

Kant's Perspectival Solution to the Mind-Body Problem

Or, *Why Eliminative Materialists must be Kantians*

Stephen R. Palmquist

Department of Religion and Philosophy,
Hong Kong Baptist University, Hong Kong, China
stevepq@hkbu.edu.hk

Abstract

Kant's Critical philosophy solves Descartes' mind-body problem, replacing the dualism of the "physical influx" theory he defended in his early career. Kant's solution, like *all* Critical theories, is "perspectival," acknowledging deep truth in both opposing extremes. Minds are not *separate* from bodies, but a manifestation of them, each viewed from a different perspective. Kant's transcendental conditions of knowledge portray the mind not as *creating* the physical world, but as necessarily structuring our knowledge of objects with a set of unconscious assumptions; yet our *pre-conscious* (*pre-mental*) encounter with an assumed spatio-temporal, causal nexus is entirely physical. Hence, today's "eliminative materialism" and "folk psychology" are both *ways of considering* this age-old issue, neither being an exclusive explanation. A Kantian solution to this version of the mind-body problem is: eliminative materialism is good science; but only folk psychologists can consistently *be* eliminative materialists. Indeed, the mind-body problem exemplifies a feature of all cultural situations: dialogue between opposing perspectives is required for understanding *as such* to arise.

Keywords

Immanuel Kant – mind-body problem – eliminative materialism – neurophilosophy – perspectives

Eliminative Materialism as a Challenge to Kant's Apparent Dualism

Readers of Immanuel Kant's great work, *Critique of Pure Reason* – whether or not they regard themselves as Kantians – tend to assume that his position entails some form of mind-body dualism.¹ After all, his ground-breaking “transcendental idealism” argues that the *mind* imposes necessary and universal conditions onto the *physical* world as we know it. Kant's insistence that good philosophy must start with the assumption that space and time are mental constructs (dubbed “forms of intuition”), his frequent appeal to a distinction between the “phenomenal” and “noumenal” worlds, and his incessant use of terms referring to a complex array of mental faculties, all suggest that he was merely *adopting*, rather than transcending, the legacy bequeathed to him by Descartes. Indeed, many would regard Kant's position as a dualist *alternative* to the anti-Cartesian monism, known as “eliminative materialism” (hereafter “EM”), that some philosophers, most notably Patricia Churchland, have defended in recent years.

EM emerged during the last quarter of the twentieth century as one of the most radical and controversial theories of mind-brain identity. At its heart is an *assumption* about the role of concepts that refer to mental states and a *prediction* regarding the future progress of neuroscience. The assumption is two-fold: first, dualism is grounded in what proponents of EM call “folk psychology” (i.e., the commonsense view that our mental states, especially our desires and beliefs, are *causes* of the behavior that we regard as self-consciously *ours*);² second, folk psychology functions as a *theory* when used in such an explanatory

1 The examples that could be cited here are myriad. Particularly ironic among the illustrious list of those who see Kant as a dualist are those whose attempt to defend an account of the *embodied* mind that could have been significantly buttressed by recognizing that Kant can be read in a very different way. One of the best examples of this subgroup is the book by George Lakoff and Mark Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought* (New York: Basic Books, 1999). For a critique of their approach, see my book review in *Journal of Scientific Exploration* 24, 2 (Summer 2010): 323-327.

2 This use of the term “folk psychology” assumes that the common-sense understanding of the causal role of mental states is similar to other folk theories that have appeared to be self-evident, until hard scientific evidence demonstrated otherwise. For example, folk cosmology claimed that the earth is flat and that the sun revolves around it, until Copernicus' alternative theory gained empirical confirmation. Proponents of EM cite various examples of folk theories that have been debunked. See for example, Patricia S. Churchland, *Neurophilosophy: Toward a Unified Theory of the Mind-Brain* (Cambridge, MA: The MIT Press, 1986), 290 and 300-301. She defines a “folk theory” in general as an “intuitive framework” for understanding some aspect of our human experience (288).

way. The prediction EM makes is that brain research will eventually become sufficiently sophisticated to enable neuroscientists, perhaps with the help of some empirically-minded philosophers, to develop a complete theory of the causal relationship between brain functioning and human behavior. Once this occurs, folk psychology will be reducible to brain science, so that references to our “desires” and “beliefs” will have no more literal truth than references to the sun “rising” over the eastern horizon of the earth. In short, the core prediction of EM is that the human brain is smarter than it is complicated.³

Among the many empirical facts cited to support EM is that the human mind-brain enables us to perform many highly complex tasks *without* the involvement of a conscious “ego” directing its processes. While examples of such facts could be taken from a wide variety of common human experiences, from speaking to the mechanics of walking or driving a car, an example that is particularly relevant in philosophical circles is given in a recent magazine article detailing the life and thought of the contemporary British philosopher, Derek Parfit. The author describes Parfit’s normal approach to writing as follows:

He doesn’t believe that his conscious mind is responsible for the important parts of his work. He pictures his thinking self as a government minister sitting behind a large desk, who writes a question on a piece of paper and puts it in his out-tray. The minister then sits idly at the desk, twiddling his thumbs, while in some back room civil servants labor furiously, come up with the answer, and place it in his in-tray.⁴

While perhaps not everyone has the good fortune to be able to come up with creative philosophical ideas in this way, we *do* all deal with a wide variety of other complex tasks in much the same way, hardly giving a single (conscious) thought to what we are doing. How is this possible? If EM is correct, the answer lies *solely* in the workings of the brain: we will eventually realize that such examples describe the norm, not an exception, because even the ego is actually one of the “civil servants” whose tasks collectively constitute the human brain.

Almost as if she were attempting to account for this biographical description of creative thinking, Churchland describes the brain as consisting of “suitably orchestrated throngs of stupid things,” observing that “it seems quite shocking

3 Churchland, *Neurophilosophy*, 316; she readily admits, however, that “the brain . . . *may* be more complicated than it is smart” (374).

