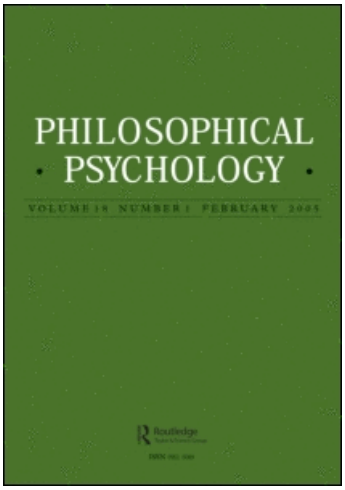

Philosophical Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713441835

Neutral monism reconsidered

Erik C. Banks

Online publication date: 23 April 2010

To cite this Article Banks, Erik C.(2010) 'Neutral monism reconsidered', Philosophical Psychology, 23: 2, 173 — 187
To link to this Article: DOI: 10.1080/09515081003690418
URL: http://dx.doi.org/10.1080/09515081003690418

Philosophical Psychology
Vol. 23, No. 2, April 2010, 173–187

Neutral monism reconsidered

Erik C. Banks

Neutral monism is a position in metaphysics defended by Mach, James, and Russell in the early twentieth century. It holds that minds and physical objects are essentially two different orderings of the same underlying neutral elements of nature. This paper sets out some of the central concepts, theses and the historical background of ideas that inform this doctrine of elements. The discussion begins with the classic neutral monism of Mach, James, and Russell in the first part of the paper, then considers recent neo-Russellian versions in the second half. The chances for a revival of neutral monism are probably slight; its key ideas and starting points lie far from those in contemporary philosophy of mind. A better route might be through the philosophy of science and a deeper understanding of causation.

Keywords: Causation; Mach; Monism; Neutral; Russell; Wide

1. Introduction

Neutral monism is a position in metaphysics and philosophy of mind, the classic version of which was defended by Ernst Mach, William James, and Bertrand Russell (Banks, 2003).[1] Neutral monism arose in a climate of revolt against the 'billiard ball' mechanical philosophy of nature and paralleled the rise of a scientific psychology in the late nineteenth century. Mach's original theory of elements was developed as part of a project to eliminate unnecessary metaphysical notions of 'mind' vs. 'body' and to unite physics and psychology in a general scientific view which could then be applied in any field of inquiry. The elements that made up minds and bodies were in themselves *neither* mental sensation qualities, *nor* physical objects in space and time, but rather neutral between both domains and distinguished provisionally by isolating out one kind of variation of elements as 'mental' and one kind of variation as 'physical'. In reality, there was no such dualism, either of elements or variations (Mach, 1959, pp. 16, 310–311; see also James, 1977; Russell, 1914/1984, 1921, 1927/1959).

William James developed neutral monism further into his own direct realist epistemology in the *Essays on Radical Empiricism* (1977). Russell, originally an

Erik C. Banks is Associate Professor of Philosophy at Wright State University.
Correspondence to: Erik C. Banks, 384 Millett Hall, Wright State University, 3640 Colonel Glenn Hwy, Dayton OH 45435, USA. Email: erikb2000@netzero.net

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/10/020173-15 © 2010 Taylor & Francis
DOI: 10.1080/09515081003690418

opponent of Mach and James, in a delightful twist, abandoned his theory of acquaintance and converted to the view himself by 1918 (Hatfield, 2002, p. 205) and used the elements—sometimes under the name of 'event-particulars' (p. 208)—for a construction of space and objects from linked perspectives, where perspectives were obtained by connecting elements to each other via their causal relations. On Russell's view, objects could be expressed as sums of their effects or interactions with observers and with other objects. The object is the sum of the event-particulars in its causal history, not a substratum in which intrinsic properties of the thing are grounded, a view very much in line with Mach's conception of an object as a stable function of elements:[2]

> According to the view that I am suggesting a physical object or piece of matter is a collection of all those correlated particulars which common sense would regard as its effects or appearances in different places. (Russell, 1921, p. 101)

As Michael Heidelberger (2004, 2007) has shown, Mach's theory of the elements emerged as a reaction to his mentor G.T. Fechner, the founder of psychophysics. Fechner had urged that sensations and the physical brain were to be taken as the ''inner'' and ''outer'' aspects of a single reality, like the mutually exclusive aspects of a penny, whose two faces mutually exclude one another from view exactly because they are two sides of the same object. Mach, however, objected to this lingering dualism of ''inner'' and ''outer'' aspects and was thus led to one kind of elements:

> There is no rift between the psychical and the physical, no inside and outside, no ''sensation'' to which an external ''thing'' different from sensation, corresponds. There is but one kind of element, out of which this supposed inside and outside are formed.[3]

Mach also worked in a climate of revolt against the seventeenth-century mechanical view of nature (Einstein, 1949) and the object-property ontology lurking in its background. According to the object-property view, the object is primary; it grounds the thing's intrinsic properties, which are then expressed in different situations as the object's dispositions to behave, or appear, this way or that. A thing's relational dispositions are its qualities, which, because they vary from situation to situation, are considered less fundamental than the object with its intrinsic properties. As an empiricist, however, Mach regarded the dispositional qualities as not only more fundamentally known, but also more fundamental in being than either the object or its so-called intrinsic properties. Objects became functions of their experimentally accessible dispositions and nothing more: there was no inaccessible substratum and no intrinsic properties, except insofar as some dispositions were more stable than others. Prime suspect: seventeenth century 'billiard ball' atoms, which for Mach were not atoms as we understand them today, but idealized, indivisible things in themselves incapable of interaction with the outside world (Banks, 2003). Mach's ideal for physics, then, was to replace the object-property ontology with a world of powers, or world elements, linked to each other in functional relations, a power-ontology which Mach believed held all the way down to a basic structure of nature even more fundamental than atoms (Banks, 2003).

