# *The Extended Mind and Religious Thought*

*with Leslie Marsh, "Mindscapes and Landscapes: Exploring the Extended Mind"; Mark Rowlands, "The Extended Mind"; Lynne Rudder Baker, "Persons and the Extended-Mind Thesis"; Teed Rockwell, "Minds, Intrinsic Properties, and Madhyamaka Buddhism"; Joel W. Krueger, "Empathy and the Extended Mind"; Leonard Angel, "Quintuple Extension: Mind, Body, Humanism, Religion, Secularism"; and Matthew Day, "Constructing Religion without The Social: Durkheim, Latour, and Extended Cognition"*

## MINDS, INTRINSIC PROPERTIES, AND MADHYAMAKA BUDDHISM

*by Teed Rockwell*

*Abstract.* Certain philosophers and scientists have noticed that there are data that do not seem to fit with the traditional view known as the Mind/Brain Identity theory (MBI). This has inspired a new theory about the mind known as the Hypothesis of Extended Cognition (HEC). Now there is a growing controversy over whether these data actually require extending the mind out beyond the brain. Such arguments, despite their empirical diversity, have an underlying form. They all are disputes over where to draw the line between intrinsic and relational causal powers. The second-century Buddhist philosopher Nagarjuna deals with similar issues when he argues for a middle way between the two positions that were known in his time by the terms *eternalism* and *nihilism*. Eternalism, like MBI, asserts that the mind is a permanent enduring substance (although the two theories disagree as to how long mind endures). Nihilism argued that the mind had no intrinsic existence, and today some argue that HEC could lead us to a similar conclusion. Nagarjuna's argument for a middle way between these two extremes is similar to an argument that can be made for HEC. We can accept that neither the brain nor any other single physical item is identical to the mind without falling down the slippery slope that leads to "The mind does not really exist, and therefore we are one with everything." Nagarjuna was correct to say that the mind has conventional reality—that the mind exists even though there is no sharp border between the mind and the world.

Teed Rockwell is Lecturer in Philosophy, Sonoma State University, Nichols 363, 1801 East Cotati Avenue, Rohnert Park, CA 94928; email mcmf@sfo.com.

[*Zygon*, vol. 44, no. 3 (September 2009)]
 ISSN 0591-2385

*Keywords:* extended mind; hypothesis of extended cognition; mind/brain identity theory; Nagarjuna

---

Those of us who defend the Hypothesis of Extended Cognition (HEC) get criticized from two different perspectives that, to use a political metaphor, could be called *conservative* and *radical*. Because HEC was born in the cognitive science community, most of the criticism comes from epistemological conservatives—those who want to conserve the idea that the mind is best described as being in some sense identified with the brain. These critics want to be assured that there is some place where the mind stops and the world begins and believe that the brain is the best place to draw the line. Outside the orthodox cognitive science community, there are readers from the radical epistemological "left" who welcome HEC as some version of the claim that we are "one with everything." The most articulate and cautious of these radicals is David Skrbina, who argues that if I were to follow through with my own logic I would accept "a kind of full-blown panpsychism" (Skrbina 2006). It is possible that I could be persuaded to agree with Skrbina about this, depending on how we define our terms and what level of reality he is willing to grant to discrete individual minds. That, however, is a topic for another time. In this essay I concern myself with those who see my position (whether approvingly or disapprovingly) as a kind of muddled monistic mysticism. These casual readers serve an important function in the debate by providing a *reductio ad absurdum* argument against HEC for the conservatives. If HEC really required us to abandon all distinctions between mind and world, it could not be the next paradigm in cognitive science. On the contrary, it would require us to abandon cognitive science altogether.

One reason that my version of HEC sometimes receives this radical interpretation is that I believe the mind is best described as a behavioral field rather than a single item such as a brain or a body. There is also the fact that I occasionally describe this behavioral field with somewhat evocative language that might be appealing to the radicals, such as "Consciousness could be a pattern which, like a vibration started by throwing a stone in the water, ripples through the world even though there is a biological creature at its center" (Rockwell 2005, 103). However, it is my intention to position myself in a kind of middle way between these radical and conservative extremes, even though my position is more radical than that of some other HEC theorists. For example, Andy Clark's version of HEC tries to give fairly hard and fast criteria for identifying the mind with certain kinds of external cognitive "scaffolding," such as the notebook that aids the memory of someone with Alzheimer's disease (Clark and Chalmers 1998, 17). Unlike him, I believe that drawing a single line between the self and the world outside the brain is probably even more misleading than

trying to draw the line at the brain. Consequently, I think we should abandon the idea that there is a single place where the line can always be drawn. This is what makes some of my readers accuse me of rejecting "the analytic distinctions of self and world" (McCarthy 2006, 85) and thus embracing the radical "we are one with everything" position.

