished by their intrinsic nature" (*AMi*, 151).[40] What is that supposed to mean exactly? Does it imply that they can, sometimes, be distinguished by their intrinsic nature? Or does it mean that they can be distinguished in some other way? Or does it mean that Russell had changed his mind about the nature of *images* from 1919 to 1921? For even

[38]This point is already well-appreciated by Pincock (2019, 322–25). Pincock and I may disagree about whether Russell was a *neutral monist* in 1919, but only because we disagree about what it meant for Russell to have called *images* "purely mental entities" in 1919.

[39]Compare *AMi* (117).

[40]Compare *CP8* (286).

the most careful readers of Russell, these passages present something of a puzzle.[41]

The question at hand is directly related to my broader aim of understanding Russell's *doxastic sentimentalism*, because Russell treated *images* and *sensations* as the most basic contents of belief. Russell adopted a roughly Brentanian schema for understanding the structure of belief: each belief possess an act character (what Russell called an *attitude*), a content (which Russell, characteristically, understood to be a *proposition*), and an object (the fact making the belief true or false, which following Meinong, Russell referred to as the belief's *objective*).[42] The world-historical significance of Russell adopting this tripartite schema cannot be overstated. Through Russell, this schema is responsible for the "propositional attitudes" of the 20th Century's philosophy of mind. *Propositions*, i.e., whatever can be true or false, were identified by Russell in *AMi* as "the contents of actual and possible beliefs" (*AMi*, 241).[43]

This is not to say that Russell was a Brentanian. The opening pages of *AMi* begin not merely with a statement of *neutral monism*, but with a rejection of Brentano's Thesis, namely Brentano's characteristic claim that all and only mental phenomena possess intentionality (or as Russell puts Brentano's Thesis, that mental states possess "essential reference to an object").[44] Russell reports that he and G. E. Moore had previously committed themselves to two of Brentano's three schematic elements: a mental act/subject, and an object to which that act/subject is related, rejecting only the "Brentanian" category of the mental content. But by 1919 Russell was no longer satisfied with that two-part theory, claiming that the act/subject is "schematically convenient, but not empirically discoverable" (*CP8*, 294).[45] This, in turn, meant wholesale rejection

[41] My own answers to these questions, respectively and respectfully, will be: "no, not via introspection", "yes, by their causes", and "no".

[42] Franz Brentano (1838-1917). Russell attributes not only the language of the "objective", but also this tripartite structure to Brentano's "Austrian successor Meinong" (*AMi*, 16); see "On Propositions" (*CP8*, 294). See also "On Sensations and Ideas" (CP8, 252). Baldwin nicely catches that Russell has made a minor interpretive mistake on this point: the basic schema is more appropriately attributed to one of Brentano's other students, Kazimierz Twardowski (1866-1938). See Baldwin (2003, 441).

[43] See "On Propositions" (*CP8*, 278, 296).

[44] For this rejection and its reasons see *AMi* (13–21). See "On Propositions" (*CP8*, 294–96). Russell reports that he had endorsed Brentano's Thesis, "until very lately" (*AMi*, 15).

[45] See *AMi* (17–18).

of "Brentano's view", i.e., the view that treats mentality as essentially characterized by an "objective reference". Even the most basic elements of mental life must then be re-conceptualized: sensation, e.g., could no longer be treated by Russell as a relation of *acquaintance* between subject and *sense-datum*. Instead, neutral *sensations* (i.e., spacetime events) could be treated as elements of the world (under a physical *construction*) and/or elements of a mind (under a mental *construction*). Russell's broadly Humean discovery that the subject is a "logical fiction" allowed him to reject the very notion of the "mental act", as not merely fictitious but "unnecessary and fictitious" (*AMi*, 17). However else he may have understood it, Russell took his conversion to *neutral monism* to be bound up with a rejection of Brentano's Thesis.

Nevertheless, and despite this conversion, Russell still appealed to "Brentano's" tripartite schema when theorizing about belief.[46]

> In Lecture I we criticized the analysis of a presentation into act, content and object. But our analysis of belief contains three very similar elements, namely the believing, what is believed and the objective. The objections to the act (in the case of presentations) are not valid against the believing in the case of beliefs, because the believing is an actual experienced feeling, not something postulated, like the act (*AMi*, 233).

Let us focus for a moment on the content-side of the schema, namely, the *propositions*. Russell writes that, "The content of a belief may consist of words only, or of images only, or of a mixture of the two, or of either or both together with one or more sensations" (*AMi*, 236).[47] It is perhaps more familiar to think of *propositions* as composed of words. For example, when one believes that "Two times five equals ten", or "London is the capital of England". Or when one believes "There is a tiger in the classroom next-door". Typically, propositions like those are associated with declarative sentences. But, clearly, the content of a belief need not be so wordy.[48] The latter example may have content composed entirely

[46]See Apostolova (2017, 325–27), who provides a similar reconstruction. See also Pears (1975): "His [Russell's] rejection of the subject made it necessary to find something else to put in its place. He put what he calls 'the content of the belief' in this position" (1975, 235). On my reading, Pears is not quite correct on this point; what Russell instead put in this position is the *attitude*, i.e., the feeling of believing.

[47]See "On Propositions" (*CP8*, 296–97).

[48]In Appendix C to the second edition of *Principia Mathematica*, Russell writes, "Some people maintain that a proposition must be expressed in words before we can believe it;

of *images* of tiger and classroom. The belief about London could include more or less vague *images* of The Palace of Westminster (or perhaps not). Maybe your belief about twice five being ten is accompanied by an *image* of two groups of five objects? (Or maybe not.) Whether a belief includes an *image,* or not, it is one of Russell's insights in *AMi* that *images*, even when unaccompanied by words, can be propositionally structured, i.e., they may still be true or false. So, they may still be believed. It matters significantly, when identifying a particular belief, whether one has pictured the tiger within the classroom, or has instead pictured a different animal, or has instead pictured a tiger elsewhere. An *image* may depict a window located to the left of a doorway. A different *image* may depict a window to the right of a doorway. So those *images* are *propositions*, expressing spatial relations of windows to doorways. Even wordlessly, *images* express some (almost certainly multiple) *proposition(s)*.[49]

According to Russell in *AMi*, these *image-propositions* are the most basic type of *proposition*. He attributes beliefs (with image-propositional content) to non-human animals. "There is no reason why memory-images, accompanied by that very simple belief-feeling which we decided to be the essence of memory, should not have occurred before language arose; indeed, it would be rash to assert positively that memory of this sort does not occur among the higher animals" (*AMi*, 242). Here he gives an example of a pigeon failing the mirror self-recognition test: the pigeon believes that his own reflection is a rival pigeon, and becomes agitated. Without quibbling over what may or may not count as a "higher" animal, Russell's claim is that animals no more sophisticated than pigeons have genuine (let us scrupulously avoid calling them "literal") beliefs. This is a claim that would be significantly less plausible were the propositional contents required for believing always verbal.