With the object disposed of, the conscious ego was slated for elimination next. Rather than thinking of sensation/elements as defined by the subject who apprehended them, elements were concrete realities in their own right, existing independently of the act of perceiving them. (According to Hatfield, 2002, pp. 205–214, this strong stand urged Russell away from Brentano's act-content psychology, towards the functionalist psychology of Mach and James.) Sensations are thus not modes of appearance of an object to a subject, not intrinsically subjective items arising from an act of representation (see Harman, 1990). Instead the subject, like the object, is only a provisionally coherent construction out of elements and functional variations which may be of great complexity, giving the illusion that they are bound to a substratum, when no such entity exists except as a functional connection.

> If a knowledge of the connexion of the elements (sensations) does not suffice us, and we ask, *Who* possesses this connexion of sensations, *Who* experiences it? then we have succumbed to the old habit of subsuming every element (every sensation) under some unanalyzed complex, and we are falling back imperceptibly upon an older, lower, and more limited point of view. (Mach, 1959, p. 26)

Mach wrote that *das Ich ist unrettbar*, literally 'unsalvageable', and James, in the title of his 1904 article, asked "Does Consciousness Exist?" with the implicit answer: no.

2. What are Elements? Four Theses

This historical climate of ideas above left its stamp on the following theses about the nature of elements. I do not say that these theses are articulated by every classic neutral monist, but they are implicit in all three of them:

1. *Monism*: the mental and physical domains are part of a greater natural domain of elements and their functional variations.
2. *Neutralism*: elements are neither mental nor physical; rather minds and physical bodies are functionally related complexes of elements. Certain functional variations of the elements are called 'mental' and others are called 'physical', but there is no underlying duality of variations.[4]
3. *Psychophysical Identity*: every sensation, such as green, is also a physical element (s/e), a neural energy in the brain. Not every element is a sensation, or even a possible sensation.[5]
4. *Powers*: elements are powers with causal force. They are concrete qualities *and* dispositional ways of affecting things in their various causal, or functional, roles. The concrete quality instantiates the dispositional, relational role. Every element is naturally embedded in its functional role.[6]

There are numerous differences between this view and traditional views of sensations as epiphenomenal qualia, secondary qualities or inert or passive sense data. Neutral monist elements pack a punch, following from the idea that a natural quality is always a way of affecting or being affected by something (except here the 'ways' of affecting are more fundamental than the 'somethings'). Mach's elements

push, pull and impede one another in dynamical complexes and can only be isolated from these complexes provisionally (Banks, 2003). We don't start from a basic set, as in logical constructions à la Carnap, but rather isolate an element from a complex empirically, by holding it constant against the variations of the other elements that always accompany it. Reciprocally, a complex is simply a function of its elements and their variations. As Mach often expressed it, if the elements $(\alpha, \beta, \gamma, \ldots, \omega)$ change, there may be functions $f(\alpha, \beta, \gamma, \ldots, \omega)$, $g(\alpha, \beta, \gamma, \ldots, \omega)$, $h(\alpha, \beta, \gamma, \ldots, \omega)$ that do not change and which can represent invariant 'islets' such as egos and objects of perception. By contrast, qualia have no powers, are identified by the subjective criterion of "what it's like" to experience them, and are grouped into similarity classes by the subject's ability to discriminate them (see Clark, 1993). They are widely assumed to lack all effect on physical chains of events, giving them a dubious status within science and inviting elimination.[7]

As Mach and Russell emphasized, however, a sensation of green is identical with a certain neural energy of the brain (which may be very complex). The look of this agglomeration of neurons, and their collective neural energy, to the brain's possessor *is* the sensation quality. This is what Russell meant when he said that "what is experienced may itself be part of the physical world and often is so" (Russell, 1914/1984). We can thus represent the sensation of green as either a sensation quality or a physical element of nature in its own right participating in physical events and variations (**e**). In itself it may be considered either way (**s**/**e**) depending upon whether we want to talk about it in its psychical or its physical variations. One might also note in passing that sensation qualities would certainly pack a punch if they were identical with neural energies, which participate fully in physical events and cause physical changes.