This is a misinterpretation, however, because I also insist that "To say that the mind emerges from the brain-body-world nexus does not mean that there is no world, only a mind. The line between the self and the world must always be drawn somewhere. . . . That is what it means to live in a world" (Rockwell 2005, 104). I do not identify the mind with the entire brain-body-world nexus because I believe that the line between the self and world must be drawn somewhere *at any given moment.* But this does not necessarily imply that there is a single place that the line can be drawn for all conscious creatures, or for a single conscious creature throughout its history. A great deal of useful scientific work can be done by drawing the line at the skull, but the books that defend HEC describe scientific work that needs to draw the line in a variety of other places. I think the best way to account for both mainstream neuroscience and this other more problematic work is to see the boundary between self and world as flexible. That is why I feel the mind is best described as a behavioral field rather than as an organ in the skull.

In many ways, these two reactions to HEC are similar to the conflicting positions of Hindus and Buddhists in the debates about mind in second-century India. The radical mystical position is similar to the position called *nihilism*, and the conservative mind/brain identity theory is similar to the positions called *eternalism.* There were two eternalist positions represented in that debate, one Hindu and one Buddhist. For the Hindus of that time, mind was Atman, the eternal oversoul that was real, enduring, and permanent. The Buddhist form of eternalism, *Abhidharma*, claimed that although minds and other substances are not ultimately real, there are ultimately existing entities called *dharmas*, which are the fundamental elements that make both mind and world possible. The second-century Indian philosopher Nagarjuna, founder of the Madhyamaka school of Mahayana Buddhism, said that Abhidharma Buddhism did not go far enough, because he believed that "the set of ultimately existent things is an empty set" (Arnold 2005). Nagarjuna and his commentators used the argument called "the interminability of dependent origination" (Arnold 2005) to argue that not even the dharmas can be ultimately existent and eternal because each thing is dependent for its existence on something else, which is in turn dependent on something else, and so on *ad infinitum*. However, if the mind is not eternal, this implies that there is no afterlife and that all of the rules of karma and morality do not exist—the position that was called nihilism. Nagarjuna's middle way was so called because it rejects both extremes, nihilism and eternalism.

To identify the mind/brain identity theory (MBI) with eternalism may seem almost paradoxical. The MBI is a materialist position that argues that the mind is a single lump of protein in the skull. Eternalism, not unlike Platonism, argues that the mind is an eternal abstract principle that transcends the physical realm. Nevertheless, MBI does claim that the mind endures the way an Aristotelian substance endures and has properties and borders that define it as a distinct entity. Because the MBI theorist believes that a mind is nothing but a lump of flesh between the ears, each mind will pass out of existence eventually. Nevertheless, the assumption is that while an organism lives, its brain possesses the intrinsic mental properties that are sufficient for consciousness and cognition. Both the modern debate and the second-century Indian debate thus are concerned with whether the mind has intrinsic qualities. The eternalists and MBI theorists both think it does, although they have very different ideas as to what those intrinsic qualities consist of. The ancient nihilists and the modern "one with everything" crowd think that it doesn't and therefore conclude that the mind does not exist as a separate entity.

There is a middle way between unbridled holism and the mind/brain identity theory. We can acknowledge that the mind has what Nagarjuna called conventional reality, even if it is not a permanent enduring substance with a "stable fixed essence" (Berger 2007). A great deal of useful scientific work can be done by drawing the line between mind and world at the skull, but the research described in the pro-HEC texts needs to draw the line in a variety of other places. We can respect both kinds of scientific work by recognizing that there are many distinctions between mind and world that have Nagarjuna's kind of conventional reality. Most HEC theorists argue that the exact borders between the mind and the world cannot be sharply drawn at the skull. I go one step further and argue that there is probably no single place where the borders can be sharply drawn. However, I also say that there are pragmatic reasons for the various places the border is drawn, and therefore these fluctuating borderlines should be accepted as conventionally real even if they are not ultimately real.

## The Data in Question

Most research on the biological embodiment of mind consists of establishing correlations between neural events and behavior. This is obviously well worth doing, because there are a lot of those correlations, and knowing about them gives us power over all sorts of mental afflictions. However, sometimes the neuroscience by itself seems to tell only a fragmented and incomplete story. For such data, it seems necessary to also refer to other events in the body and/or world to make the neural events seem cognitive at all. The foundation of most books on HEC is the detailed description and interpretation of these kinds of data. Part of the appeal of HEC is that

there really are a lot of these data, and they cry out for some kind of new theory to explain them. For this article I consider a few examples of the data that make up the bulk of chapters 2, 3, and 10 of Rockwell 2005. These examples should be sufficient to illustrate the underlying form that can be found in almost all of the reinterpretations inspired by the insights of HEC: a shifting from a description of intrinsic properties of neurons and neural networks to a description of the relational properties that emerge from a more extensive brain-body-world network.