4 Larissa MacFarquhar, “How to be Good,” *The New Yorker* (September 5, 2011): 42-53; quote from page 44.

that one's cleverness should be the outcome of well-orchestrated stupidity."⁵ Churchland (and EM in general) claims this requires us to *unlearn* our belief in a controlling "self." However, these very insights regarding the workings of the brain might point not to complete eradication of the human self, but to its redefinition, in much the same way that Jung attempted to replace Freud's conviction, that the conscious "ego" ought to be in control of one's mental life, with the new conviction that the path to genuine psychological health is the ability to trust in our awareness of a "Self" that is greater than any conscious "ego." On the interpretation of Kant that I shall present below, this higher "Self" can be identified with *the body*. Neuroscience will begin to dovetail nicely with philosophy when it recognizes that the brain itself *trusts the body* for virtually all of its functioning, just as Jung says the ego must trust the Self-archetype.⁶

Because Kant never considered the challenge EM poses to philosophers, at least not in such an extreme form, we cannot be certain just how he would have responded. My goal here, therefore, is not to argue that Kant would either endorse or reject EM per se, but to defend two claims: (1) transcendental philosophy can be conceived, like EM, as *anti-dualist*; and (2) without being supplemented by a key feature of Kant's theory of mind (one that is not inconsistent with any essential feature of EM – namely, his claim that reflective *perspectives* are necessary for the possibility of human understanding), EM cannot be consistently maintained. If my argument is correct, then at the very least, Kantians and eliminative materialists ought to remain in close dialogue, if representatives of either camp wish to arrive at an accurate understanding of the workings of the human mind-brain.⁷

Many would expect Kant, with his reputation as an adherent of Cartesian dualism, to be among the last modern philosophers to whom we might turn for insight into the question of how best to defend a theory of mind-brain *identity*. However, Kant scholars over the past few decades have made considerable progress in dispelling the myth of Kant the naive Cartesian dualist.⁸ While not explicitly addressing the mind-brain issue, Henry Allison's 1983 book, *Kant's*

5 Churchland, *Neurophilosophy*, 407.

6 For further discussion of Jung's position, see Stephen Palmquist, *Dreams of Wholeness: A Course of Introductory Lectures on Religion, Psychology and Personal Growth*² (Hong Kong: Philopsychy Press, 2008 [1997]), chs. 6-7.

7 Given this more modest (but still quite ambitious) goal, I suspect that, despite my best efforts to argue otherwise, some proponents of EM might respond in the end that all I have done is to confirm their suspicions that Kant was a dualist after all.

8 A good discussion of the myth of Kant the Cartesian dualist can be found in Arthur W. Collins, *Possible Experience* (Berkeley: University of California Press, 1999), 5-9 and *passim*. See also Helge Svarre, *Body and Practice in Kant* (Dordrecht: Springer, 2006).

Transcendental Idealism,⁹ has persuaded many contemporary Kant-scholars that what looks like an untenable *ontological* dualism defined by two distinct *worlds* (the phenomenal and the noumenal) is properly understood as two distinct *perspectives* on one and the same world: the physical world of our everyday experience. My 1993 book, *Kant's System of Perspectives*,¹⁰ formalized the perspectival approach and extended it to Kant's entire Critical System, arguing that Kant's ultimate goal in philosophizing was to articulate an idea of human nature as a *unified* whole. In that work, however, I did not provide a detailed explanation of how Kant's new, perspectival approach transcends Cartesian dualism. Not unlike the charge often leveled against EM, Kant appears to be *changing the subject* of philosophy so radically that he often simply ignores the old mind-body problem. Fortunately, 1993 also saw the publication of Alison Laywine's thoroughgoing study demonstrating how the issue of soul-body interaction was the chief philosophical focus of Kant's pre-1770 (what I call his "pre-Copernican"¹¹) writings.¹² (What is left unclear by that study is just how Kant's mature Critical philosophy allowed him to *transcend* his early dualist position.) And since then a number of excellent studies have been published, arguing that Kant was far more body-oriented than has often been assumed.¹³ These have helped to dispel the impression of many previous commentators,

-
- 9 Henry E. Allison, *Kant's Transcendental Idealism: An Interpretation and Defense* (New Haven: Yale University Press, 1983). For a critique of Allison's position, see Kenneth R. Westphal, "Freedom and the Distinction Between Phenomena and Noumena: Is Allison's View Methodological, Metaphysical, or Equivocal?," *Journal of Philosophical Research* 26 (2001): 593-622.
- 10 Stephen R. Palmquist, *Kant's System of Perspectives: An architectonic interpretation of the Critical Philosophy* (Lanham, MD: University Press of America, 1993).
- 11 In *Kant's Critical Religion: Volume Two of Kant's System of Perspectives* (Aldershot: Ashgate, 2000), §11.1, I argue that the convention of calling the years before 1770 Kant's "pre-Critical" period is a misnomer: a proper understanding of Kant's term, "Critical" (i.e., its grounding in a three-step process of resolving philosophical problems, whereby one considers two opposing positions then resolves the opposition by referring to a third position that synthesizes the first two), enables us to trace Critical strains in his philosophy from his very first publication to the last. The new feature introduced in (or around) 1770 is what he refers to in the second Preface to the first *Critique* as his *Copernican hypothesis* (Bxxvif, Bxxiin).
- 12 Alison Laywine, *Kant's Early Metaphysics and the Origins of the Critical Philosophy* (Atascadero, CA: Ridgeview Publishing Company, 1993), 52 and 159.
- 13 See especially Susan Meld Shell, *The Embodiment of Reason: Kant on Spirit, Generation, and Community* (Chicago: Chicago University Press, 1995). Svare, *Body and Practice in Kant*, 3n, lists several other key authors who have also emphasized the importance of embodiment for Kant.

that Kant was not only a dualist but, largely because of the somewhat strict moral theory defended in his second *Critique*, also a body-hater.¹⁴

Again, my purpose here will be not be so grand as to claim that Kant himself was a confirmed physicalist, explicitly defending mind-brain identity. This would be too extreme, for his discussion of mind-body interaction in the Paralogisms (especially A381-96) presents an *idealist* solution, consistently using language that seems to assume that mental states really exist. Rather, more modestly, my purpose will be to explore whether Kant's mature Critical philosophy is *consistent with* mind-body identity, and whether Kant has any insights that might challenge those who affirm an extreme form of EM to correct potential defects in their own philosophy of mind.