Very well, but why do sensory qualities and the neural energies of populations of neurons look so different, so much so that some philosophers claim on purely *conceptual* grounds, by way of thought experiments, that they could never be the same thing?[8] When we investigate the brain we find very complexly specialized and cross wired cells with their electro-chemical connections, but nothing that looks at all like green. Mach's answer to this ran as follows:

> When I see a green leaf (an event which is conditioned by certain brain processes) the leaf is of course different in form and color from the forms and colors which I discover in investigating the brain, although all forms and colors are of like nature in themselves, being neither psychical nor physical. The leaf considered as dependent on the brain-process is something psychical, while this brain itself represents something physical. (Mach, 1959, p. 62)

Russell also pointed out that the experience of neural energy to the brain itself was, and should be, different from the experience of what it's like to investigate that brain indirectly, and externally, with measuring devices or surgery. What looks like a sensation to the possessor looks like a tangle of neurons to an external observer. There is nothing special about this, of course, *all* natural events are different when investigated externally from what they are in themselves, it is just that only in this one

case of the brain do we actually have access to both points of view. It must be admitted that this response still seems inadequate. It tables the problem without explaining why this interior aspect of neural energy is, or perhaps only seems to be, *so very different* from its exterior aspect (McGinn, 1991; Papineau, 2002). It doesn't seem possible that a cluster of cells carrying out complex electrochemical signaling *could* look like green even from the inside, so this is where explanations should start, not finish. We'll take this up again below.

3. Causal Maps, Causal Closure, Micro-vs-Macro Causation

For all of the classic neutral monists, if we are to represent all of the facts of perception, we employ a device which I will call a 'causal map' of the sensation/elements as well as elements which are not anyone's sensations. Mach calls it a 'functional presentation' (Mach, 1959, p. 27); James calls it simply 'space', and Russell calls it a 'system of perspectives' (see Banks, 2003, chapter 9). We assign each element and each element-sensation a place in a multidimensional graph. We then represent all functional/causal relations as arrows on a graph connecting elements at the vertices. The sensation/elements like green are represented *alongside* (not inside[9]) the physical elements of the brain process being investigated externally, either in the sensations of an observer or the elements produced by a scan. These elements of the brain process in an observer or measuring device are external representations of the brain, while the neural energy of the brain process, as it is in itself, is represented by either the green qua sensation, or qua neural energy of the cluster, if we want to emphasize its physical connections to the elements of the brain process.

Notice that a causal map does not prejudge the question of exactly what the causal connection is between the elements of the external brain process and the qualities of the sensation/elements causally annexed to them. They could be one-to-one, one neuron to one quality, which seems unlikely, or polyadic, one sensation annexed to an entire population of neurons in certain structural configurations and firing patterns. This depends upon experiment, but given what is known about the physiology of sensation, it seems more likely that a polyadic configuration of neurons would correspond better to what we call a given sensation.[10]

But now a second issue arises for the causal map representation. If sensation/elements possess causal force and directly affect physical chains of events, does this not violate the principle of the causal closure of the physical,[11] long recognized by philosophers and physicists alike? As Edwin Boring once characterized the problem:

> The brain, being material, must be regarded as a closed system: whatever energy is delivered to it must be given up again in motion or lost in heat. There seemed, indeed, to be no way of breaking such a system open to introduce mental links . . . . Mental events might parallel its action, but could not cause or be caused by it. (Boring, 1942, p. 86)

As often happens, the trouble is with the understanding of exactly what physics does and does not say about causal closure. Mach responds to the point in the

Analysis of Sensations (1959), speaking as a physicist versed in the interpretation of physical laws (in his famous *Mechanics* for example):

> I cannot here refrain from expressing my surprise that the principle of the conservation of energy has so often been dragged in in connexion with the question of whether there is a special psychical agent. On the assumption that energy is constant, the course of physical processes is limited but not necessarily determined with perfect uniqueness. That the principle of conservation of energy is satisfied in all physiological cases merely tells us that the psyche neither uses up work, nor performs it. For all that, the psyche may still be a partly determinant factor. When a philosopher asks the question . . . he usually misses the point of the principle of the conservation of energy and the stock reply of the physicist has no intelligible meaning in a case so far removed from the scope of his ideas. (Mach, 1959, p. 55n)

A principle like the conservation of energy is a *general phenomenological law*, a blanket statement, about the initial and final states of a process in a closed system. It is silent about mechanisms, does not say *how* the process runs, nor does it prevent us from interpolating another process between the initial and final states, so long as there is no net energy absorbed or contributed by it. Mach also addressed this point in his *Principle of the Conservation of Energy* (1911) remarking that the closure of a system says nothing about its *specific* transformations from state to state, configuration to configuration, only about the energy going in and going out.[12] He points out in the earlier book that little islets or closed systems of activity "shut up in themselves" might occur in an overall system (like the brain) and the conservation law says nothing about how closed systems form, or what effects such formations might have. If, for example, the sensations or behavior of an organism supervenes not on the energy per se but on these intermediary closed islets, configurations of neural energies, specific changes not predicted by, or described by, energy data, could make a recognizable difference to the behavior of the organism and still be in accord with the principle.