In 1985, neurochemists F. O. Schmitt and Candace Pert discovered what they called a parasynaptic system, which operates independently of any neural networks by using chemicals they call information substances. These information substances communicate without synaptic firing because they are released into the extracellular fluid and then float until they reach a receptor molecule with which they are compatible. Thus, even though no neural network connects the two organs, the specificity of the receptors makes it possible for them to exchange information chemically. For example, there is a neuropeptide called angiotensin for which there are receptors in both the limbic system (a neural structure located in the skull) and the kidneys. There is evidence that by releasing and receiving angiotensin the limbic system and the kidneys create a balance between thirst and the body's need for water (Pert 1987, 83). Activity in the limbic system also has been found to correlate with a variety of other emotional states, and it contains very high concentrations of neuropeptide receptors.

Many of these neuropeptides include substances that, when secreted by the endocrine glands, are called hormones. Insulin, traditionally considered to be a secretion of the pancreas, is also made and stored in the limbic system, as are its receptors. That the endocrine glands and the limbic system share what is in effect a common chemical alphabet makes it highly plausible that they communicate with each other chemically without having to rely on synaptic connections. This is one of many facts that led Pert to conclude the following: "In the beginning of my work, I matter of factly presumed that emotions were in the head or the brain. Now I would say they are really in the body as well. They are expressed in the body and are part of the body. I can no longer make a distinction between the brain and the body" (1987, 84). Why did Pert feel compelled to make this claim? Because it is the relationship between the nervous system and the rest of the body that produces the emotions. Although the patterns intrinsic to the nervous system are clearly necessary for the experience of emotions, they are not sufficient to produce the emotions by themselves.

The connection between the kidneys and the limbic system is one example of this kind of communication, but after Pert's 1987 article was published several others were discovered. The vasoactive intestinal peptide can be found in the frontal cortex of the brain, the thymus gland, the gut, the lungs, some immune cells, and parts of the autonomic nervous system

(Pert 1997, 208). Pert speculates that because the frontal cortex is the part of the brain associated with long-range planning, the chemical connection between the gut and the frontal cortex could partially account for the cognitive efficacy of "gut feelings." We cannot see most of this communication if we study the intrinsic properties of the nervous system, which by definition includes only synaptic connections. To account for the patterns constituted by these chemical exchanges we have to think in terms of a system of relations that includes the nervous system, the endocrine system, and the immune system (the spleen, bone marrow, and lymph nodes).

From this perspective it becomes possible to give concrete physiological detail to William James's proposition that emotions exist in the body, not in the brain. James saw emotions as analyzable into a constellation of particular physical feelings (James 1890, 443–49). Apparently the release of a particular neuropeptide can produce a particular set of sensations in certain organs and muscles because of the existence of receptors in their cells, and that specific neuropeptide is released when certain emotions are felt. Norepinephrine, for example, is released from the locus coeruleus in the hindbrain, and when that region of the brain is electrically stimulated it produces strong sensations of pleasure in rats and humans (Pert 1997, 138). If the cells in the body contain different amounts of different kinds of receptors for each ligand, and a different ligand (or combination of ligands) is released for each emotion, it would be reasonable to say that this chemical exchange of information embodied that emotion. This would mean that emotions as a class of mental phenomena would actually exist in the entire body, not just in the brain. To say that one felt a certain emotion would mean not only that one was in a certain brain state but also that the cells distributed throughout one's body had received the particular dosage of ligands that was responsible for that emotion. The emotion therefore is not intrinsic to a closed system of neural networks but emerges as a relational property of the nervous system and the rest of the body.

## Is Causation Different from Embodiment?

One common response to such an argument is that these various hormonal activities may *cause* mental states, but they do not *embody* mental states. If we want to explain Pert's data and still accept the MBI theory, we can say that the angiotensin in the kidneys caused a change in the limbic neurons, but the resulting neural activity embodied the pain. Similarly, if there is a chemical interaction between the gut and the forebrain during the experience of gut intuitions, this could be described by saying that the chemical interactions in the gut caused the intuition, but the neural activity in the forebrain embodied the intuition. In any attempted counterexample, we can easily draw a distinction between the nonneural activity that causes mental states and the neural activity that embodies them. The

mind–nervous system identity thus appears to be safe from any critique of this sort.