Without entering into the semantic niceties of Russell's account of the meanings of *images* and their relations to the meanings of words (niceties

if that were so, there would not, from our point of view, be any vital difference between believing and asserting" (1925, 662).

[49] In this paragraph I am attempting to countenance Russell's claim that an *image* can be a type of *proposition*, but I find it highly dubious that each *image* would instantiate only a single *proposition*. Individuation of the so-called "image-propositions" is under-discussed by Russell, in my humble opinion, and probably would require some of the semantic "niceties" mentioned in the paragraph after next, i.e., those that I am about to skip over.

that I happily skip over here), we can easily account for the origination of the *images* themselves: they were supposed by Russell to have arisen from the *sensations* themselves. Russell espoused a version of Hume's Copy Principle: "*. . .all our simple ideas in their first appearance are deriv'd from simple impressions, which are correspondent to them, and which they exactly represent*" (Hume 2000, 9). While endorsing neither Hume's notions of *idea* nor *impression*, Russell treated all simple *images* as "derived from" simple *sensations*. This made him a Humean in at least three important respects. First, like Hume, Russell understood *images* and *sensations* to be simple or complex. And like Hume, he did not claim that every *complex image* is an exact copy of a *complex sensation*. He merely claimed that all of the simple elements of each complex *image* had its origin in some simple *sensation*. Second, Russell claimed that *images* resemble *sensations*. Without endorsing Hume's notion of *force & vivacity*, this resemblance came complete with a Russellian claim that the "copies," typically, are ". . .more dim or vague or faint. . ." versions of their otherwise similar-looking *sensations* (*AMi*, 154).[50] But third, and most importantly, Russell endorsed the Humean notion that the relationship between *images* and *sensations* is fundamentally causal. Sometimes a *complex image* is effected all at once by some *complex sensation* (in cases of an associated memory). But in all cases the constituents of a *complex image* were once effects of some *simple sensation*. Russell adopts a technical term for *sensations* that are direct (i.e., sensory) causes of their resembling *images*; he calls those "prototypes" (*AMi*, 179–80, 184–87, 207–209, 219, 289–91).[51]

On my interpretation of Russell's *AMi* (1921) and "On Propositions" (1919), *images* are neutral spacetime events, just like the *sensations* that caused them.[52] Attention to the fact that Russell endorsed Hume's Copy Principle helps explain why *images* and *sensations* "cannot always be distinguished by their intrinsic nature". This is so for two simple reasons: because they resemble one another, and because they have no difference in their intrinsic nature. They differ principally in their immediate causes. For these reasons it is possible, though not typical, to be introspectively confused about whether or not one is perceiving a content or merely remembering it, or merely hallucinating it rather than

[50] See "On Propositions" (*CP8*, 292).

[51] See "On Propositions" (*CP8*, 293, 305).

[52] Koç Maclean has already put the point most eloquently: "The doom of dualism is overturned when we look at how Russell defines images" (2014, 123–24).

perceiving it or remembering it. Of course, typically, such differences can be ascertained introspectively. But not always. Establishing the real differences requires the appeal to causes.

This is to say, given Russell's interpretation of the Copy Principle, there is always a sharp distinction to be made between an *image* and a *sensation*.[53] The difference is the causal one.[54] Russell argued for this point quite explicitly, entertaining Humean proposals that *images* and *sensations* could instead be differentiated by Hume's specific feeling of *force & vivacity*, or by a similar "feeling of reality". Russell rejected both of these rival (even more sentimentalist!) attempts at differentiation. See (*AMi*, 145–54). *Images* are caused by *sensations*, and not vice versa. That is an *extrinsic* difference rather than an *intrinsic* one. We can imagine a future psychology, advanced beyond our own, capable of deriving all psychological laws from physical laws. In that case Russell would no longer claim that *images* are "exclusively mental", because they could then be used in physical *constructions* rather than (as in Russell's day, but also our own) solely mental *constructions*. Still, even in that remote possible future there would remain the *extrinsic* difference between the two types of neutral spacetime events, insofar as the *sensations* would remain the causes of the *images*, and not vice versa.

My reconstruction thus far, concerning Russell's *images*, has been on the content-side of the "Brentanian" schema distinguishing content from *attitude*. But already it has required periodic hinting at differences in *attitude*. To complete my interpretation of Russell's *doxastic sentimentalism*, I now face the believing itself. It would be a stretch for me to say that *doxastic sentimentalism* ever took center stage in Russell's writings, but it does seem to have drawn more of his attention after his study of psychology in Brixton. In "On Propositions" he employed the same quotation from James' *Psychology* that he had used in his *Theory of Knowledge* manuscript, i.e., the third paragraph of Chapter XXI beginning, "*In its inner nature, belief, or the sense of reality, is a sort of feeling more*

[53] See any of the dualist readings of Russell listed above, which invariably treat *images* as irreducibly mental. Also see Engelmann (2012), who interprets *images* simply as the *sensations* that are used in the mental *constructions* (2012, 283–311, especially 285). This was not Russell's account; Russell treated *images* and *sensations* as having no *intrinsic* difference, but insisted upon their causal differentiation.

[54] Here I again follow Koç Maclean (2014, 124–25).

allied to the emotions than to anything else".[55] But gone now is any explicit reference to Hume. And gone now is the attempt to tie this Jamesian doctrine back to Hume, for purposes of criticism. Instead, by 1919, Russell simply attributes sentimentalism to James, and then follows that immediately with:

> In the main, this view seems inevitable. When we believe a proposition, we have a certain feeling which is related to the content of the proposition in the way described as "believing that proposition". But I think various different feelings are collected together under the one word "belief," and that there is not any one feeling which pre-eminently *is* belief (*CP8*, 298).

With due deference to My Lord Russell, the view being expressed here is hardly "inevitable". If it were "inevitable" it would have been considerably more ubiquitous in the literatures of philosophy and psychology over the past century. And if it were "inevitable", Russell might have espoused it himself prior to 1913. I am also tempted to add: were it "inevitable", Russell himself might not have labeled doxastic sentiments "undesirable alien[s]", whose photographs ought to be circulated amongst "the authorities at the frontier".

