

## **Brain in a Vat or Body in a World?**

### **Brainbound versus Enactive Views of Experience<sup>1</sup>**

Evan Thompson

Department of Philosophy, University of Toronto, 170 St George Street, Toronto,

ON M5R 2M8 Canada

evan.thompson@utoronto.ca

Diego Cosmelli

Escuela de Psicología, Pontificia Universidad Católica de Chile, Av. Vicuña

Mackenna 4860, Macul, Santiago, Chile

dcosmelli@uc.cl

In press for a special issue of *Philosophical Topics* on embodied cognition.

Is consciousness all in the head, or more precisely all in the brain? Or is the body beyond the brain an essential part of the biological basis of consciousness? To put the question another way, does the body belong to the “minimal realizing system” for conscious experience or is the minimal realizing system for consciousness confined to the brain?

According to the “enactive” view of experience, consciousness is a life-regulation process of the body interacting with its environment.<sup>2</sup> Perception, action, emotion, imagination, memory, dreaming—these are modes of self-regulation that depend directly on the living body and not just the brain. According to the enactive view, the body shouldn’t be seen as a mere outside causal influence on an exclusively neuronal system for consciousness because the minimal requirements for consciousness include a living body, not just neuronal events in the skull.

The enactive view stands in tension with the widespread view in the neuroscience of consciousness that consciousness is brainbound. Many neuroscientists and philosophers would say that your brain directly determines what you experience, but your body affects what you experience only via its influence on your brain. According to this way of thinking, the body is strictly inessential for conscious experience; for example, in principle, a disembodied brain in a vat could have the same kinds of subjective experiences or states of phenomenal consciousness as an embodied brain.

Our focus will be on the tension between these two views—between what we’ll call the *enactive view* and the *brainbound view*, or *Enactive* and *Brainbound* for short.<sup>3</sup> As we’ll see, these views work with different conceptions

of consciousness, and these different conceptions of consciousness motivate different working assumptions about the biological basis of consciousness. Our aim is to persuade you to prefer Enactive to Brainbound. Specifically, we'll argue that a strong case can be made for thinking that the biological system for consciousness isn't limited to the brain. Consciousness isn't a strictly brain phenomenon but an organism phenomenon. The minimal biological basis for consciousness includes nonneural factors and physiological processes beyond the skull.

### **Brainbound Basics**

We begin with Ned Block's statement of what he calls the "orthodox view" of the brain basis of consciousness, the view we're calling Brainbound.<sup>4</sup>

Block presents Brainbound in the context of criticizing Alva Noë's version of the enactive view.<sup>5</sup> According to Noë, perceptual experience is a mode of temporally extended skillful interaction with the world. To perceive is to explore one's environment by exercising in a practical and bodily way one's sensorimotor knowledge of how sensory appearances vary as a function of movement. It follows, for Noë, that, "To perceive like us... you must have a body like ours." As he says, "If perception is in part constituted by our possession and exercise of bodily skills... then it may also depend on our possession of the sorts of bodies that can encompass those skills, for only a creature with such a body could have those skills."<sup>6</sup>

Block contends that this view misidentifies what *causes* experience with what *constitutes* experience. Although perceptual experience depends causally on having a body and exercising sensorimotor know-how in movement, exercising

this bodily knowledge doesn't constitute or directly determine perceptual experience. Rather, what directly determines a given perceptual experience is a specific pattern of brain activity. If the right neural pattern were to occur, so too would the subjective experience, regardless of the wider context of bodily activity in the environment.

Block interprets Noë as making a *metaphysical* claim about the subpersonal or physical basis of perceptual experience—that it includes the body and not just the brain. This interpretation strikes us as off-key. Noë's enactive view and enactive views in general offer an *explanatory* framework for perception. We use phenomenological considerations about perceptual experience to constrain how we think about the subpersonal mechanisms of perception. First, we start by trying to get the phenomenology right, or to put it another way, by trying to conceptualize perceptual experience properly at the personal level. Here the central idea is that to perceive is to be in an interactive relationship with the world, not to be in an internal state that happens to be caused by the external world. Second, we argue that, given this conception of perceptual experience, we can't specify the mechanisms of perception only in terms of what goes on in the brain without including the body and its dynamic sensorimotor coupling with the environment. Therefore, it makes little sense to restrict the physical basis of perceptual experience to the brain without including the body.