Kant on the Brain in *Dreams of a Spirit Seer*

Before examining how the *Critique of Pure Reason* defends an alternative to Kant's youthful dualism, and how his transcendental idealism can be understood as overcoming dualism entirely, let us look briefly at a book that I believe marks *the* crucial turning-point in his philosophical development: *Dreams of a Spirit-Seer Elucidated by Dreams of Metaphysics* (1766). In his earliest writings, Kant defended a theory of "physical influx," whereby the soul has quasi-material characteristics, such as impenetrability.¹⁵ Having learned of Swedenborg's amazing visions in the early 1760s, Kant read some of the mystic's explanations of how such visions occur. In *Dreams* we witness Kant grappling with the uncomfortable fact that Swedenborg held a vulgar version of the same dualist metaphysics that he himself had been promoting. As a result, Kant tells us toward the end of this much-neglected early work, he found it necessary to *give up* his formerly naive reliance on "the *butterfly-wings* of metaphysics" – presumably this included his dualist theory of mind-body interaction – and remain "on the humble ground of experience and common sense," where we "devote ourselves to what is useful."¹⁶ He then foreshadows his intention to

14 As the examples of this tendency are legion, a single reference will suffice: Laura Hengehold, in *The Body Problematic: Political Imagination in Kant and Foucault* (University Park, PA: The Pennsylvania State University Press, 2007), 7-8, says Kant "protected his sense of moral integrity and autonomy against inclinations associated with the body."

15 See Laywine, *Kant's Early Metaphysics*, 25-42, for a detailed account of this position. I summarize and assess Laywine's arguments in *Kant's Critical Religion*, Appendix II.2.

16 Immanuel Kant, *Dreams of a Spirit-Seer Elucidated by Dreams of Metaphysics*, tr. David Walford with Ralf Meerbote, in *Theoretical Philosophy 1755-1770* (Cambridge: Cambridge University Press, 1992), 368.

perform a thoroughgoing critique of reason's cognitive powers that will determine "the limits imposed upon [science] by the nature of human reason."¹⁷

Dreams of a Spirit-Seer, Kant's last major publication prior to the appearance of his *Critique of Pure Reason* some 15 years later,¹⁸ advances a cutting criticism of his formerly cherished theory of physical influx. Without rehearsing the details of that theory or why Kant came to recognize it as fundamentally flawed, I shall limit my attention here to a single passage from Chapter I, where he explicitly discusses the question of whether it makes sense to locate the "soul" (and its mental capacities) in a particular part of the body. His poignant answer appeals to the limits of the *brain's* role in human thinking and is worth quoting at length:

The body, the alterations of which are *my* alterations – this body is *my* body; and the place of that body is at the same time *my* place. If one pursued the question further and asked: Where then is *your* place (that of the soul) in this body? Then I should suspect there was a catch in the question. For it is easy to see that the question already presupposes something with which we are not acquainted through experience, though it may perhaps be based on imaginary inferences. The question presupposes, namely, that my thinking 'I' is in a place which is distinct from the places of the other parts of that body which belongs to my self. But no one is immediately conscious of [occupying] a particular place in his body; one is only immediately conscious of the space which one occupies relatively to the world around. I would therefore rely on ordinary experience and say, for the time being: Where I feel, it is there that *I am*. I am as immediately in my finger-tip as I am in my head. It is I myself whose heel hurts, and whose heart beats with emotion . . . No experience teaches me to regard some parts of my sensation of myself as remote from me. Nor does any experience teach me to imprison my indivisible 'I' in a microscopically tiny region of the brain, either so as to operate from there the

17 Kant, *Dreams*, 369. Only when these "boundary-stones" are "securely fixed" (369), Kant assures us, will metaphysics finally become "that which it is far from being at the moment . . . , namely, the *companion of wisdom*."

18 In 1768 Kant published his short essay "*Von dem ersten Grunde des Unterschiedes der Gegenden im Raume*", and in 1770 he wrote the "Inaugural Dissertation" (*De mundi sensibilis atque intelligibilis forma et principiis*) that was required when he took up a professorship at the University of Königsberg. While the latter on its own cannot properly be regarded as a "major work," many of its most important arguments did form the backbone of the first *Critique*, eleven years later. Other than these two works, Kant published nothing more than a few short notices during this 15 year period.

levers governing my body-machine, or so as myself to be affected in that region by the workings of that machinery.¹⁹

While this quotation on its own is not enough to demonstrate that by 1766 Kant had shaken off the cloak of dualism, it does provide some crucial inspiration for the view I shall support here; in any case, it demonstrates that at this point in his life, during his early 40s, he had explicitly identified any thought of *himself* as having, first and foremost (if not entirely), a *physical* reference – to his fingers and feet as well as to his head and heart.

Without further elaboration and interpretation, the foregoing passage presents admittedly thin evidence for an anti-dualist position, especially since Kant seems to affirm that his “soul” consists of an “indivisible ‘I’” that might be metaphysically distinct from the body. Nevertheless, he is here identifying his *awareness* of that “I” with his awareness of his *body* and explicitly states that the “I” itself might be a product of merely “imaginary inferences” – a position tantalizingly consistent with EM. Taking this hint as a starting-point, I shall argue that, far from being simply a further development of Cartesian dualism, Kant’s Copernican revolution in philosophy advances an alternative to *any* form of mind-body dualism by demonstrating that, from the point of view of any legitimate natural science, I just *am* my body.