So, here we witness a shift in Russell's own attitudes, at the very least his change of heart about *doxastic sentimentalism*. But even more significant is the shift in his exposition of the contents of the doctrine itself. Consider the following version of sentimentalism, still weaker than Hume's, but nevertheless significantly strengthened:

> *Russell's doxastic sentimentalism*: For all x, if x is believed then the manner of x's conception requires one of seven feelings: *assent*, *memory*, *expectation*, *conviction*, *doubt*, *non-belief*, or *disbelief*.

Unlike Hume, Russell did not posit a single feeling or complex of feelings identical to believing. We recall that Hume's claim was not merely that believing is a feeling (i.e., Hume's identity thesis), but also that believing is one particular feeling amongst others. Hume's sentimentalism was particularly strong in this second dimension too, namely, insofar as he claimed that there is only one feeling of believing (i.e., Hume's singularity thesis). Another way to put this same point is to say that it is one thing to claim that believing has a "cognitive phenomenology", but it is quite another to claim that believing is

[55] Again, Russell's quotation is from James' *Principles of Psychology, Vol. 2* (1981, 913).

only one animal in the phenomenological zoo. Recall that I contrasted *strong doxastic sentimentalism*, above, with the "weakest" version of the doctrine, i.e., the claim that believing is merely "accompanied" by a feeling (of perhaps a variety of different types). *Weakest doxastic sentimentalism* makes only some sort of causal claim (and in Russell's case in 1913, it was a very relaxed sort of causal claim) about the relation of believing to feelings; it posits feeling as neither necessary nor sufficient for belief. So now we must also contrast Hume's sentimentalism with Russell's, but also by highlighting that the latter came to posit doxastic sentiments of a variety of different types. (Recall that in 1913 Russell discussed only *conviction* as amongst the beliefy-feelings.) By 1919 Russell's sentimentalism has moved significantly closer to Hume's in one of these dimensions, but away from it in the other: believing is (rather than merely accompanies) feeling, but that means one or the other of a small set of different feelings.

I should also say that, there appear to be places where Russell endorsed Hume's identity thesis, i.e., where he appears to have claimed that believing simply is a feeling.[56] Nevertheless, it is more accurate to attribute to Russell a type of sentimentalism that merely posits some feeling or other (i.e., one of the seven listed above) as necessary for believing, rather than identical to it. This is because typical beliefs were analyzed by him into one of the seven characteristic feelings listed above, but also into a propositional content, and into a relation between those feelings and their *proposition(s)*. Typically, believing has both feeling and content (not to mention whatever facts make them true or false).[57] So officially, we should say that Russell's sentimentalism declares some

[56] E.g., *AMi* (58, 233, 249–50); "On Propositions" (*CP8*, 298–99, 300). See also his "Miscellaneous Notes" [1919], where he defines *memory* as "a specific feeling or sensation, constituting belief about the past. . ." (*CP9*, 20).

[57] See *AMi*, (251); compare also "On Propositions" (*CP8*, 297). As originally reported by James, nitrous oxide intoxication provides an exception. Russell avers, with James, that in cases of intoxication "a man's very soul may sweat with conviction", but Russell adds that this means that *conviction* becomes intense without having any propositional content to which it is attached. He speculates that this phenomenon, wherein a feeling of believing is detached from any content, and perhaps goes hunting for one, are "what passes for revelation or mystic insight" (*AMi*, 252). Russell is quick to qualify this last claim as mere speculation. But we notice *conviction* reappearing here, in 1921. See Russell's discussion of *familiarity* unattached to its typical object, illustrated by Turgenev's *Smoke* (*AMi*, 168–69). Regardless, such cases are supposed to be exceptions.

feeling or other as necessary for believing, but not identical to it.[58] Still in this respect, Russell's sentimentalism grew considerably closer to Hume's, i.e., insofar as it came to posit a tighter (logical) relation between believing and feeling, and not merely the looser (causal) one.

An example might be useful here. Consider one of Russell's own: "there is an egg for breakfast". This same *proposition* can provide the content for either a *memory*, or a tenseless belief, or an *expectation*, depending upon the *attitude* taken toward it, i.e., depending upon the feeling one has in relation to the content. Russell explicitly argues that temporality is frequently to be found in belief's feeling, rather than its propositional content. And he attempted to do justice to the important difference between remembering that one has had an egg for breakfast and merely believing (non-temporally) that one has had an egg for breakfast. Imagine, for example, that you couldn't remember your own breakfast, but that you had a trusted companion (perhaps your spouse) who told you: "it was definitely an egg". Propositional contents may, of course, contain pastness (or futurity) as a part of their content, e.g., when we (non-temporally) *assent* to a *proposition* like, "Caeser conquered Gaul". But that is very different, phenomenologically, from actually remembering that time when "Caeser conquered Gaul". And that, in turn, is very different from remembering that "Caeser conquered Gaul" is true, i.e., from remembering to *assent*. A unique feature of Russell's account here is that *assent* is construed as non-temporal, in opposition to both *memory* and *expectation* which are, for him, inherently temporal feelings. Russell complained that analyses of time are frequently fouled by the fact that natural language does not readily lend itself to a tenseless mode of expression for *assent*.

Nevertheless, those are all happy occasions, i.e., examples of eggs for breakfast and imperial triumphs. Less optimistically, egg-wise, it is possible to doubt that there was/is/will be an egg for breakfast, or simply not believe it, or even downright disbelieve that there was/is/will be an egg for breakfast. None of these more pessimistic *attitudes* should be confused for the others.[59] And Russell quite explicitly argued that

[58] Brentano has his revenge here, because a belief without content is not really a belief in anything, i.e., not really a belief at all. It is merely a sweaty *conviction*.

[59] About the second of them Russell had the following to say: "The state of mind in which we merely consider a proposition, without believing or disbelieving it, will then appear as a sophisticated product, the result of some rival force adding to the image-proposition