But if by 'causal closure' we mean something *other* than a blanket conservation principle, this is where we can get into trouble, for there is really no agreed upon definition of cause (Hausman, 1998). 'Cause' may refer to micro-causation *only*, effecting change according to fundamental laws of physics (Kim, 1998). Or 'cause' may also refer to 'wide' causation or macro-causation, i.e., the effect of macro-types of configuration on each other and on micro-events. Notions of macro-causation, variously defined, are available in Dretske (1993), Garfinkel (1990), Lockwood (2007), Rosenberg (2004), Yablo (1992, 2003, 2009), and Yoo (2006). The effect of causation by macro-types is not to cause micro-events directly and one to one, but to limit micro-caused events in these configurations to other broad type-classes of configuration, or to exclude them. Properties of configurations of neurons could include macro-symmetry or asymmetry, spatio-temporal patterns or islets of firing bounded by inhibitory areas, anatomical location in the brain, and relations to neighboring islets of neural activity. An example of wide causation of individual neural events might be the way that structural features of a configuration constrain the firing of individual neurons by forcing them to accommodate

the firing of neighboring cells in a pattern. Macro-causation does not make the effect happen in the same way that the micro-mechanisms do; a configuration does not 'fire', individual cells do. But the range of micro-physically caused events may be limited by being grouped into certain types of configuration and not others, which may be just as important a feature of neural events, especially if sensations or behavior supervene on the macro-types of configuration in which firings occur, and any higher order laws taking one macro-class of configurations to another macro-class.

The opponent of macro-causation points out that configurations of micro-states *are* already included in a complete micro-causal description of events, so, if the causal powers of configurations are produced by a constellation of micro-factors, the factors and the combination are already included in the micro-description. This is true, the principle of the causal closure of the micro-physical covers all factors and all micro-configurations—configurations of things micro-physically described—in a blanket statement, but for that very reason it is not capable of specifying laws governing the formation of macro-types of configuration (see Banks, 2009). This for the familiar reason that macro-types of configuration may not be realized by a single recognizable type of micro-configuration. Micro-causation is a general physical description of *all* micro-states in *all* possible micro-configurations: all are causally closed, all are the result of micro-factors acting together in concert in configurations. Macro-causation, however, describes how *some* class of macro-configurations *A* either fosters, limits, or rules out another class of macro-configurations *B*. This is more specific information not contained in the micro-descriptions of either events or configurations micro-physically described, hence macro-causation is not redundant, or epiphenomenal. Instead, macro-physical configurations may be taken to instantiate legitimate higher-order natural *dispositions* which are not exhausted, or screened off, by the micro-physical dispositions that realize them case by individual case. Example: I can tell you that an individual configuration *a* has a probability of 1/2 of evolving into an individual configuration *b*, but that macro-configurations of type *A*, to which *a* individually belongs, evolve into macro-configurations of class *B* with a probability of 9/10. The higher order law may even be a stronger natural disposition.[13]

As this paper is not a study of wide causation per se, I will simply note in passing that this kind of ''two levels'' approach is relevant to, and works well with, Mach's language above about 'physical' versus 'psychical' variations of the self-same elements on causal maps. As illustrated below, a physical neural cluster (**e**) is a cluster or configuration of micro-physical dispositions ($e_1 \ldots e_n$) and appears to an external investigator as a population of separate cells in a snarl of complex connections to one another, any of which may be investigated separately. These are its 'physical' variations. The neural energy of this whole cluster in itself, however, has the psychical disposition to appear as the quality green (**s**/**e**) and to interact and interfere with, or inhibit other sensations as one provisional sensory quality among others (see Figure 1).

Hence the causal map will represent both kinds of causal relation among elements some individual, some polyadic type relations, some micro, some macro,

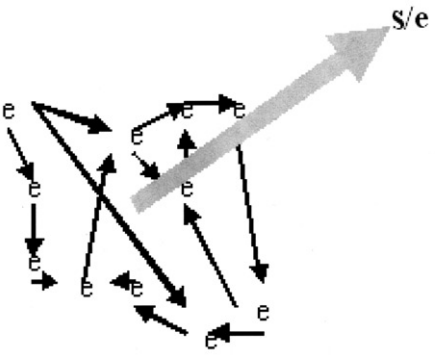

Figure 1 Macro-physical dispositions.

without preferring one over the other. Indeed Mach's flexible notion of causation-cum-functional relations allows this quite naturally. Of course we still need further proof that such macro-types really exist and then we need the laws that relate them. But it is best not to prejudge the question. Levels of explanation and modeling can be very important in the sciences and perhaps especially for biology and neuroscience. In specific, our sense organs are probably adapted to overall environmental macro-conditions, and very insensitive to micro-conditions or realizations. Lockwood (1989, 2007) gives an interesting range of scientific examples from convection cells to quantum wave-functions.

4. Russell Redux: Neo-Neutral Monism and Panpsychism

Neutral monism never really disappeared in the twentieth century. After Russell, it was taken up in a problematic way by Rudolf Carnap and the Vienna Circle, by Erwin Schrödinger, by Herbert Feigl and many others (Carnap, 2003; Feigl, 1967; Schrödinger, 2006). Recently, there has been a modest revival of the position deriving from a passage in Russell's (1927/1959) *Analysis of Matter* (see Banks, 2003, 2004; Chalmers, 1996, 2003; Holman, 2008; Lockwood, 1989; Maxwell, 1978; Rosenberg, 2004; Stoljar, 2001a; Strawson, 2006; and Unger, 1999, among others):

> Physics, in itself, is exceedingly abstract and reveals only certain mathematical characteristics of the material with which it deals. It does not tell us anything as to the intrinsic character of this material.... But by bringing physics and perception together, we are able to include psychical events in the material of physics and to give physics the greater concreteness which results from our more intimate acquaintance with the subject matter of our own experience.... As to intrinsic character, we do not know enough about it in the physical world to have a right to say that it is very different from that of percepts. (Russell, 1927/1959, pp. 10; 400–401)[14]

In other words, physics deals with the relational or dispositional characteristics of matter, which can enter into abstract mathematical laws. It does not deal with the intrinsic *character* of matter, or what concretely instantiates those dispositions.