Such unassailability should be viewed with deep suspicion, however, because it is produced by dogmatic and arbitrary proclamation. This distinction can never be either confirmed or refuted because everything we know about the nervous system is gathered by establishing causal relationships between neural states and mental states. We know that damaging Broca's area causes aphasia, so we assume that this part of the brain embodies our ability to speak. If we learn that angiotensin traveling from the kidneys to the limbic system causes thirst, why assume that only half of that process embodies the thirst and the other half "merely" caused it? Why stop there, if we are going to permit ourselves such ad hoc divisions? Why not say that most of the brain only causes mental activity but that there is a small part of the brain that actually embodies consciousness, such as the pineal gland? Do we really want to base a theory on such arbitrary foundations? As Robert Wilson (2001) rightly points out, we could use the same kind of reasoning to say that a single neuron could embody almost any mental state that it helped to cause. How can we choose any of these alternatives over the others except to claim that one of them (somehow) satisfies our so-called intuitions?

One could try to dispose of this problem with a *reductio ad absurdum*. If we abandoned the MBI theory, what would stop us from saying such counterintuitive things as "Because the tree I am observing outside my window is causally connected to my psychological state, my mind is partially instantiated by that tree"? But there are far more arguments for biting such a bullet than one might first suppose. Such a conclusion would not be resisted by Berkeley, Kant, or the philosophical traditions of Hinduism and Buddhism. And J. J. Gibson, a controversial but still respected psychologist, had a theory of perception that required the moving of mental embodiment not only outside the nervous system but completely outside the skin.

If Gibson's theory of perception is correct, our visual experience is embodied not only by the nervous system but also by the patterns of light in the perceiver's environment. In a 1955 lecture Gibson attacked the idea that "Perception is supposed to be based on sensation." He figured that because there was so much information in the array of light in our environment, there was no reason that every bit of that information had to be copied by an analogous sensation in the brain. He claimed that if there is no such analog, we could think of the perception as "direct," because the sensation was not getting between the stimulus and the perceiver. Consequently, Gibson claimed that as far as perceiving the outside world was concerned, we could "dispense entirely with the concept of sensation." He did qualify this bold claim by admitting that the concept of sensation could be useful for studying certain phenomena such as visual afterimages. But he considered these phenomena to be "a psychological curiosity," not a

fundamental component of those perceptual processes that make us aware of the world.

Is there any way of proving Gibson's theory of perception? Suppose some neuroscientists of the future are studying the brain activity that occurs when a laboratory subject has two distinctly different phenomenal experiences, such as seeing two slightly different shades of green. The shades have been identified as being spectrally different from each other, and the subject correctly identifies each of them when they are presented to her. However, a thorough brain scan reveals no differences in brain activity that can account for her ability to make the distinction. Many would insist that there must be a significant difference somewhere in brain activity, because otherwise we would have to accept that she made the distinction by means of some psychic power. However, this dichotomy is forced on us only if we assume that embodiment stops at the skin. If the differences are captured in the light array, which is physically connected to the brain, no psychic powers would be necessary—if we could accept that the light array partly embodies the sensations. The experience would be an emergent property of the relationships between the neurons and the light rather than a property of the neural network in isolation.

The sheer weirdness of such a discovery would probably trigger a willingness to consider an HEC-based paradigm. However, other widely accepted facts support HEC with similar effectiveness. What is disturbing about this thought experiment is that it seems to show that we could change consciousness without changing brain states. If someone proved that changes in brain states are not necessary for changes in mental states (and there is no reason to automatically dismiss that possibility), this would be a decisive death blow for the MBI theory. However, this would be overkill, because mind/brain identity requires that brain states be both necessary and sufficient to produce mental states. This is not just a philosopher's quibble. Even basic commonsense concepts of identity require both necessity and sufficiency. It is necessary that your car have spark plugs in order for it to be a real, functioning car. A car without spark plugs is a piece of junk, not a car. But that doesn't mean the spark plugs are identical to the car, because owning a set of spark plugs is not sufficient to enable you to drive to work. MBI theory requires that for every experience or thought that we have, every shift in qualitative nuance or cognitive deliberation, there must be something in the brain that is not only co-occurrent with that mental event but also robust and detailed enough to be entirely responsible for it. Consequently, even if it turns out that some change in brain state is necessary for every change in mental states (as MBI theorists believe), this would not mean that the brain possesses the intrinsic power to produce mental states. Except for such things as afterimages, which Gibson called "psychological curiosities," all of our actual thought and experience emerges from the interactions between a brain, a body, and world. There is no scientific

evidence that the experiences that fill our days could be produced by anything less.