In the course of defending this radical position, however, Kant also argued that there is one and only one way to *defend* it consistently, and that is to recognize that, out of this very body that *is me* arise forms of knowledge whose *validity* is entirely independent of my physical nature. This position does not make Kant a dualist, nor even an epiphenomenalist, provided we keep in mind the perspectival character of his philosophy: although these special, “transcendental” forms of knowledge are all-important from one perspective (the one that matters most to philosophy), from another perspective (the one that matters most to science) they are imaginary ideas that paradoxically cannot be said to *exist* at all! This is why Kant calls them both “transcendental” and “ideal”: we *must* assume them, for they are the source of all meaning, yet their *necessity and universality* have no empirical grounding. Because we need these forms in order to *explain* what we know, I suggest we call this aspect of his position “*explanatory idealism.*” To unpack the perplexing implications of Kant’s two-sided position, we must move beyond his pre-Copernican period and examine the key features of the *Critique of Pure Reason*.

19 Kant, *Dreams*, 324-325 (all but the second emphasis added).

Kant's Perspectival Alternative to Cartesian Dualism

Kant begins the Preface to the first edition of the first *Critique* with the tantalizing claim that human reason is *inevitably* caught up in a highly problematic way of thinking: one that it cannot dispense with, even though it is entirely groundless as a description of empirical reality. He identifies this system of thinking as “metaphysics,” saying the goal of his book will be to examine the powers and limits of human reason, thus enabling the reader at least to *be aware* of the illusory nature of traditional metaphysical ideas, even though their *effects* on our thinking will remain in place. I can here only briefly outline Kant's critical philosophy of mind, highlighting those features that most clearly demonstrate its monistic emphasis. As we shall see, the illusion Kant unveils is remarkably similar to the illusion EM predicts will someday be dispelled: that what we have become accustomed to calling our “mind” is something separate or distinct from our body.

In the first *Critique*'s first major section, the Transcendental Aesthetic, Kant famously (or on most accounts, infamously) argues that space and time are not to be viewed as absolute containers (as Newton had claimed), nor as mere systems of relations between objects (as Leibniz had argued), but rather as “forms of intuition.” Exactly what he meant by this obscure term has been, and will no doubt continue to be, a matter of considerable debate. But he explicitly states that he is *not* defending a form of idealism even remotely similar to Berkeley's, whereby space and time have no *empirical* reality outside of our perceptions (see B70-71). The fact that Kant appeals throughout the Aesthetic to various examples of perceived objects, as being *externally given* to the human subject in the process of experiencing them, suggests that his special term refers to the requirements of our *bodily functioning*.²⁰ Rukgaber defends this position in detail, arguing that Kantian “intuition” is another term for what we might nowadays call “qualitative sensation processing,” while “forms of intuition” are “an account of the structure of our embodied perspective” on the world.²¹

To set the background for the arguments of the first *Critique*, Rukgaber appeals to Kant's 1768 essay on incongruent counterparts,²² where Kant uses the right and left hands to illustrate how “directionality . . . is provided through the mediation of the body.” The same fundamental appeal to the body occurs in Kant's 1786 essay, “What does it mean to orient oneself in thinking?” Rukgaber

20 See e.g., B70n. See also note 33, below.

21 Matthew S. Rukgaber, “‘The Key to Transcendental Philosophy’: Space, Time and the Body in Kant,” *Kant-Studien* 100 (2009): 166-186; quoting from pages 167 and 166.

22 *Ibid.*, 172. Cf. note 18, above.

claims that in this essay Kant views our bodily senses as constituting “the ‘projection’ of an intuitional field” through the power of “a priori intuition,” understood now as a person’s “bodily grasp on the field of possible spatio-temporal relations that we have by virtue of being an embodied creature.”²³ In Kant’s theory “the relation of the given to the larger spatial frame” is transformed from a secondary feature of our experience to the primary issue of philosophical interest; this is a form of idealism because “the whole of space” is now regarded as “relative to our perspective” – i.e., to our “representational system.”²⁴

Along these lines, Kant argues in the first *Critique* (A26/B42) that “we can accordingly speak of space, extended beings, and so on, only from the human standpoint.” Kant’s primary concern is not with the *concept* of space (as claimed by those such as Allison and Falkenstein, who see the emphasis of Kant’s argument as primarily psychological), but with how our “representation of space” serves “as the ground of all concrete spatial determinations”: “A *priori* intuition precedes psychology, because it is the frame in which all possible experience – even psychological experience – emerges.”²⁵ The common assumption “that the form of inner intuition is identical to the flow of empirical consciousness results in multiple contradictions within Kant’s text.”²⁶ A better interpretation, Rukgaber argues, is to read Kant as “describing something like a general, indeterminate temporal background to events” – both inner (so-called “mental”) and outer (so-called “physical”) events – with the key factor defining this background (namely, “successiveness”) emerging “from our capacity for self-movement . . . In other words, time also emerges from a bodily foundation.”²⁷ Kant’s repeated references to time as being like “a line that is infinitely progressing . . . in space” are also grounded in “the feeling of our own possible movement along an axis of the body.”²⁸ Rukgaber finds the same view operating in Kant’s *Opus Postumum*, which also depicts “a *priori* intuition [as] the way in which the active body projects a structured, intuitional, spatio-temporal background against which objects emerge.”²⁹ On this reading, Kant’s basic claim in the Aesthetic is that “The forms of intuition are

23 Ibid., 173.

24 Ibid., 174 and 178.

25 Ibid., 176-177. This “framing of all possible experience” (177), he adds, is “an ideal projection tied to the formal structure of our perspective.”