One option is to adopt a suitably enhanced physicalism,[15] in which these externally described dispositions are concretely instantiated by something with "intrinsic character," presumably the same way our sensations concretely instantiate states of our physically described brains (as Russell held elsewhere; Russell, 1921).

The mathematician Max Newman (1928) pointed out that Russell's preliminary definition of physics as *purely* relational would not suffice, since (according to Russell's own theory of relations!) given any cardinal number of objects, all of the abstract combinatorial relations between them simply come along for free. *Which* among all these abstract possibilities are concretely instantiated by physics, for example by genuine causal connections between objects? Russell readily admitted this criticism (Demopoulos & Friedman, 1989; Seager, 2006), but Newman is actually endorsing Russell's real proposal for an enhanced physicalism, in which causal dispositions are concretely instantiated by qualities and are *not* just abstract combinations of objects. It is this proposal for an enhanced physicalism that most authors have taken from the quote. Rather than canvass all of the different proposals cited above, I will make use of David Chalmer's (2003) concise summary in which he dubs the Neo-Russellian position 'Type-F monism'. According to Chalmers:

> Type F monism is the view that consciousness is constituted by the intrinsic properties of fundamental physical entities: that is, by the categorical bases of fundamental physical dispositions. On this view, phenomenal or protophenomenal properties are located at the fundamental level of physical reality and in a certain sense, underlie physical reality itself.... As a bonus this view is perfectly compatible with the causal closure of the microphysical, and indeed with existing physical laws. The view can retain the structure of physical theory as it already exists; it simply supplements this structure with an intrinsic nature... (proto)phenomenal properties serve as the ultimate categorical basis of all physical causation. (Chalmers, 2003, p. 265)

These 'categorical properties' are, for Chalmers, the concrete instantiations of causal forces, and are themselves dispositions (see Banks, 2009). They have both an intrinsic, or concrete, aspect as real qualities, and they have also the relational or dispositional aspect as it is known in physics. Thesis number four above seems completely satisfied by Type F-Monism. On the other hand, these categorical properties are what Chalmers calls "(proto) phenomenal properties" inside of even the smallest parts of matter, a suggestion with disturbing panpsychist overtones. As the physical stuff combines, according to physical laws, into ever more complex structures like the brain, these tiny protophenomenal properties inside matter combine into conscious experience and sensation. Chalmers suggests that there must be principles for how these protophenomenal properties combine, which would run in parallel with, but be independent of, the way that the physical dispositions combine:

> This is... the combination problem for panpsychism. To answer it, it seems that we need a much better understanding of the compositional principles of phenomenology: that is, the principles by which phenomenal properties can be composed or constituted from underlying phenomenal properties, or from protophenomenal properties. (Chalmers, 2003, p. 266)

Why would we need two completely different, parallel orders of composition if the protophenomenal properties are exactly the occupants of the known causal roles of physics? Won't the principles of combination be just the same? I can't see how there could be a second ordering besides that by which physical dispositions ordinarily configure by known physical laws.[16] Moreover the idea that physical dispositions in matter are little "proto-sensations" seems wrong, much less proto-sensations in little "proto-egos" (Strawson, 2006). Elements are concrete instantiations of physical dispositions, but they are not all sensations, or even possible sensations, and they certainly don't depend on there being little egos to perceive them. So Type-F monism appears to lack features of classic neutral monism already.

Instead, as outlined above, why not let the phenomenal properties (sensations) be the concrete instantiations of higher order dispositions, while the lower-level physical dispositions are concretely instantiated by physical elements in matter that are not anyone's sensations, nor related to them by a line of descent? Here, Chalmers explicitly rejects the notion that a sensation quality might be intrinsic "to high level physical dispositions (e.g., those involved in neural states)":

> If the low-level network is causally closed and the high level intrinsic properties are not constituted by low-level intrinsic properties, the high level intrinsic properties will be epiphenomenal. The only way to embrace this position would seem to be in combination with a denial of microphysical causal closure, holding that there are fundamental dispositions above the microphysical level which have phenomenal properties as their grounds. (Chalmers, 2003, p. 267)

I give him credit for seeing the possibility, but he, too, has fallen for the one-size-fits-all notion of causation criticized above. This prevents him from seeing the appeal of making sensations the instantiation of higher order dispositions of configurations of neurons. Given that we are still owed a definition of causation, *why* is it impossible that a configuration of elements instantiating a cluster of micro-physical dispositions might not also collectively instantiate a disposition of the cluster as a whole?[17] There is no need for a mysterious line of descent from conscious sensation to inorganic matter by which protophenomenal properties combine by a set of weird laws into conscious phenomenal properties; we can just stick with physics and its laws, complemented by further empirical laws describing the higher order dispositions of configurations of neurons. Lockwood (1993) addresses this point explicitly:

> Their [basic dispositions of matter] potential for generating awareness could be a matter of the application of certain currently unknown laws to their familiar physical attributes (in which laws, of course, there would be an essential reference to the emergent attributes). This fairly elementary point seems to have escaped those authors who have argued that, if we are made of electrons, quarks, gluons, and the like, then—given that we are conscious—electrons, quarks and so forth must themselves be possessed of some sort of primitive proto-consciousness. As I see it, this is a complete non-sequitur (p. 281).