a positive feeling which may be called suspense or non-belief—a feeling which may be compared to that of a man about to run a race, waiting for the signal. Such a man, though not moving, is in a very different condition from that of a man quietly at rest" (*CP8*, 300); reprinted identically in *AMi* (248–49). This passage suggests that it is possible to conceive *non-belief* (or "suspense") as a distinctive feeling; so, I would have been remiss had I not included it in the list of seven that I officially attributed to Russell. However, I confess here to having been hesitant to include it, because the passage just quoted comes from a part of Russell's argument where he was criticizing one of James' views rather than presenting his own. See *AMi* (247–50); See "On Propositions" (*CP8*, 300–301). I read Russell as rejecting James' theory that belief is merely an uncontradicted *image*, but maintaining his notion that *non-belief* (or "suspense") is a distinctive feeling. Including *non-belief* in a list of doxastic sentiments has something of the air of paradox: various kinds of feelings of believing include *non-belief*? I can already hear a critic grumbling that this does not make good sense. My response can only be to remind the grumbler that we shouldn't get hung up on terminology. Based on the passage above, perhaps it would be better to think of the particular sentiment as "suspense" rather than *non-belief*? Whatever it is called, it is not meant as our default doxastic *attitude*, but instead the sentimental achievement in having one's *credences* perfectly balanced, akin to ancient *equipollence*. Regardless, if a feeling of *non-believing* doesn't belong amongst the other beliefy feelings, then it seems to me that a feeling of *disbelief* must be excluded too, by similar logic. And Russell clearly endorsed the treatment of *disbelief* as a feeling. Exactly how many feelings of believing are to be found in *AMi*? Answering this question is surprisingly difficult, and involves surprisingly deep interpretive choices. In part this is because Russell's sentimentalism extended beyond the merely doxastic. In places he appears to treat *desire* and *aversion* as feelings too. And here is a rare case where we might be able to find some daylight between "On Propositions" and *AMi*. In 1919 Russell appears (in passing) to classify *desire* as a feeling (*CP8*, 288), but by 1921 he provided a thoroughly behavioristic analysis (*AMi*, 58–76). *Aversion* is closely associated with *pain* or *discomfort*, which are sometimes described as "sensation" See, e.g., *AMi* (68). But one of Russell's main lines of argument in Lecture III was that because the question concerning whether these are "intrinsic qualities of sensations" as opposed to "causal characteristics of [their] occurrences" cannot be settled, it is therefore safer to treat *pain* and *discomfort* as the latter. See *AMi* (68–72). My conclusion is that they should not be regarded as feelings, insofar as Russell thought (at least in 1921) that they could be given behavioristic analysis. He also appeared to treat *anxiety* as a feeling. (Again, only in passing. "On Propositions" (*CP8*, 291); (*AMi*, 199). And he definitely treated *familiarity* as a feeling (*AMi*, 161–63, 168–70, 178, 252). But while Russell may have treated *desire*, *aversion*, and *anxiety* as feelings, and while he definitely treated *familiarity* as one, he classified none of these as "cognitive," and thought (mistakenly, in my view) that all believing/judging is cognitive. So, I have comfortably left those off the list above. See *AMi* (170, 244). In part the interpretive/enumerative difficulty arises because Russell did not propose to limit himself to only seven sentiments. In the chapter on memory he discusses a feeling of the *imaginary*, as opposed to the feeling of *reality*. And while these, it seems to me, (unlike *desire*, *aversion*, *anxiety*) are clearly treated as doxastic, at least in 1919-21, I have not included them on the official list because it is not obvious (at least not to me) that what Russell called "the feeling of reality" is in fact distinct from *assent*. The *reality* feeling and the *imaginary* feeling can also be found in the *Theory of Knowledge* manuscript (1913). See *CP7* (53, 61). But in that work it is also not clear that they were meant to be cognitive.

disbelief is a distinct feeling. That is to say, Russell quite explicitly denied that *disbelief* in *p* could be treated as equivalent to *assent* that not-*p*:

> There is no "not" in an image-proposition; the "not" belongs to the feeling, not to the content of the proposition. An image-proposition may be believed or disbelieved; these are different feelings towards the same content, not the same feeling towards different contents (*CP8*, 304).

Of all of Russell's expressions of sentimentalism, this one is perhaps the most striking. He is suggesting here that something akin to negation (if not logical negation itself) is part of our *attitude* towards a *proposition*, rather than an operation on or relation between propositional contents. Here we must, of course, be quite careful about the context, which was a discussion of *image-propositions.* We must keep in mind the clear cases where negations (i.e., logical negations proper) are part of the content of the *proposition*, i.e., those cases of so-called *word-propositions*. But the *image-propositions* instead merely depict relations amongst the constituents of a picture. And, according to Russel, those pictures may be assented to or disbelieved. And disbelieving was clearly treated by Russell as a distinctive kind of feeling, not to be confused for any modified version of *assent*. This all suggests (to me at least) that *disbelief*, for Russell in the period from 1919 to 1921, amounted to a distinct type

Russell wrote: "I do not mean that the subject, as a rule, definitely *judges* that they belong to sensation, but that his feelings towards them, while he is dreaming, are such as he would usually only have towards objects of sense. . ." (*CP7*, 61). This suggests that the feelings in question were to be distinguished from the judging/believing. There may also be room for phenomenological quibbling here over whether a "feeling of the imaginary" is or is not equivalent to *disbelief* (as I have chosen to treat it). And one presumes that Russell would have been interested in reducing the number of distinctly posited sentiments, wherever possible. But the interpretive/enumerative difficulty also arises because Russell himself eschewed analysis of the sort that I am presently undertaking on his behalf. He claimed in a final note to his memory chapter that "I do not wish to commit myself to any special analysis of the belief feeling" (*AMi*, 187). This same claim is repeated at *AMi* (250). In any event his number of kinds of belief-feeling was (perhaps embarrassingly, if only to me) slightly larger than his official number of kinds of belief, which included only the first three members of my official list: "memory, expectation, and bare assent." See *AMi* (244, 250). Still, one is sorely tempted to analyze, on his behalf, *conviction* and *doubt*, into merely intensified or attenuated degrees of *assent*. For what is *conviction* if not an intensely felt *assent*, and *doubt* if not a feebly felt *assent*? I have elected to resist that temptation here, because it would still only get his number down to four. And at the very least, this sort of reductive maneuvering is impossible for his account of *disbelief*.

of "experiential negation".[60] And that is a striking doctrine, especially to those of us attuned to the history of phenomenology, or to debates over psychologism in the history of logic.[61] What would be the relationship between that sort of "experiential negation", i.e., *disbelief*, and logical negation proper? Unfortunately for posterity, Russell does not appear to have ventured further down this particular rabbit hole.

6. Russell's Reasons for Sentimentalism?

It is clear where Russell picked up his *doxastic sentimentalism*. He got it from Hume's *Treatise* and James' *Principles of Psychology*. It developed out of his theories of memory and imagination.[62] And, after some

[60] There is at least one other passage where he suggests feelings affiliated with logical operations, "There may be other belief-feelings, for example in disjunction and implication;. . ." (*AMi* 250); see "On Propositions" (*CP8*, 299).