There is a philosophical thought experiment often cited to support MBI, but its persuasiveness reveals nothing but the nature and extent of our neo-Cartesian prejudices. This thought experiment, often called "the Brain in a Vat," is the materialist equivalent of Descartes' suggestion that an "evil genius" could be producing the illusion that an external world exists. In the MBI version, a mad scientist is substituted for the evil genius, and Descartes' noncorporeal mind is replaced by a brain in a vat. In the most straightforward version of the thought experiment, the brain is sitting in a vat and being causally triggered by a battery of bioelectric gizmos that are making it have worldlike experiences.

This thought experiment does not support the mind/brain identity, for two closely related reasons.

1. Until somebody gets a grant and conducts the appropriate experiments, there is no way of knowing whether or not a brain in a vat can fully duplicate all human experiences. It is far from obvious what the appropriate experiments would be, but even if that problem were solved we would simply have to wait and see. Perhaps a brain in a vat would feel disembodied even if all of the cranial activity ordinarily caused by embodiment were perfectly duplicated. Perhaps every single bit of the cranial activity ordinarily produced by the sight of a tree would produce only a pale simulacrum of a tree experience rather than the robust experience that happens when a real tree is present. Only when we make the appropriate tests (assuming the idea of appropriate tests makes sense, which I don't necessarily grant) can we determine whether or not such a brain can have all the conscious experiences we can have. Having just written a book arguing that it would not, my impulse is to say no. Others, caught up in the throes of the MBI, would probably say yes. But I see no reason why either set of intuitions would be any more useful in answering that question than in determining the chemical composition of the moons of Jupiter. These are empirical questions and cannot be answered by thought experiments.[1]

2. The Brain in the Vat thought experiment appears plausible only if we assume that experience consists of the discrete particles of sense-datum theory. If we imagine the brain-triggering devices to be something like electrodes, it is easy to also imagine that we could stimulate a brain to have experiences of green tree-shaped patches. But what John Dewey claimed, and what modern neuroscience appears to confirm, is that experience does not consist of discrete moments like frames in a film. Even when perception scientists show individual flash cards to subjects and have them report what each looks like, the subjects do not experience the cards as genuinely discrete sense data. Their experience of the so-called distinct sense data is embedded in a fundamentally unified life experience from which it cannot be separated. This life experience includes being involved in an experiment

that requires them to look at certain things and follow certain instructions, and which they find to be boring or interesting or an important contribution to science. Even if the subject is a devout Humean empiricist and believes that she is experiencing sense data, her beliefs and mood are still there to provide the context for what she takes to be sense data. Dewey argues, in other words, that there can be no experience at all without extended interactions with the world that can last for minutes or even hours and that such interactions must be constituted by skillful and flexible responses to a world that has enduring physical laws. The Brain in the Vat experiment thus would require a device that responds to us the way the world responds to us, and to which we could respond skillfully because it created experiences that obeyed all the laws of science and common sense. A device capable of creating world-experiences in a brain would have to be informationally identical to the world it was simulating in order to completely create those experiences. In other words, the device would not be a simulation of a world, it would *be* a world, and it would have to throw in a simulation of a body somewhere to complete the picture. This world-simulation system could be different from the so-called real world in ways the brain could not detect, but it would still be every bit as ontologically robust, unless you posited some sort of miracle to make up the difference. So in this case, as in our own, experience would be an emergent property of a brain, body, and world. The brain would not be creating those experiences all by itself. We have merely created functional equivalents for the body and world with silicon (or whatever); we have not dispensed with them. Even if building such a system were possible, it would not prove that the mind is identical to the brain. This is because the experiences emerge not from the neural activity alone but also from the interaction between the brain and the device.

## The Alleged Intrinsic Properties of Brains

In his *Psychosemantics*, Jerry Fodor claims that it is impossible to do psychology scientifically unless we assume that "causal powers . . . in the psychological case . . . supervene on local neural structure. We abandon this principle at our peril; mind/brain supervenience/identity is our only plausible account of how mental states could have the causal powers that they do have" (1987, 44). Fodor claims that these kinds of assumptions are inherent in the very concept of science—that science must operate by classifying things into natural kinds, each of which possesses intrinsic causal powers.