26 Ibid., 180.

27 Ibid.

28 Ibid., 181.

29 Ibid., 183. In this final (unfinished) work, Kant was to some extent responding to Schelling (184), in the sense that he posits “a deep connection between the materiality of the subject and the materiality of nature.”

not mental operations performed on sense data but are the formal structure of spatio-temporal relations in which objects stand in relation to the body.”³⁰ As such, this first stage of Kant’s theoretical system sets out the basis for “a rudimentary physics” that is primarily *ontological* and thus *physical* in its emphasis, not merely epistemological and mental.³¹

One of the most important implications of the Transcendental Aesthetic is that, as far as our pre-conceptual experience is concerned, so-called “self-knowledge” does not occupy any privileged position in relation to our knowledge of the external world.³² Descartes had argued, of course, that our knowledge of the external world is mediated through our senses and therefore always open to doubt, whereas our knowledge of our own thoughts is immediate and therefore indubitable. By contrast, Kant’s theory of space and time as forms of intuition leads him to distinguish between what he calls *outer sense* and *inner sense*. Insofar as introspection gives us access to perceptions and/or thought-processes that *occur in time*, Kant portrays self-knowledge as empirical and acknowledges that it, too, may well turn out to be mistaken.³³ We may *think* we observe ourselves believing or desiring a certain thing, but evidence to the contrary might emerge, demonstrating that our presumed self-knowledge is not infallible after all. This crucial insight has important implications for Kant’s later, moral and religious writings, where he insists that our knowledge of a person’s inner conviction (*Gesinnung*) and of the motives that make it what it

30 Ibid., 185.

31 Rukgaber admits that Kant’s position has elements of both. He backs up his radical claim by pointing out that in his 1783 sequel to the first *Critique*, appropriately titled *Prolegomena to Any Future Metaphysics that will Present itself as a Science* (bold type added), Kant explicitly claims that the arguments of the Transcendental Aesthetic answer the question: “How is nature possible in general *in the material sense . . . ?*” (*Prolegomena*, 318; emphasis added). Kant’s answer, according to Rukgaber (“The Key to Transcendental Philosophy,” 186), is “that our perspective, part of which is our being embodied material beings, is a condition on the possibility of nature itself.”

32 Churchland briefly mentions this interesting point (in *Neurophilosophy*, 248-249), but does not develop it.

33 Kant refers to perceptual illusions on several occasions; see e.g., B69-70, A295/B351-352, A297/B353-354. At one point, he explicitly states that the situation regarding inner sense and outer sense is “exactly the same” (B67); tr. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998). That his discussion of the senses is meant to be understood as a discussion of the *body* is implied in numerous passages, such as when he describes “a representation of sense” as a “force of nature” (A294/B350). In direct opposition to Descartes, he insists that “the senses do not err” (A293/B350); for as part of the physical mechanism as nature, “they do not judge at all.”

is, *even in the case of our own motives*, is never perfectly clear and distinct³⁴ – as Descartes and the whole rationalist tradition had taken such inner observations to be. While we will not be able to explore the implications of this point for moral philosophy here, it is important to note that Kant's doubts regarding the moral agent's ability to judge his or her own self-worth are grounded in what can only properly be regarded as a *monistic* view of human sensibility: although we do have *two forms* of sense, inner and outer, *both* are necessarily and universally subjected to the spatio-temporal form of all human intuition that defines the human standpoint; any knowledge-claims made on the basis of either one must therefore be contingent and open to revision.

The implications of this *bodily* grounding of all mental functioning can be traced throughout virtually every step of Kant's theoretical system. In the next main section of the first *Critique*, the Transcendental Analytic of Concepts, Kant appeals to the human imagination as the power that enables us to form concepts out of the raw material presented to us by what he calls "the manifold of intuition" – i.e., the unorganized content that is the pre-conscious predecessor of our conscious perceptions. If we accept that this "manifold" refers to our *bodily* functioning, then Kant's arguments take on a distinctly anti-Cartesian flavor. In the first edition, Kant describes the following three-step process (A94): "(1) the *synopsis* of the manifold *a priori* through sense; (2) the *synthesis* of the manifold through imagination; finally (3) the *unity* of this synthesis through original apperception." The first step is entirely physical, for as we have seen, the *a priori* forms (space and time) arise out of our body's encounter with the world around it. The second step argues that our bodily sensations must be weaved together in such a way that we form images of them; nothing in Kant's text prevents us from seeing this process as entirely a function of the brain. In the second edition Kant distinguishes between two aspects of this second step: the "productive" and the "reproductive" imagination. The purpose of this distinction is to point out that our mind-brain processes these sensation-based images in two very different ways: first, by *producing* them; second, by *re-producing* them, in the form of what we normally call "memory." Neuroscience has provided strong evidence that the brain has a great deal to do with both processes.

The third step will be the basis, toward the end of this article, for a challenge I will pose to EM. The synthesis of bodily sensations to produce images, that in turn form the content of our memory, does not suffice to produce knowledge

34 In *Religion within the Bounds of Bare Reason* (1793/1794), for example, Kant says "the depth of the heart (the subjective first basis of [a person's] maxims) is inscrutable" (51; tr. Werner S. Pluhar [Indianapolis: Hackett, 2009]).

of objects all on its own. For that, we must bring this synthesis to a “unity,” a requirement that is fulfilled by a power Kant calls “apperception” – i.e., a metaphorical “perception” of one’s self as a knowing subject. That Kant associates this power with our sense of “I,” even employing Descartes’ term, “ego,” is one of the primary reasons so many readers merely assume he is adopting Cartesian dualism. But this common way of reading Kant’s theory of apperception ignores the fact that, from its first introduction and throughout his theoretical system, the kind of apperception Kant defends is *transcendental*. By calling this “I” the *transcendental ego*, Kant means to imply that there is nothing substantial about it; that is, it does not exist in space and time, so it is not a mental or “soulish” correlate to our physical existence. Like every a priori element of his theoretical system, the transcendental ego is *ideal*; this means that, as far as our physical existence is concerned, its status is that of a *necessary presupposition* – a concept we impose onto our experience in order to *explain* what we have come to know. In this case, it is the presupposition that every image produced by my imagination (which might end up meaning: by the relevant part of my mind-brain) is *my* image and, when connected to a concrete synopsis (cf. the first step), *my perception*.