[61] Gottlob Frege (1848-1925), at almost exactly this same moment, argues against "two different ways of judging, of which one is used for the affirmative, and the other for the negative. . .", and on grounds that Russell himself would find quite familiar, namely the "*Ersparung an logische Urbestand teilen*", i.e., the "savings in logic's original parts" Frege (1918–1919, 152–55), translated in McGuinness (1984, 383–84). In fairness to Russell, Frege's *thoughts* are more akin to *word-propositions* than *image-propositions*, and it is particularly the latter that require this kind of "negation" that goes beyond the content. See Heidegger's claim that logical negation is intelligible only through an antecedent understanding of the unity of being, elucidated by Shirley (2010, especially 81–90, 141–54).

[62] See the "feeling of pastness" (*CP7*, 173) appealed to as part of his analysis of memory in 1913. At that time Russell analyzed our capacity to experience something as having happened in the past using a definition of "sensation" as an "acquaintance with particulars given as simultaneous with the subject," and argued that because there can't be an intrinsic difference between a present object and a past object, our immediate memory must also be a special kind of relation: "Hence, by our usual criterion, since immediate memory is intrinsically distinguishable from sensation, it follows that it is a different relation between subject and object" (*CP7*, 58, 73, respectively). For discussion of this theory see Apostolova (2017, 307–33, especially 325–27). Apostolova's paper is one of only a very few that explicitly discuss Russell's sentimentalism, particularly in its relation to memory. Germane to my argument, above, could Russell's feeling of *pastness* have been another early doxastic sentiment? While it was certainly a feeling, and while it was certainly important for the theory of memory in 1913, there is an important reason to think that Russell did not treat *pastness* as rising to the level of a belief/judgment. In his theory of memory from the period he was careful to distinguish what he called "perceptive memory" from "judgment memory," arguing that the feeling of *pastness* "is quickly succeeded by the more stable and more easily revived *judgment* of memory" (*CP7*, 173); compare *CP7* (71–72). So, while the feeling in question was supposed to be a precursor to the full-blown judgment, the latter, for Russell, were characteristically treated as "cognitive" rather than merely "perceptive".

reconstruction, I hope it is now also a bit clearer when that happened. Russell espoused the *weakest doxastic sentimentalism*, albeit grudgingly, in his *Theory of Knowledge* manuscript (1913). He endorsed sentimentalism more explicitly in print, and with considerably less reservation, and with a full panoply of doxastic sentiments, in "On Propositions" (1919) and *The Analysis of Mind* (1921). But why he did so, especially in the latter case, may remain a bit murky.

Russell may have thought that a sentimental theory of believing was required on pragmatic grounds, i.e., because of belief's tight connection with action. We recall that in 1918 Russell had claimed that a behaviorist analysis of belief "hangs together" with *neutral monism*. He self-consciously, and after considerable reflection, adopted *neutral monism*. So perhaps he believed that by doing so he must in one fell swoop adopt behaviorism too?

But that cannot have been Russell's reason, because at the very moment he began espousing *Russell's doxastic sentimentalism* (as I've defined it above) he also continued to deny that belief could be defined by its consequences for action.

> I think it must be conceded that a mere image, without the addition of any positive feeling that could be called "belief," is apt to have a certain dynamic power, and in this sense an uncombated image has the force of a belief. But although this may be true, it accounts only for some of the simplest phenomena in the region of belief. It will not, for example, explain memory. Nor can it explain beliefs which do not issue in any proximate action, such as those of mathematics. I conclude, therefore, that there must be belief-feelings of the same order as those of doubt or disbelief, although phenomena closely analogous to those of belief can be produced by mere uncontradicted images (*AMi*, 249–50).[63]

Not only is this an outright rejection of behaviorism, in the reductive sense that I have focused upon above, it is also recognition by Russell that sentimentalism and behaviorism about belief are incompatible, i.e., they are inconsistent theories of believing. The two doctrines are inconsistent because the one treats believing as (or as closely affiliated with)

So, sticking to my methodological principle, i.e., reading Russell's theories of judgment as equivalent to his theories of belief (and vice versa), I have not interpreted this particular feeling as a doxastic one.

[63] This passage is almost identical in "On Propositions" (*CP8*, 300–301), merely omitting reference to *desire* and *images* as signs.

feeling(s), while the other defines believing as nothing but behaviors or behavioral dispositions.[64] Having abandoned the multiple-relation theory of judgment, *pace* MacBride (2013), it was open to Russell to choose either theory of believing, but not both. Faced with that choice, Russell read deeply in the psychology of his day, particularly Watsonian behaviorism, and clearly chose sentimentalism.[65]

But why? We should not overlook the anti-behaviorist reasons he gave for his choice. At least two objections to behaviorism are suggested in the passage quoted immediately above. Both criticisms (of behaviorism) might also be tentatively offered, at least on Russell's behalf, as indirect arguments for sentimentalism. We can reconstruct the first from his reference to *memory* in the third sentence of the passage. According to Russell a reductive behavioral analysis cannot account for the difference between remembering a past content and merely assenting to that past content. For example, remembering that one had an egg for breakfast is not the same, at least not phenomenologically, to assenting (non-temporally) to the fact that one had an egg for breakfast. Behaviorism can make no sense of that difference. Which is to say, behaviorism cannot account for the distinctive experiential quality of *memory*, which is not merely *assent* to facts about the past.

His second argument (referred to in the quotation above) regards beliefs that do not issue in any "proximate action". Here mathematical beliefs again make an appearance. Recall that in 1913 Russell appealed to multiplication tables to illustrate a certainty that lacked *conviction* (because there are certainties that admit of no possibility of *doubt*). Here mathematical beliefs play spoiler in a quite different argument: Russell thought that some mathematical beliefs are sufficiently abstract as to make no difference for action. Because standard arithmetical beliefs are too practical to serve as examples here, instead consider something romantic, like believing Fermat's Last Theorem. It is not at all obvious

[64] It is slightly misleading to write "*intrinsic* nature of believing," because on Russell's account a belief is an event that may someday be used in a physical *construction*, and not merely in mental ones, i.e., it could be just as much a part of the one kind of "logical fiction" as the other. *Intrinsically*, they are spacetime events, causally ordered (according to the best sciences of our day) using only the (i.e., up until now) psychological laws.

[65] On this point I remain unreconciled with Misak (2016), who understands Russell's theory of belief to be broadly pragmatist. "But while Russell may have become a pragmatist about belief, he has not become a pragmatist about truth" (2016, 149).

how believing there are no positive integers a, b, c, n that satisfy the equation $a^n + b^n = c^n$ for any value of n greater than 2, would have a practical significance for one's everyday life.[66] Nevertheless, we believe Fermat's Last Theorem. Which is to say, behaviorism cannot account for some (abstract, mathematical) beliefs.