The concept of intrinsic causal powers appears to make sense if one focuses exclusively on certain examples. Knives really do seem to possess sharpness, and gunpowder does seem to be intrinsically explosive. It thus seems natural to assume that there is a clear distinction between intrinsic causal powers, which somehow inhere or are predicable to the objects de-

scribed by natural-kind terms, and the extrinsic causes that push the buttons that release those powers. This concept enables us to see causal powers as attributes of objects and to see the universe as a network of objects interacting with each other. According to this view, when an object encounters another object, each releases the other's intrinsic causal powers, rather like the way bumper cars at an amusement park ring bells in their bumpers when they collide. This distinction is also what makes the distinction between cause and effect work. When object A activates one of its causal powers, it has the effect of releasing the causal powers of B, which in turn have an effect on A or C or whatever and so on, with each effect becoming a cause for the next effect in the chain.

There is no reason to assume that just because this distinction often is useful to us it reflects something intrinsic about the world, independent of all of our concerns and projects. A nexus of causes is just a nexus of causes, and all of them are equally responsible for the events occurring, regardless of our thinking of them as internal or external. The explosion is every bit as dependent on the oxygen as it is on the match, despite our tendency to take the oxygen for granted. Causal properties are fundamentally relations, not monadic predicates, and sometimes the simplification that enables us to refer to some of them as intrinsic powers is not going to be useful. We say that an object has a certain causal power because there are so many different situations in which it produces a certain effect. Knives are considered to be intrinsically sharp because they can participate in butter-cutting events, paper-cutting events, wood-cutting events, and so forth. But exactly how many such situations must there be before we ascribe the causal power to the object rather than to the situation as a networked whole? We have no reason to assume that there is always (or ever) a single metaphysically correct answer to that question.

Nagarjuna argued that causal powers can never be coherently described as being intrinsic.

> Nagarjuna finds that, were objects to have a stable, fixed essence, the changes brought about by causes would not be logically intelligible or materially possible. Let us say, along with the school of Universal Existence, that the effect pre-exists in the cause, or for example, that the burning of fire and the thirst-quenching of water are inherent in the kinds of substances fire and water are. But if the effects already exist in the cause, then it would be nonsensical to speak of effects in the first place, because in their interaction with other phenomena the pre-existent causes would not produce anything new, they would merely be manifesting the potential powers already exhibited . . . if fire and water are stable substances which possess fixed natures or essences, then what sort of relation could they bear to other objects which have entirely different fixed natures? (Berger 2007)

Nevertheless, Nagarjuna does not take the nihilist position that all causal talk is empty and worthless. On the contrary, because he rejects the ultimate reality of any substances with intrinsically fixed natures or essences, he is willing to ascribe some level of reality to many kinds of causal talk.

Tao Jiang, describing Nagarjuna's position, says that "Nagarjuna . . . rejects the conception of 'primary' reality, and embraces 'secondary' reality—the only reality for him. More importantly, in so doing he radicalizes the very conception of 'secondary' reality by cleansing from it any 'primary' element, such as substance and essence, etc." (Jiang 2008). Jiang thus sees Nagarjuna as a kind of protopragmatist who encourages us to change our ontologies whenever that brings greater understanding, because there is no fundamental ontology underlying everything else. If there is no ultimate reality, so-called appearances are the only reality. The Abhidharma Buddhist claims that mind and all other substances are ontologically dependent on the dharmas. The MBI theorist believes that mental experiences are ontologically dependent on brain states and nothing else, even though other causal factors are partly responsible for almost all of our experiences. Both of these positions assume that there is a line between the "real" intrinsic properties and the causal relations that trigger those so-called intrinsic properties—a metaphysical claim that both Nagarjuna and I reject. This distinction between intrinsic and relational is pragmatically necessary, but there is no single way of deciding where it should always be drawn. Nothing is intrinsically intrinsic, as it were. Intrinsicality is itself a relational property.

When scientists use the terminology of a different scientific specialty, they usually redefine the objects in one science as a system of relations in another. A chemist, for example, may refer to the intrinsic causal properties of sulfur or magnesium, while a particle physicist will see those intrinsic chemical properties as being relations between subatomic particles. The borderline between objects and relations will shift depending on what one is talking about. The only way that one can actually *explain* an object's "intrinsic" causal powers (rather than merely describe or refer to them) is to analyze the object into its parts and then talk about the relations between the parts. To say that certain properties are intrinsic is simply to say that for the moment we are going to refrain from analyzing them. Because discourse has to start somewhere, every science will talk about entities with intrinsic causal powers, and such talk will describe the relations between those entities. But this does not eliminate the possibility (or the necessity) of other scientific specialists redescribing these so-called fundamental entities by analyzing them into relations among other "fundamental" entities.