I would say that instead of a bogus 'composition problem' for neutral monism, the principle of causal closure and the lack of any nuanced definition of 'cause' seems to

be the real problem. There is a need for a better understanding of multi-level causation and the variety of causal dispositions that exist to be concretely instantiated in nature by elements and element/sensations, both lower and higher level 'attributes', on the causal map of nature. This is a problem for the philosophy of science, which is not a bad result; at least we are learning where to locate the problem correctly.

5. Conclusion

The odds of convincing contemporary philosophers of mind to reconsider neutral monism are slight. The concepts of neutral monism are estranged from those in use in the contemporary discussion, in particular notions of qualia, causation, and causal closure in particular are areas where misunderstandings could easily blossom. And then there is the question: *which* neutral monism? Classic, or the neo-monist views which veer increasingly panpsychist, a much less interesting and probably false view. A better suggestion might be to define elements as concretely instantiated dispositions, and to fill out the picture of nature to include higher level attributes and their concretely instantiated dispositions as element/sensations.

Notes

[1] The following historical account is only a sketch, which leaves out other interesting versions due to Avenarius and Clifford. The full version is told in Banks (2003).

[2] Both Mach and Russell were skeptical of the notion of cause, with good reason, and preferred to express the notion of causal law by means of mathematical functions (solutions to differential equations). There are difficulties with such a view but they cannot be discussed here. I will conflate causal and functional connections in the paper, but this should not indicate that I think they are the same thing. Mathematical functions allow for flexibility in modeling causal relations (for example in causal maps below) and a model, or causal map, can be tested by experience for robustness and to screen off illusory correlations; see Spirtes, Glymour, and Scheines (2002).

[3] Mach (1959), p. 310; Mach was also deeply persuaded by philosopher J.F. Herbart to adopt Herbartian energetic qualities as his elements (Banks, 2003, 2005).

[4] Unfortunately the word 'element' is used for both the specific physical variations and objects and for the overall general class of elements and variations, which includes sensations and physical objects. World-elements might be a better designation for the most general class (as in Banks, 2003).

[5] This continues to be a basic misunderstanding of neutral monism, i.e., that it is a form of phenomenalism or idealism. See Mach (1959), p. 62. For the notion of mind-independent world elements in Mach, James and Russell, see Banks (2003). See also Stubenberg (2005), who points out that "the view that Russell's neutral elements are mental—in the sense of being confined to an existence in the mind—is quite popular and quite wrong".

[6] Mach called this principle the "functional dependence of elements on one another." James claimed that his pure experience elements were embedded in "energetic" relations, and that "real relations" were instantiated by bits of pure experience. Russell said his event particulars were causal "effects" collected into objects and perspectives (Russell, 1921, p. 101). The conception of sensation qualities as causal forces, bound in real relations to one another,

is where neutral monism parts ways with traditional British Empiricism and contemporary notions of ''qualia.'' Mach's notion of element-sensation owes more to Kant, whom Mach read as a boy. Kant's *Eindrücke* or impressions are forces, or as Kant calls them 'moments' of force, already bound up in laws or functions and are not Humean atoms of experience (see Banks, 1992).

[7] John Heil (2002) has recently proposed an ontology of pure powers, which replaces the object property ontology in his scheme. To my mind, epiphenomenal qualia are a gross violation of the spirit of evolutionary theory and natural selection. Why are these qualia so richly structured in manifold representations like color-space and sound space if these representations play no evolutionary role in producing behavior?

[8] Of course our concepts might be deeply flawed due to our ignorance, and thus thought experiments based upon them may be binding on thoughts and arguments using these concepts, but not on reality (Stoljar, 2007).

[9] Spatially projecting sensations into the interiors of brain tissue was called the ''error of introjection'' by Mach's colleague Richard Avenarius, whose work has been neglected by historians of philosophy perhaps because of being overshadowed by Mach.

[10] Assuming we have some way of separating ''the'' sensation out provisionally from the complex of other sensations to which it belongs; they are not atoms or sense data. Sensations are specific to certain neurons in some cases, which can be misleading because neurons also have many of their properties through connections to surrounding neurons and the functional pattern of structural connections and firing patterns (see Kandel, Schwartz, & Jessell, 2000). The individual neuron does the actual triggering so to speak, but within a nexus of other conditions that constrain its firing and guide it.

[11] The principle that the domain of physical events is closed under the causal relation the way that the natural numbers are closed under addition: the sum of every two numbers is a number, every physical effect *E* has a physical cause *C* (see Kim, 1998).

[12] Some authors (such as Hitchcock, 1995) have pointed out that the law of the conservation of energy does not define the closed system in which energy is conserved, but rather physics accepts a circular criterion: a system is closed if energy is conserved in it; energy is conserved in a system if it is closed. It is true that one could appeal to conservative forces and mechanisms in the micro-description of a system in any particular case, but there is no such mechanism or micro-structural feature closed conservative systems have in general, in keeping with the phenomenological mechanism-independence of the law.