In addition to those two arguments, Russell also cited an argument originally made by Brentano (*AMi*, 247).[67] There is a clear difference between believing that there will be (i.e., *expecting*) an egg for breakfast and merely entertaining the idea that there will be an egg for breakfast. Such a difference cannot lie in the content of the two *attitudes* (which is the same in both cases.) So, if behaviorism is true, then that difference must lie instead in egg-behaviors or dispositions to egg-behaviors, present in the one case but not the other. But any difference in effect must mean that there was some difference in cause. So, there must have been some *intrinsic* difference between the two *attitudes*, i.e., whatever was necessary to cause the behavioral differences. But if behaviorism is true, then believing is nothing but the behaviors or dispositions to behaviors (i.e., it is not any *intrinsic* difference between the two *attitudes*). Therefore, behaviorism is false.

In the manuscript notes Russell used for the preparation of *AMi*, he cites Brentano, applying this last argument to behaviorism in the sense defined above, and writes "This objection is fatal" (*CP9*, 9). In the published version of "On Propositions" Russell softens that claim and instead prints the argument with the remark that it "seems fatal" (*CP8*, 298). And by the time *AMi* itself is printed, while this argument remains prominent, it is instead merely announced as "another, more theoretical, ground for rejecting the view…", i.e., without any comment about its degree of lethality (*AMi*, 247). Russell continues to offer it, but appears to grow increasingly cautious about how deadly it may be.

Of the three arguments, I myself think the first is strongest. As lethal as Brentano's argument or the argument about abstract mathematical beliefs may (or may not) be, what they really force upon *reductive behaviorism about belief* is not a death so much as a transformative

[66] Assuming the everyday life is not the everyday life of a mathematician! (See my criticism of this argument immediately below.) Clearly Andrew Wiles' belief of Fermat's Last Theorem had some practical consequences.

[67] See also Brentano (1959, 40–41); translated by Rancurello et. al. as Brentano (1995, 202–3).

choice. Recall that we defined the doctrine disjunctively: "all believing is reducible to either an action or a disposition to an action". So now, under mortal threat, disjoin them. Call the claim that all believing is reducible to an actual action, "crude reductive behaviorism about belief". And call the claim that all believing is reducible to a disposition, "reductive dispositionalism about belief". Brentano's argument might kill *crude reductive behaviorism about belief*, but it proves far less fatal to *reductive dispositionalism*.[68] Distinguishing actions from dispositions can also satisfy the Brentanian argument by preserving a difference between a belief and its consequences. And with a suitably sophisticated understanding of dispositions (this will almost certainly need to be one that includes merely possible behaviors) a similar sort of response could be given even to Russell's argument about beliefs that supposedly make no difference for action. The dispositionalist argues that they do make a difference, if only for some possible actions. Whether or not behaviorism was killed, we must at least acknowledge that, in its functionalist and dispositionalist guises, it remained quite lively over the course of the 20th Century.

I should also emphasize, here in conclusion, that while these arguments were provided directly against (one type of) behaviorism, they offer only indirect support for Russell's sentimentalism. I speculate that Russell's more direct evidence was phenomenological, i.e., a structure in his own experience of *memory*.[69] Obviously, for a philosopher like Russell, an appeal to introspection could never be sufficient. Still, I suspect it played a role here, given his hesitance regarding behaviorism. I suspect it reports one of the "facts" that he thought must be accounted for. There are, of course, many other possible theories of belief, neither sentimentalist nor behaviorist. Direct argument against one can only ever be (at most) indirect evidence for another. It's equally important to

[68] For a more sophisticated argument against *reductive dispositionalism*, see Collins (1999, 75–88).

[69] And here one final "thank you" to the reviewers, for pushing me on this point too. Rejecting behaviorism, in the narrowly defined sense, still leaves open a variety of alternative doxastic theories. In claiming that the first of Russell's three arguments was the strongest I have been suggesting that his phenomenological evidence may have been his true motive. And as one of the reviewers has pointed out, following my line of argument to a clearer conclusion than I was able to articulate: "Russell's other arguments likely provide *supplemental* support which is needed due to the fact that introspection is limited, fallible, and susceptible to pretheoretical bias".

recognize that at this particular historical juncture there were a variety of avenues the new psychology could have taken, some neither Jamesian nor Watsonian. We should not forget that Russell adopted sentimentalism, self-consciously, from James and rejected behaviorism as he found it vocalized (not to mention subvocalized) by Watson. Russell's arguments were mustered against what he took to be, at one moment in time, sentimentalism's chief rival.

And none of this was Russell's final word on the relationship of believing to behavior. By the time he published *The Analysis of Matter* (1927), his understanding appears to have modulated yet again.

> Beliefs, in the psychological sense, seem to emerge out of previously existing habits, and to be, at first, little more than verbal representations of habits formed before words could be uttered (Russell 1954, 142).

While this may seem a step closer to behaviorism, I would caution that even this step would not be the full distance. The claim articulated here in 1927 remains entirely consistent with Russell's earlier treatment of the simplest form of believing as sentimental and *image*-based, i.e., entirely consistent with his rejection of *reductive behaviorism about belief* (defined in either of the ways disjoined above).[70] *AMi* clearly gives a broadly naturalist story about the development of believing. Though I have not emphasized it above, in my own retelling, Russell's Humeanism made liberal appeal to "habit". And that may, indeed, amount to a kind of "behaviorism". But it is not the behaviorism that I have been discussing in this paper. Taking the wider and more synoptic view now, in conclusion, I think it would be a mistake to read Russell's claims about habit as in opposition to, let alone some sort of retraction of, his sentimentalist doctrines. (Just as it would be a mistake to read them as if they were opposed in the writings of someone like Hume.) To say that beliefs "emerge" out of habits, or that they are habits' "verbal representations", is to posit a developmental relationship between them, if not a strictly causal one. In the spirit of the once-thought-"fatal" Brentanian argument: to say that beliefs "emerge" from previously

[70] More problematic (for me) is Russell's 1926 entry "Knowledge, Theory of" in the *Encyclopedia Britannica*, reprinted in *CP9* (194–202). The theory of belief in that entry shows no trace of sentimentalism. It is completely behaviorist, in precisely the sense with which I have been concerned. I am not at all sure what to make of it, honestly.

existing habits is to commit oneself to beliefs not being identical to previously existing habits.