The first chapter of the first *Critique*’s Transcendental Dialectic explicitly *rejects* Descartes’ theory that the soul is a substance, equal and opposite to the body. Many interpreters have assumed that Kant’s arguments in the Dialectic reject the idea of the soul altogether. However, Julian Wuerth persuasively demonstrates that those arguments focus on attacking the *substantiality* of the soul and that Kant himself consistently affirms a belief in the *other* basic characteristics of the soul throughout his *corpus*.³⁵ This does not necessarily contradict the argument I am presenting here, but merely raises the question: if the soul is *not* a spiritual “substance,” then *what is it?* Kant says that if this refers to the soul “in itself” (A684/B712), then “the question would have no sense at all.” My claim is that EM’s response (i.e., that the soul is an illusion created by certain brain states that have yet to be discovered and/or completely understood) is not inconsistent with Kant’s position, at least as an account of what a *science* of the soul might look like. For in a similar way, Kant replaces the Cartesian soul with what he calls the *empirical ego*. A much-neglected implication of Kant’s transcendental idealism is that this empirical ego is subject to the limiting conditions of time and space just as much as are our bodily sensations. Kant himself tends to assume that it is more closely connected to our inner sense than

35 Julian Wuerth, “The First Paralogism, its Origin, and its Evolution: Kant on How the Soul Both Is and Is Not a Substance,” in Stephen R. Palmquist (ed.), *Cultivating Personhood: Kant and Asian Philosophy* (Berlin: Walter de Gruyter, 2010), 157-166.

to our outer sense, and therefore treats the temporal limitation as more basic than the spatial limitation. Admittedly, this could pose a problem for someone who wished to defend both Kant's philosophy of mind and EM. But I suspect Kant, if presented with this potential conflict, would give a response entirely consistent with EM: we must leave the outcome of this issue to be decided by empirical research.

Interestingly, one inference that Kant himself drew from his theory of the empirical ego, for which he has often been criticized, actually renders his position *more* amenable to EM: Kant denied that empirical psychology could ever become a science.³⁶ As various critics have pointed out (see note 36), Kant's position on this issue did not make sufficient room for the possibility of either a behaviorist approach or approaches focusing on empirical manifestations of the so-called *unconscious*, such as the Freudian and Jungian use of *dream texts* as objective content for scientific analysis (see note 6). But his main point was that *introspection* cannot be the basis for a reliable empirical psychology, for the very reason that it suffers from all the contingencies of any spatio-temporal experience, yet lacks (on its own) the external manifestation that enables scientists to test and verify it using objective methods. (I shall comment further on the latter possibility, below.) Believing he had destroyed the foundation for a science of *rational* psychology by refuting Descartes' notion of a substantial soul, Kant assumed that the only thing left for *empirical* psychology to research would be the contents of human introspection, which could not, as such, ever be accessible to external observation, as is required for a science to be empirical. As we have seen, the primary claim of EM is that folk psychology is a *theory* (i.e., an attempt to regard desires and beliefs as *scientifically justifiable explanations* of human behavior) that is weak and unreliable, and that it will therefore probably end up being replaced some day by a complex, neurologically-based account of why we do what we do. For my purposes here, it will suffice merely to express an opinion: Kant's rationale for predicting that psychology (as introspection) would never establish itself as an empirical science is compatible with, if not identical to, EM's prediction that folk psychology will eventually be reduced to neuroscience.

36 Kant's most explicit denial occurs in the Preface of his *Metaphysical Foundations of Natural Science* (1786); his discussion of empirical psychology in the first *Critique* is more open-ended (A848-849/B876-877; see also A347/B405-406 and A682-684/B710-712). Thomas Sturm, "Kant on Empirical Psychology: How Not to Investigate the Human Mind," in Eric Watkins (ed.), *Kant on the Sciences* (New York: Oxford University Press, 2001), 163-184, argues that Kant's denial is based on a very narrow conception of what constitutes a *natural science*, and on a limited view of what empirical psychology entails.

Backtracking from the Dialectic in the text of the first *Critique*, we find several other key aspects of Kant's theoretical system that are explicitly body-based.³⁷ Of these, by far the most explicit comes toward the end of the *Analytic of Principles*, in the section called the Refutation of Idealism. Kant there argues that our self-awareness arises first and foremost out of our *bodily experiences*, and that our awareness of having a mental life arises only as an offshoot from this physical grounding. Kant added this section in the second (1787) edition in order to refute early interpreters who had mistakenly identified his *transcendental* idealism with Berkeley's *empirical* idealism. According to the latter, the very existence of the physical world can be called into question, because we have intimate knowledge only of our own mental life. But for *Kantian* idealism, exactly the opposite holds: we know the *physical* world most intimately, so if we wish to doubt either side of Cartesian dualism, the empirical reality of our *mental life* must fall by the wayside. Here Kant's position virtually foreshadows EM.