[13] It is sometimes suggested that the macro-causal factors are only roughly law-like or *ceteris paribus* law-like with lots of exceptions (see Davidson, 1993) but I see no reason this must be so, for the macro-law relating types could be stronger. I also cannot deal here with the barrage of arguments thrown up by Jaegwon Kim (1998) against macro-causation, except to say that macro-causation can be said to codetermine a given event, without conflicting with micro-causation, if it helps bring about events in a different, but complementary way. Neither are macro-causes screened off by their micro-causal realizations, as Kim claims. A correlation of *A* and *B* is established by the fact that the probability of *A*&*B* is different from the Probability of *A* alone times the probability of *B* alone. A correlation is 'screened off' by a factor *c* if, by conditionalizing on *c*, we can reduce the correlation of *A* and *B* conditional on *c*, *A*&*B*/c, to the probability of *A*/*c* alone times the probability of *B*/*c* alone. But there is no in general micro-physical way to specify the *c*-screeners *as a class*, only a macro-causal way of doing so. You could screen off each macro-causal correlation one by one, but not be able to characterize the screener off condition in general. David Papineau (1993) similarly prejudges the issue by arguing that if a macro-causal factor or type is causally efficacious, say in producing behavior in an organism, their micro-causal realizations must have something in common as a class. But that presumes the macro-type will be couchable in micro-level descriptions.

[14] Michael Lockwood (1989) points out that this quote is a very close paraphrase of a section of the Paralogisms in Kant's *Critique of Pure Reason*. It also echoes some of the things Kant says about sensation and the *Ding an sich* (Banks, 2005).

[15] For a definition of standard physicalism see Gillet and Loewer (2002) and Stoljar (2001b). Roughly it is the idea that any world minimally identical to ours in all (micro) physical respects, consisting of the particles and fields defined by physics, is identical in all other respects.

[16] If they are not the instantiations of the physical dispositions but only 'ground' those dispositions without playing a direct causal role then the problem is worse and they become not epiphenomena but endo-phenomena (Banks, 2009; Blackburn, 1992).

[17] Elsewhere (Banks, 2009) I argue that concretely instantiated dispositions of configurations *are* what we feel as the qualities of our experience. A structural property of a neural cluster, like its asymmetry, shape, or functional pattern of firing is *felt*, dispositionally, as the quality of the sensation in my view, there is no superaddition necessary.

References

Banks, E. (2003). *Ernst Mach's world elements: A study in natural philosophy*. Dordrecht: Kluwer.

Banks, E. (2004). Philosophical roots of Ernst Mach's economy of thought. *Synthese*, *139*, 23–53.

Banks, E. (2005). Kant, Herbart and Riemann. *Kant Studien*, *96*(2), 208–234.

Banks, E. (2009). Russell's hypothesis and the new physicalism. *Proceedings of the Ohio Philosophical Association*. Available online at <http://proceedings.ohiophilosophy.org/006005>.

Blackburn, S. W. (1992). Filling in space. *Analysis*, *52*, 62–63.

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. New York: Appleton Century Crofts.

Carnap, R. (2003). *The logical structure of the world and pseudoproblems in philosophy*. (R. A. George, trans.) Chicago: Open Court.

Chalmers, D. (1996). *The conscious mind*. Oxford: Oxford University Press.

Chalmers, D. (2003). The place of consciousness in nature. In S. Stich & F. Warfield (Eds.), *Blackwell guide to philosophy of mind* (pp. 102–142). Malden, MA: Blackwell.

Clark, A. (1993). *Sensory qualities*. Oxford: Oxford University Press.

Davidson, D. (1993). Thinking causes. In J. Heil & A. Mele (Eds.), *Mental causation* (pp. 3–18). Oxford: Oxford University Press.

Demopoulos, W., & Friedman, M. (1989). The concept of structure in Russell's Analysis of Matter. In C. Wade Savage & C. Anthony Anderson (Eds.), *Rereading Russell: Essays in Bertrand Russell's metaphysics and epistemology*. Minneapolis: University of Minnesota Press.

Dretske, F. (1993). Mental Events as Structuring Causes of Behavior. In J. Heil & A. Mele (Eds.), *Mental causation* (pp. 121–136). Oxford: Oxford University Press.

Einstein, A. (1949). Autobiographisches. In P. Schlipp (Ed.), *Albert Einstein: Philosopher–scientist* (pp. 1–96). Chicago: Open Court.

Feigl, H. (1967). *The mental and the physical: The essay and postscript*. Minneapolis: University of Minnesota Press.

Garfinkel, A. (1990). Reduction. In R. Boyd, P. Gaspar, & J. D. Trout (Eds.), *The philosophy of science* (pp. 443–459). Cambridge, MA: MIT Press.

Gillet, C., & Loewer, B. (Eds.). (2002). *Physicalism and its discontents*. Cambridge: Cambridge University Press.

Hatfield, G. (2002). Sense-data and the philosophy of mind: Russell, James and Mach. *Principia*, *6*, 203–230.

Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, *4*, 31–52.

Hausman, D. (1998). *Causal asymmetries*. New York: Cambridge University Press.