7. Conclusions

Regardless, we are now finally in a position to appraise the relationship between *Russell's doxastic sentimentalism* and his *neutral monism.* That relationship is anfractuous, historically. In the second decade of the 20th Century Russell's attitudes about each, not to mention his understandings of the contents of each, were shifting, and shifting independently of one another. Nevertheless, the two doctrines clearly had a common source, i.e., the inspiration of William James. My hope is that their relation, theoretically, has been brought into somewhat further focus with the aid of the analyses above: at the very least I hope to have shown how they may be interpreted as consistent with one another.[71]

While Russell adopted sentimentalism at exactly the moment he posited *images* as contents for the simplest forms of believing, and while those *images* were posited by him as "neutral" in their *intrinsic* nature, they need not have been posited as neutral in their *intrinsic* nature. Imagine instead, that the ontology of Russell's *Analysis of Mind* (1921) had remained wholly dualist. In that case much that Russell argued about believing would, of necessity, have been different. Beliefs in that case would not have been treated as neutral spacetime events, i.e., they would not have been party to *constructions* of mental or physical entities. But much of his theory of believing, e.g., 243–52 of *AMi*, could have remained mostly intact. The feelings and the *images* would in that case have been construed as irreducibly mental. But the "Brentanian" schema

[71] On this point, see Tully (2003): "In retrospect it seems clear that, even when he had finished *The Analysis of Mind,* Russell had not yet worked out a satisfactory analysis of belief, along with a closely related theory of propositions, that would clearly fit within the existing framework of Neutral Monism. Although"On Propositions" officially disowns his older multiple relation theory of judgment, according to which propositions are parts of actual judgments, the replacement theory he offered both there and in *The Analysis of Mind* seems bizarre, and in view of the constraints he was putting on himself, it probably was bound to be" (Tully 2003, 353). Also see again Tully (1988), where he writes that the introduction of feelings, "appears to contaminate, rather than merely complicate, the thesis of neutral monism which Russell had adopted" (Tully 1988, 209–24, especially 218–19). Setting aside whether Russell's theory of believing was "bizarre", or "contaminate[d]", I hope to have at least shown how the two doctrines "clearly fit".

itself, and the analysis of believing as a feeling in relation to a *proposition*, and the seven sentiments, could all have remained the same.

Contrariwise, to be a *neutral monist* requires only that particular episodes of believing, and those actions resulting from them, be treated as events, and that any physical or mental "stuff", i.e., any supposedly physical or mental entities, be treated as complexes built out of elements that are themselves neither mental nor physical. Believing need not be treated as requiring any relation whatsoever between a feeling and some content, or as requiring any feeling whatsoever. (Presuming, as we typically do, that not everything mental need be a feeling.) A philosopher other than Russell could have just as easily posited belief as an affectless *attitude*, or as a behavioral disposition, so long as those other sorts of mental state were conceived consistently with the underlying metaphysics.[72]

This is to say that *Russell's neutral monism*, as defined above, is consistent with either *Russell's doxastic sentimentalism* or *reductive behaviorism about belief*, as defined above. It is consistent with many other theories of belief as well. Perhaps we must choose whether our ultimate theory of believing will be behaviorist or not (as Russell thought we must), but neither behaviorism nor sentimentalism are necessitated by the underlying metaphysics. So, neither theoretically nor temporally are the doctrines of *Russell's neutral monism* and *Russell's doxastic sentimentalism* connected. They are connected only genealogically, i.e., only in the sense that they both developed under the influence of James. This is my secondary conclusion yet again (as it was prematurely at the end of Section 2). Russell's conversion to *neutral monism* happened in late 1918, during his study of psychology in the Brixton Prison. But he had already expressed, albeit grudgingly and only in its weakest form, *doxastic sentimentalism* in his *Theory of Knowledge* manuscript of 1913.

Doxastic sentimentalism was an element of James' psychology that Russell could have considered more closely, given that it is not "inevitable". Russell's conversion to *neutral monism* happened under the pressures of significant philosophical problems. It was the result of considerable reading, research, thinking, and argument. His adoption of James' sen-

[72] A true sentimentalist would probably deny that there could be such a thing as an "affectless attitude". For a *thoroughgoing sentimentalist*, i.e., one who treats not only believing but all consciously experienced mental states as feelings, an "affectless attitude" is a contradiction in terms.

timentalism, by contrast, happened (somewhat like feelings themselves) in the background; by comparison it was startlingly under-considered and under-rationalized. Still, as I have attempted to document above, it was not entirely invisible and unreasoned: it is possible to reconstruct at least three Russellian arguments against behaviorism. None of those provide direct evidence for sentimentalism, and even Brentano's "fatal" argument remains controvertible. Still, we should not hold the finitude of Russell's attention against him, however we might wish it had been directed. He is one of our tradition's greatest thinkers. We should follow him, as in so many other ways, by trying to further discover his reasons.

Acknowledgements

I gratefully acknowledge Nick Denning, Debra Jean Harris, and Ivan Welty, who read early drafts of this paper and were characteristically thoughtful and helpful in reply. The two anonymous reviewers of this manuscript for the *Journal for the History of Analytical Philosophy* were the best that I have ever had in my academic career.

References

Alter, Torin, and Yujin Nagasawa, eds. 2015. *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford: Oxford University Press.

Apostolova, Iva. 2017. "Russell's Two Theories of Memory." *Russell: The Journal of Bertrand Russell Studies* 37: 307–33.

Baldwin, Thomas. 2003. "From Knowledge by Acquaintance to Knowledge by Causation." In *The Cambridge Companion to Bertrand Russell*, edited by Nicholas Griffin, 420–48. Cambridge: Cambridge University Press.

———. 2018. "Truth in British Idealism and Its Analytic Critics." In *The Oxford Handbook of Truth*, edited by Michael Glanzberg, 125–49. Oxford: Oxford University Press.

Banks, Eric C. 2014. *The Realistic Empiricism of Mach, James, and Russell*. Cambridge: Cambridge University Press.

Bostock, David. 2012. *Russell's Logical Atomism*. Oxford: Oxford University Press.

Brentano, Franz. 1959. *Psychologie Vom Empirischen Standpunkt, Bd. II [1874]*. Hamburg: Felix Meiner Verlag.

———. 1995. *Psychology from An Empirical Standpoint*. Translated by Antos C. Rancurello, D. B. Terrell, and Linda L. McAlister. New York: Routledge.