Nor is downward analysis the only way to shift the definition of what is intrinsic. To see an object as possessing intrinsic causal powers is to see it as an autonomous closed system of relations that interacts with other autonomous systems. As Nagarjuna correctly points out, however, this idea contains an inherent contradiction. Once an "autonomous" system interacts with another "autonomous" system, the two systems are no longer genuinely autonomous. Instead they become parts of a larger system that could be seen as a unified whole with "intrinsic" causal powers of its own,

and so on *ad infinitum*. Nagarjuna called this principle "the interminability of dependent origination." Because there are many different ways that parts can be assembled into wholes, or wholes analyzed into parts, each science designates different entities as possessing so-called intrinsic causal powers. We HEC theorists have become somewhat aware of this conclusion through studying outstandingly anomalous scientific data, which has gradually eroded our faith in the intrinsic mental powers of the brain. Nagarjuna articulated this principle thousands of years before modern science.

Fodor is probably correct when he claims that psychology and neuroscience assume that brains have the intrinsic causal power to produce mental states. But this may be only an indication that these sciences have not freed themselves from the naive Aristotelian assumptions of common sense. If we assume, as Aristotle did, that biological science is only supposed to catalog dispositional properties, then it is true in some sense that the brain has the power to cause mental states. But it also is true in the same sense that opium has the power to cause sleep, and no one considers that to be a sufficient scientific explanation any more. It seems likely that the cognitive science of the future will dispense with this kind of talk about causal powers altogether. Modern science rejects purely dispositional explanations for practical reasons. Newtonian science, which explains dispositions by analyzing them into relations between smaller and/or external parts, has much more predictive power than Aristotelian descriptions of intrinsic dispositions. This is why Newtonian science is the basis of modern engineering and science. It is much more effective and informative to give the chemical structure of opium than to say it has intrinsic dormitive powers.

Nagarjuna saw the conceptual inadequacies in the idea of intrinsic dispositions centuries before the existence of modern science. He pointed out that if fire burned intrinsically "it would be forever aflame; flames could be ignited without a cause" (Garfield 1995, 190). His point is that we cannot make a distinction between a potential fire and an actual fire unless we acknowledge that the disposition is incomplete in some way. If the fire "were eternally in flames, starting it would be meaningless" (Garfield 1995, 190). Consequently, the assertion that fuel has the intrinsic disposition to burn is true only in a manner of speaking, and false if we take it to be absolutely true. What Nagarjuna is really doing here is predicting the course of modern science. Aristotelian causality was intrinsic and dispositional and thus could be only descriptive. Nagarjuna's causal paradoxes show that Aristotelian explanations cannot be genuinely explanatory—that something more is needed if we are going to have genuine knowledge of causes. It is convenient to say that certain properties are intrinsic dispositions, but it is self-contradictory to say that this is an absolute metaphysical fact. One of the reasons that modern science has made so much progress is that, unlike Aristotelian science, it refuses to accept dispositional explanations as the end of the causal story.

I have argued that we ought to take a pragmatic madhyamaka-style view even toward the intrinsic causal properties of physics (Rockwell 2008). The concepts of modern physics, such as sound wave, pressure, heat, and electromagnetism, rely far more on processes and fields than on aggregates of particles. I believe that there is no scientific reason for asserting that the processes are more fundamental than the particles and that our insistence that everything must be controlled by the intrinsic properties of tiny objects is the last holdover of Aristotelian metaphysics in our post-Newtonian world. This is the conclusion we must accept if we consistently apply Nagarjuna's arguments to modern science.

The modern reductionist view, which is probably the majority opinion among scientists and philosophers, is that this process of uncovering dependent arising has got to bottom out somewhere. If we study brains, we must analyze the brain into neurons, then analyze the neurons into parts, then the parts of those parts, and so on until we eventually discover a fundamental item (a particle, a string) that alone possesses the genuine intrinsic causal powers that propel everything else in the universe.

HEC is concerned only with cognitive science, not physics, so defenders of HEC do not need to deny the existence of these fundamental particles. They can take a position similar to Abidharma Buddhism and accept that the categories of psychology and cognitive science are not as fundamental as the categories of microphysics. Just as the Abidharma Buddhists believed that the categories of mind and world are reducible to the more fundamental constituents they called dharmas, so the reductive physicalist must deny that brains and neurons are fundamental entities with intrinsic properties. Scientific reductionists should not attribute intrinsic properties to the medium-sized objects described by neuroscience and psychology because that would work against the claim that microphysical particles possess the intrinsic causal properties that control neurons and brains. There is no denying that attributing intrinsic properties to the brain is sometimes useful, but genuine scientific progress requires recognizing that these so-called intrinsic powers have only what Nagarjuna calls a conventional reality. To deny them any reality at all would be to make the mistake that Nagarjuna called nihilism. To assume that they must be absolutely real would be to make a mistake analogous to what Nagarjuna called eternalism.