A likely objection to this way of reading Kant is that the second *Critique* presents the noumenal world as the basis for our *moral* life, and morality as the source of the meaning of human life in general, so Kant could not have countenanced EM's total *reduction* of the mental to the physical. A complete response to this objection would be too lengthy to be attempted here. But in a nutshell, a Kantian defender of EM might respond that in the second chapter of the first *Critique's* Dialectic (in the Third Antinomy), Kant acknowledges deep problems in treating freedom as a *cause* within the nexus of the empirical world. Whatever Kant meant by "free will" in the second *Critique*, and by the claim that it is the one "fact" of *practical* reason, he did not mean that noumenal causality *interrupts* the phenomenal causality required by empirical science. Our moral convictions, however their noumenal standpoint is to be philosophically justified, do not give us license to disrupt the causal nexus that the *Analytic of Principles* demonstrates to be a necessary condition for empirical knowledge. Instead, Kant's justification would begin by reminding us that even the transcendental principles that enable us to explain our empirical

37 One of the most interesting of these key features, though beyond the scope of this article to explore, is Kant's theory of the twelve categories and their schematization as "principles" that define the mental presuppositions we must adopt, in order to explain what we know about the physical world. Typically regarded as the most immaterial stage of Kant's entire theoretical system, this theory lies at the very core of transcendental idealism's grounding in the body, argues Svare (*Body and Practice in Kant*, especially chapters 9-11).

knowledge are *ideal* (i.e., not empirically based), so that in this respect the phenomenal is neither more nor less securely grounded than the noumenal.³⁸

Two Challenges to Eliminative Materialism

In concluding these reflections on Kant's body-centered approach in the first *Critique*, we must recall where Kant started in the first Preface, by acknowledging the *inevitability* of metaphysical illusions. To avoid being *fooled* by such illusions, Kant says we must learn to think perspectivally. This means recognizing that, when we create ideas, they can carry a heavy load of meaning from one (explanatory) perspective even though from another (physical or substantive) perspective, the object of such ideas might not literally "exist." Once we grasp the perspectival nature of Kant's affirmation of the body, we can recognize that his position challenges EM in at least two ways.

First, the passage quoted in the second section, above, from *Dreams of a Spirit-Seer*, suggests that Kant would be extremely skeptical about EM's basic prediction ever coming to pass, as long as scientific research focuses exclusively on the *brain*. For Kant thought it was a fundamental mistake to assume that "I" somehow *live in my brain*. The brain's functions may be – nowadays we can say *are* (at least as far as we know) – essential to the experience of what we call "mental states"; but to hold out the hope that the mind will someday be *reduced to the brain* is no more plausible today than was the hope of *locating the soul in the brain* in the days of Descartes and Kant. Ironically, on this point the typical eliminative materialist is in danger of looking more like Descartes than Kant does! The first challenge to EM, therefore, is to recognize that the *whole body* is the source of our feeling of having a distinct "mental life," so that unless neurology is extended and supplemented by an understanding of how each part of our body complements and feeds into the brain in such a way that *the body controls the brain just as much as the brain controls the body*, the prediction that forms the basis of EM is doomed to remain unrealized. In other words, it is not and never will be *the brain on its own* that turns out to be smarter than it is complicated, but rather, the brain and body functioning as a creative feedback loop and thereby constituting the *whole person*.

Churchland reports a 1950 experiment with monkeys that may have some bearing on this first challenge. Scientists located an area of a monkey's brain

38 For a perspectival interpretation of Kant's infamous theory of noumenal causality, see section IV of my article, "Kant's Ethics of Grace: Perspectival Solutions to the Moral Problems with Divine Assistance," *The Journal of Religion* 90, 4 (October 2010): 530-553.

that controlled thumb movement and removed it, producing temporary paralysis of the thumb; when the monkey eventually regained thumb mobility, researchers found that the cortical tissue surrounding the initial lesion had taken over this function; researchers then removed the newly modified tissue to produce a second paralysis and the same modification of surrounding areas gradually occurred again as thumb mobility was restored. The conclusion reached by such studies, Churchland reports, was that “strict localization of function was doubtful.”³⁹ Taking on board the first Kant-inspired challenge to EM would require neuroscientists at least to consider the possibility that the monkey’s *thumb* influenced the brain and caused its modifications, not merely the brain on its own adapting to the loss of tissue. From Kant’s Critical standpoint, it is not only the brain that controls the thumb or the thumb that controls the brain (a matter that is to be resolved by empirical research), but above and beyond both of these causal influences, *I* (i.e., the possibly illusory yet necessary idea of my whole *body* as a person) control them both.

The second challenge relates to how EM portrays the mind or mental events, once they arise. The strict defenders of EM appear at times to be claiming that, once neuroscience reaches its heyday and the brain is sufficiently understood, we will realize that *people do not have beliefs and desires*. This is the source of the many attempts that have been made to argue that EM is fundamentally self-contradictory or self-refuting.⁴⁰ What Kant’s system of perspectives (especially as interpreted in the context of *Dreams of a Spirit-Seer*) can contribute to this debate is a realization that the mind or soul or “I” that arises out of the body *is entirely a product of imagination*, yet this imaginary object is the very thing that makes possible our ability to explain what we know to be real. The latter claim, of course, is unlikely to be attractive to most defenders of EM. When considering the possibility that the mind-brain relation could be analogous to the wave-particle relation in the physics of light, for example, Churchland declares that in order for this analogy to hold, “it would have to turn out that mentality is somehow a *fundamental property of matter*, and for that there is not the

39 Churchland, *Neurophilosophy*, 171.

40 For an excellent explanation of why transcendental or quasi-transcendental arguments against EM are bound to fail, see Kenneth Taylor, “How Not To Refute Eliminative Materialism,” *Philosophical Psychology* 7, 1 (June 1994): 101-125. Toward the end of his essay, Taylor offers this life-line to anyone who still thinks a transcendental argument against EM is possible: a transcendental argument might succeed if it claimed “that folk psychology does not, after all, purport to be a robustly causal theory and that supposing otherwise leads one to incoherence.”

smallest shred of evidence."⁴¹ This claim is somewhat shocking, coming from someone so well-versed in the details of empirical science. For in the field of quantum physics it is almost commonplace to find such a possibility affirmed, even among the most respected physicists.⁴² Indeed, taking quantum mechanics into consideration would render my argument even more plausible; hence, a version of EM's prediction that takes on board this second challenge would be that future advances in both neuroscience *and* quantum mechanics will enable these two disciplines to draw on each other's insights in such a way as to reveal that, at the very deepest level of the brain's operation, brain states are an expression of free-flowing energy that will turn out to be the best possible scientific explanation of what "mental" really means – but not in a way that destroys the legitimacy of *referring to* mental states as explanatory concepts. Establishing this may not require proof that mentality exists as a fundamental *property* of matter, but it will require evidence that mentality exists as a fundamental *perspective* of matter – e.g., that quarks somehow "experience" freedom.⁴³