Collins, Arthur W. 1999. "Behaviorism and Belief." *Annals of Pure and Applied Logic* 96: 75–88.
Engelmann, Mauro L. 2012. "Wittgenstein's New Method and Russell's The Analysis of Mind." *Journal of Philosophical Research* 37: 283–311.
Frege, Gottlob. 1918–1919. "Die Verneinung. Beiträge Zur Philosophie Des Deutschen Idealismus I," 152–55.
Glanzberg, Michael. 2018. *The Oxford Handbook of Truth*. Oxford: Oxford University Press.
Griffin, Nicholas. 2003. *The Cambridge Companion to Bertrand Russell*. Cambridge: Cambridge University Press.
Hickerson, Ryan. 2020. *Feelings of Believing: Psychology, History, Phenomenology*. Lanham, MD: Lexington Books.
Hochberg, Herbert. 2000. "Propositions, Truth, and Belief: The Wittgenstein-Russell Dispute." *Theoria* 63: 3–40.
Holman, Emmett L. 2008. "Panpsychism, Physicalism, Neutral Monism and the Russellian Theory of Mind." *Journal of Consciousness Studies* 15: 48–67.
Humberstone, I. L. 1996. "Intrinsic/Extrinsic." *Synthese* 108: 205–67.
Hume, David. 2000. *A Treatise of Human Nature. [1739-40]*. Oxford: Oxford University Press.
James, William. 1981. *The Principles of Psychology, Vol. II. [1890]*. Cambridge, MA: Harvard University Press.
Koç Maclean, Gülberk. 2014. *Bertrand Russell's Bundle Theory of Particulars*. London: Bloomsbury.
Landini, Gregory. 2011. *Russell*. London: Routledge.
Linsky, Bernard. 1999. *Russell's Metaphysical Logic*. Stanford, CA: CSLI Publications.
MacBride, Fraser. 2013. "The Russell--Wittgenstein Dispute: A New Perspective." In *Judgement and Truth in Early Analytic Philosophy and Phenomenology*, edited by Mark Textor, 206–41. London: Palgrave Macmillan.
McGuinness, Brian. 1984. *Gottlob Frege: Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Basil Blackwell.
Misak, Cheryl. 2016. *Cambridge Pragmatism: From Peirce and James to Ramsey and Wittgenstein*. Oxford: Oxford University Press.
———. 2019. "Russell and the Pragmatists." In *The Bloomsbury Companion to Bertrand Russell*, edited by Russell Wahl, 59–74. London: Bloomsbury.
Owen, David. 2000. *Hume's Reason*. Oxford: Oxford University Press.
Pears, David. 1975. *Questions in the Philosophy of Mind*. New York: Harper \& Row.
———. 1989. "Russell's 1913 Theory of Knowledge Manuscript." In *Rereading Russell: Essays in Bertrand Russell's Metaphysics and Epistemology*, edited by C.

Wade Savage and C. Anthony Anderson, 169–82. Minneapolis: University of Minnesota Press.

Pincock, Christopher. 2019. "Neutral Monism." In *The Bloomsbury Companion to Bertrand Russell*, edited by Russell Wahl, 312–33. London: Bloomsbury.

Ramsden Eames, Elizabeth, and Kenneth Blackwell, eds. 1984. *The Collected Papers of Bertrand Russell, Vol. 7: Theory of Knowledge: The 1913 Manuscript*. London: George Allen \& Unwin. Abbreviated as CP7.

Russell, Bertrand. 1914. "N the Nature of Acquaintance II. Neutral Monism." *The Monist* 24: 186–87.

———. 1919. "On Propositions: What They Are and How They Mean." *Proceedings of the Aristotelian Society Supplementary* 2: 1–43.

———. 1921. *The Analysis of Mind*. London: George Allen \& Unwin.

———. 1922. "Dr. Schiller's Analysis of The Analysis of Mind." *The Journal of Philosophy* XIX: 645–51.

———. 1954. *The Analysis of Matter [1927]*. New York: Dover Publications.

———. 1959. *My Philosophical Development*. London: George Allen \& Unwin.

———. 1968. *The Autobiography of Bertrand Russell: 1914--1944*. Boston: Little, Brown and Company.

———. 1995. *The Analysis of Mind [1921]*. London: Routledge. Abbreviated AMi.

Savage, C. Wade, and C. Anthony Anderson, eds. 1989. *Rereading Russell: Essays in Bertrand Russell's Metaphysics and Epistemology*. Minneapolis: University of Minnesota Press.

Shirley, Greg. 2010. *Heidegger and Logic: The Place of Lógos in Being and Time*. New York: Continuum.

Slater, John G., ed. 1986. *The Collected Papers of Bertrand Russell, Vol. 8: The Philosophy of Logical Atomism and Other Essays: 1914-19*. London: Unwin Hyman. Abbreviated as CP8.

———, ed. 1988. *The Collected Papers of Bertrand Russell, Vol. 9: Essays on Language, Mind and Matter 1919-26*. London: Unwin Hyman. Abbreviated as CP9.

Stubenberg, Leopold. 2015. "Russell, Russellian Monism, and Panpsychism." In *Consciousness in the Physical World: Perspectives on Russellian Monism*, edited by Torin Alter and Yujin Nagasawa, 58–90. Oxford: Oxford University Press.

———. 2018. "Neutral Monism." Edited by Edward N. Zalta. The Stanford Encyclopedia of Philosophy. 2018. https://plato.stanford.edu/archives/fall2018/entries/neutral-monism/.

Stubenberg, Leopold, and Donovan Wishon. 2023. "Neutral Monism." Edited by Edward N. Zalta and Uri Nodelman. The Stanford Encyclopedia of Philosophy, Spring 2023 Edition. 2023. https://plato.stanford.edu/archives/spr2023/entries/neutral-monism/.

Textor, Mark. 2013. *Judgement and Truth in Early Analytic Philosophy and Phenomenology*. London: Palgrave Macmillan.

Tully, Robert. 1988. "Russell's Neutral Monism." *Russell: The Journal of Bertrand Russell Studies* 8: 209–24.

———. 2003. "Russell's Neutral Monism." In *The Cambridge Companion to Bertrand Russell*, edited by Nicholas Griffin, 332–70. Cambridge: Cambridge University Press.

Wahl, Russell, ed. 2019. *The Bloomsbury Companion to Bertrand Russell*. London: Bloomsbury.

Ward, James. 1906. *Naturalism and Agnosticism: The Gifford Lectures Delivered Before the University of Aberdeen in the Years 1896--1898. [1899]*. London: Adam and Charles Black.

Whitehead, Alfred North, and Bertrand Russell. 1925. *Principia Mathematica, Vol. I. Second Edition*. Cambridge: Cambridge University Press.

Wishon, Donovan. 2015. "Russell on Russellian Monism." In *Consciousness in the Physical World: Perspectives on Russellian Monism*, edited by Torin Alter and Yujin Nagasawa, 91–120. Oxford: Oxford University Press.

———. 2021. "Radical Empiricism, Neutral Monism, and the Elements of Mind." *The Monist* 104: 125–51.

Journal for the History of Analytical Philosophy

VOLUME 12, NUMBER 6 (2024)

ISSN 2159-0303

jhaponline.org