My version of HEC says that cognitive science should take a middle way and recognize that the borders between self and world have only a conventional reality. This means that we should stop assuming that there is one answer to the question "Which physical item is the mind?" and accept that the borders between the mind and the world are both ambiguous and flexible. They are *ambiguous* because (1) at any given moment there is no absolutely sharp line where the self stops and the world begins, just as you cannot draw a line accurate to the millimicron that divides the United States and Canada, and (2) different sciences must draw the line

between self and world at different places because these sciences have different goals and purposes. The best place for neuroscience to draw the line might even be at the nervous system, at least some of the time. But I believe that there is no fact of the matter as to which scientific line is *the* correct border. The borders are *flexible* because (3) they fluctuate according to the projects and purposes of not only the mind doing the studying but also the mind being studied. That is why I refer to the mind as a behavioral field, because, like a magnetic field, its borders fluctuate over time. Unlike a magnetic field, these fluctuations depend on the self's goals and purposes. Because there is no principled way of drawing the line between causation and embodiment, that is, between the so-called intrinsic causal powers of the mind and the external causes of the world, the line between the mind and the world fluctuates as we shift between different purposeful human activities.

This position seems to be at odds with common sense, but it has been defended by some twentieth-century philosophers, especially Martin Heidegger and Dewey. This was what Heidegger was describing when he talked about the readiness-to-hand of tools when they are in skillful use. When I am using a hammer, the hammer and I are one. When the hammer breaks, and becomes merely present-at-hand, the hammer and I are no longer one. When I stop hammering and drive off in my car, the car and I are one. To acknowledge that the line between self and world is purely pragmatic leads to the understanding that all of this may be true not just metaphorically but literally.

## Notes

I thank MIT Press for permission to reprint substantial selections from Rockwell 2005.

1. There are replies to two other versions of the thought experiment in Rockwell 2005, 67–69. Some have suggested that the mind is identical to the brain because brain states would completely embody mental states even if the brain states were produced by miracles or chance. However, there is no way we can possibly know this.

## References

Arnold, Dan. 2005. "Madhyamaka Buddhism." *The Internet Encyclopedia of Philosophy*. *http://www.iep.utm.edu/b/b-madhya.htm#H6.*

Berger, Douglas. 2007. "Nagarjuna." *The Internet Encyclopedia of Philosophy. http://www.iep.utm.edu/n/nagarjun.htm#H3.*

Clark, Andy, and David Chalmers. 1998. "The Extended Mind." *Analysis* 58:10–23.

Fodor, Jerry. 1987. *Representations.* Cambridge: MIT Press.

Garfield, Jay. 1995. *The Fundamental Wisdom of the Middle Way.* Oxford: Oxford Univ. Press.

Gibson, J. J. 1955. *Perception as a Function of Stimulation*. Lecture given at the Institute of Experimental Psychology, Oxford, with glosses by U. T. Place. *http://www.california.com/~mcmf.*

James, William. 1890. *Principles of Psychology.* New York*:* Dover.

Jiang, Tao. 2008. "The Incompatibility of Two Conceptions of Reality: Dependent Origination and Emptiness in the Mulamadhyamakakarika." Lecture given at Columbia University Department of Religion, November 7.

McCarthy, Christine. 2006. "Review of *Neither Brain nor Ghost*." *Education and Culture: The Journal of the John Dewey Society* 22:2.
Pert, Candace. 1987. "Neuropeptides, the Emotions and BodyMind." In *Consciousness and Survival*, ed. J. S. Spong. 82–87. Sausalito, Calif.: Institute of Noetic Sciences.
———. 1997. *Molecules of Emotion*. New York: Simon and Schuster.
Pert, Candace, and F. O. Schmitt. 1985. "Neuropeptides and Their Receptors: A Psychosomatic Network." *Journal of Immunology* 135:820–26.
Rockwell, Teed. 2005. *Neither Brain nor Ghost: A Non-Dualist Alternative to the Mind/Brain Identity Theory*. Cambridge: MIT Press, Bradford Books.
———. 2008. "Processes and Particles: The Impact of Classical Pragmatism on Contemporary Metaphysics." *Philosophical Topics* 36:239–58.
Skrbina, David. 2006. Review of "*Neither Brain nor Ghost.*" *Journal of Consciousness Studies* 13:6.
Wilson, Robert. 2001. "Two Views of Realization." *Philosophical Studies* 104:1–31.