Whatever EM means by stating that folk psychology will be reduced to neuroscience (or, in light of the first challenge, to the holistic interplay between neuroscience and human anatomy in general), it cannot mean that *we no longer have* any of the things we now call beliefs and desires; it can mean only that in the post-reductionist era, all well-educated persons will realize that, as Kant demonstrated, *all* our ideas (when viewed from the theoretical standpoint) have to be regarded as illusions as far as their relevance to empirical science is concerned, even though these same ideas remain no less meaningful and no less true (when viewed from the practical standpoint) as far as their relevance to human morality is concerned. Only by accepting this second challenge and interpreting EM perspectivally can its proponents effectively respond to criticisms such as Taylor's (see note 40). Admittedly, in this article I have not

41 Churchland, *Neurophilosophy*, 375.

42 To cite but one of many possible examples, Freeman J. Dyson states in his 1985 Gifford Lectures, *Infinite in All Directions* (New York: Harper & Row, 1988), 8: "Therefore, I say, speaking as a physicist, scientific materialism and religious transcendentalism are neither incompatible nor mutually exclusive. We have learned that matter is weird stuff."

43 For further details on how quantum mechanics can be consistent with Kant's philosophy, see my articles: "Kantian Causality and Quantum Quarks: The Compatibility between Quantum Mechanics and Kant's Phenomenal World," *THEORIA: An International Journal for Theory, History and Foundations of Science* 28, 2 (May 2013): 283-302; and "Bohm's Quantum Causality and its Parallels in Kant's Ideas of Reason", in *Death and Anti-Death, Volume 13: Sixty Years After Albert Einstein (1879-1955)*, ed. Charles Tandy (Palo Alto, Ca.: Ria University Press, 2015), 99-128.

presented a detailed account of exactly how an argument for limiting the scope of EM should be constructed. What I have argued, however, is that such an argument should start from a Kantian (perspectival) form of non-dualism and that its goal should not be to *overthrow* EM but to moderate its overly extreme claims through perspectival reasoning.

Taking on board these two Kant-inspired challenges would not require the rejection of any essential feature of EM. What it does require is a subtle yet profound revision of the way its reductionist aims are typically presented to the public, a turn not unworthy of being called “Copernican”: instead of viewing neurophilosophy *merely* as a program for *reducing* the mind in general and conscious mental processes in particular to nothing but raw neuroscience and the brain functions it describes, proponents of EM must also *raise neuroscience to the level of the mind*. The 1999 movie, “Bicentennial Man,” serves as an interesting thought-experiment regarding what might be involved in such a reversal of the reductionist’s theoretical paradigm. In it, the lead character, a robot (played by the late Robin Williams), decides after about two hundred years of service to a series of human masters that he wants to become human. In so doing, he does not need to change the way his neural circuitry is constructed (because the futuristic science of his day had already made him remarkably human-like). Rather, he must change certain *illusions* about what his existence entails; most poignantly (especially given the actor’s recent tragic death), he must accept the reality that as a human, he must die. Once he agrees to raise such illusions (motivated by his love for another) to the level of transcendental ideas, his status as a bona fide human being is confirmed.

Only by affirming the meaningfulness of explanatory idealism, of the sort Kant defends under the rubric of the “transcendental,” can one consistently *believe* or *assert* that EM is true. In order to *explain* their position, eliminative materialists will then be *compelled* to recognize that the matter cannot be quite so monochrome as extreme physicalists typically assume. This perspectival shift does not compromise our ability to take seriously whatever empirical results the science of the future may bring our way;⁴⁴ but it will prevent us from succumbing to the Cartesian (*and* extreme EM) error of treating the

44 An interesting research project would be to produce a computerized simulation of infant learning according to Kantian parameters. The project could be designed so that its assumptions were compatible with those of EM, insofar as the simulated environment would not distinguish between brain states and mental states. The key to making such a project “Kantian” would be to give the simulated agent not only a human-like set of neurological structures, but also a growing sense of self-awareness through perception of its *whole body* as a unit out of which its experience arises.

(non-brain) body as something the mind-brain somehow *owns*. No philosopher urges us more than Kant does that we are to be ruthlessly rigorous in giving science its proper place and in therefore being open to the results of empirical inquiries; yet in so doing, the role of a good *philosopher* of science is to ensure that we can encourage such a tough-minded approach without adopting background assumptions that would imply we have lost our minds altogether.⁴⁵

Biography

Stephen Palmquist (D.Phil., Oxford, 1987) is Professor of Philosophy and Religion at Hong Kong Baptist University. His more than 160 publications include multiple articles in 11 journals: *Journal of Chinese Philosophy*, *Kant-Studien*, *Philosophia Mathematica*, *Journal of Religion*, *Review of Metaphysics*, among others. His 10 books include *Kant's System of Perspectives* (1993), *Kant's Critical Religion* (2000), *Cultivating Personhood* (2010), and *Comprehensive Commentary on Kant's Religion within the Bounds of Bare Reason* (2016).

45 An earlier version of this article was presented as an invited paper at the Second International Conference on Comparative Studies in Mind, organized by the Institute for Chung-Ang Philosophical Studies and held at Chung-Ang University in Seoul, South Korea, in December 2011. I would like to thank Kwon Jong Yoo and Ho Young Lee for organizing the conference. Thanks also to Patricia Churchland and other conference participants for their helpful feedback on my initial presentation, as well as to Aaron Bunch for feedback on an earlier draft and to Richard Mapplebeckpalmer for introducing me to several of the sources used in this article.