Heidelberger, M. (2004). *Nature from within: Gustav Theodor Fechner and his psychophysical worldview*. Pittsburgh: University of Pittsburgh Press.
Heidelberger, M. (2007). Functional relations and causality in Fechner and Mach. In S. Mori, T. Miyaoka, & W. Wong (Eds.), *Fechner Day 2007* (pp. 33–38). Tokyo: International Society for Psychophysics.
Heil, J. (2002). *From an ontological point of view*. Oxford: Oxford University Press.
Hiebert, E. (1960). *The historical roots of the principle of the conservation of energy*. Madison: University of Wisconsin Press.
Hitchcock, C. (1995). Salmon on explanatory relevance. *Philosophy of Science*, *62*(2), 304–320.
Holman, E. (1986). Maxwell and materialism. *Synthese*, *66*, 505–514.
Holman, E. (2008). Pan-psychism, physicalism, neutral monism and the Russelian theory of mind. *Journal of Consciousness Studies*, *15*, 48–67.
James, W. (1977). Essays in radical empiricism. In J. J. McDermott (Ed.), *The writings of William James*. Chicago: University of Chicago Press.
Kandel, E., Schwartz, J., & Jessell, T. (2000). *Principles of neural science*. New York: McGraw-Hill.
Kim, J. (1998). *Mind in a physical world*. Cambridge, MA: MIT Press.
Lockwood, M. (1989). *Mind, brain and the quantum*. Oxford: Basil Blackwell.
Lockwood, M. (1993). The Grain Problem. In H. Robinson (Ed.), *Objections to physicalism* (pp. 271–291). Oxford: Oxford University Press.
Lockwood, M. (2007). *The labyrinth of time*. Oxford: Oxford University Press.
Mach, E. (1911). *History and root of the principle of the conservation of energy*. (P. E. B. Jourdain, trans.) Chicago: Open Court.
Mach, E. (1959). *The analysis of sensations*. (C. M. Williams & S. Waterlow, trans.) New York: Dover.
Mach, E. (1960). *The science of mechanics*. (T. McCormack, trans.) LaSalle: Open Court.
Maxwell, G. (1978). Rigid designators and mind–brain identity. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science* (Vol. 9, pp. 365–403). Minneapolis: University of Minnesota Press.
McGinn, C. (1991). *The problem of consciousness*. Cambridge: Blackwell.
Montero, B. (1999). The body problem. *Nous*, *33*, 183–200.
Newman, M. (1928). Mr. Russell's causal theory of perception. *Mind*, *37*, 137–148.
Papineau, D. (1993). *Philosophical naturalism*. Oxford: Blackwell.
Papineau, D. (2002). *Thinking about consciousness*. Oxford: Oxford University Press.
Richardson, A. (1990). How not to Russell Carnap's Aufbau. *Proceedings of the Philosophy of Science Association*, *1*, 3–14.
Rosenberg, G. (2004). *A place for consciousness*. Oxford: Oxford University Press.
Russell, B. (1921). *The analysis of mind*. London: Allen & Unwin.
Russell, B. (1957). *Mysticism and logic*. New York: Anchor Books. (Original work published 1917)
Russell, B. (1959). *The analysis of matter*. New York: Dover. (Original work published 1927)
Russell, B. (1984). Theory of knowledge. In *The Collected Papers of Bertrand Russell (Vol. 2)*. London: Allen & Unwin. (Original work published 1914)
Schrödinger, E. (2006). *What is life? With mind and matter and autobiographical sketches*. Cambridge: Cambridge University Press.
Seager, W. (2006). The 'intrinsic natures' argument for panpsychism. *Journal of Consciousness Studies*, *13*, 129–145.
Spirtes, P., Glymour, C., & Sheines, R. (2001). *Causation, prediction and search*. Cambridge, MA: MIT Press.
Stoljar, D. (2001a). Two conceptions of the physical. *Philosophy and Phenomenological Research*, *62*, 253–281.
Stoljar, D. (2001b). Physicalism. *Stanford Encyclopedia of Philosophy*.
Stoljar, D. (2007). *Ignorance and imagination*. Oxford: Oxford University Press.

Strawson, G. (2006). Realistic monism: Why physicalism entails panpsychism. *Journal of Consciousness Studies*, *13*, 3–31.
Stubenberg, L. (1998). *Consciousness and qualia*. Amsterdam: John Benjamin's.
Stubenberg, L. (2005). Neutral monism. *Stanford Encyclopedia of Philosophy*.
Unger, P. (1999). The mystery of the physical and the matter of qualities. *Midwest Studies in Philosophy*, *23*, 75–99.
Wackermann, J. (2007). Psychophysics or algebra of pure experience? In S. Mori, T. Miyaoka, & W. Wong (Eds.), *Fechner Day 2007* (pp. 51–56). Tokyo: International Society for Psychophysics.
Yablo, S. (1992). Mental causation. *Philosophical Review*, *101*, 245–280.
Yablo, S. (2003). Causal relevance. *Philosophical Issues*, *13*, 316–329.
Yablo, S. (2009). *Thoughts*. Oxford: Oxford University Press.
Yoo, J. (2006). Mental Causation. *The Internet Encyclopedia of Philosophy*.