Let's go back to Brainbound. Block reads the enactive view as making a metaphysical claim about what he calls the "minimal constitutive supervenience base" for experience. In his words:

The issue of the constitutive supervenience base for experience is the issue of what is—and is not—a *metaphysically necessary part of a metaphysically sufficient condition* of perceptual experience. That is, it is the issue of what is—and is not—part of the *minimal* metaphysically sufficient condition for perceptual experience (the minimal supervenience base). Noë's enactive view says that the skilled active body is part of that minimal condition (minimal supervenience base), whereas the view which I hold and which I have labeled the orthodox view, is that nothing outside the brain is part of it.<sup>7</sup>

Stated this way, the orthodox view is a metaphysical view about the relation between subjective experience and the brain. This metaphysical view is often combined with a certain conception of consciousness and how neuroscience should investigate it. According to this conception, the first main task for the neuroscience of consciousness is to find the *neural correlates of consciousness* (the NCC), specifically the *minimal neural correlates* for the phenomenal contents of consciousness.<sup>8</sup>

A minimal neural correlate for a given conscious experience, such as the visual experience of the color red, is the minimal set of neuronal events and mechanisms jointly sufficient for that conscious experience.<sup>9</sup> According to the NCC research program, the discovery of the NCC should be the first prime goal of the neuroscience of consciousness.

Combining the NCC research program with Brainbound (or what Block calls the orthodox view) gives us a certain picture of consciousness and the brain.

To explain this picture, however, we need to introduce some conceptual distinctions.

The first distinction is between *state consciousness* and *creature consciousness*, that is, between phenomenally conscious states and phenomenally conscious creatures.<sup>10</sup> The standard way to explain these notions is to say that a phenomenally conscious state is a state for which there's something it's like for the subject to be in that state. Such states are individuated in terms of their phenomenal content or their phenomenal character. A phenomenally conscious creature is a creature for which there's something it's like to be that creature.<sup>11</sup> Phenomenally conscious creatures are subjects of experience.

In cognitive and clinical neuroscience, studies of the neural substrates of consciousness tend to focus either on state consciousness or creature consciousness.

State-based studies contrast the reportable awareness of a given phenomenal content (such as seeing a face) with the lack of reportable awareness of that content (for example, in a masking experiment), or with the reportable awareness of a different phenomenal content (such as a house; for example, in a binocular rivalry experiment). In these studies, subjects are awake, aware, and able to report the changing contents of their phenomenal consciousness.

Creature-based studies focus either on the contrast between phenomenal consciousness and its absence (under anaesthesia or during coma) or between different global or background states of consciousness (such as wakefulness and dreaming).

The concepts of *creature consciousness* and *background state consciousness* are thus closely related. Background states of consciousness, such as wakefulness

and dreaming, are domain general, not modality specific. They characterize one's overall phenomenal perspective as a conscious subject. In this way, they are properties of conscious creatures.<sup>12</sup>

Speaking roughly, we can say that state-based studies indicate that the reportable awareness of distinct phenomenal contents depends on the activation of particular cortical regions and networks,<sup>13</sup> whereas creature-based studies indicate that subcortical systems such as the thalamus and brainstem are crucial for the presence of consciousness compared with its absence, and for the transitions and differences between wakefulness and other background states such as dreaming.<sup>14</sup>

The next distinction we need to make is between the *core realization* and the *total realization* of a given phenomenally conscious state.<sup>15</sup> In general, the core realization of a property or capacity suffices for that property or capacity only when properly placed in the context of a larger system that constitutes the total realization.<sup>16</sup> Block proposes that “the core NCC is the part of the total NCC that distinguishes one conscious state from another—the rest of the total NCC being considered as the enabling conditions for the conscious experience.”<sup>17</sup>

According to this proposal, the total NCC comprises the neural substrates of creature consciousness, that is, of a subject's being able to experience any phenomenal states at all. These substrates can be treated as enabling conditions, however, in relation to the core NCC, which distinguishes one conscious state from another in terms of its specific phenomenal content.

Yet this formulation still isn't complete. In general, the total realization of a property or a capacity suffices for that property or capacity only given the appropriate *background conditions*.<sup>18</sup> Similarly, the total NCC suffices for creature

consciousness only given certain background conditions, which normally include nonneural parts of the body and the environment.

Let's connect these points to the metaphysics of Brainbound or the orthodox view. For Block, the core NCC of a given experience is equivalent to the minimal constitutive supervenience base for that experience. In other words, fixing the core NCC fixes the phenomenal content of the experience. As Block says, "if the relevant brain state were to come about—*somehow*—the experience would be instantiated."<sup>19</sup> In his example, were a disembodied, freestanding brain in the relevant state to arise through the chance fluctuations of microphysical particles, it would instantiate the experience (that is, it would instantiate a conscious state with the same phenomenal content, though not the same intentional or representational content).

The classic philosophical expression of this idea, of course, is the brain in a vat. According to this thought experiment, a disembodied brain placed in a life-sustaining vat and stimulated in the right way by a supercomputer would have experiences with the same kind of phenomenal content or subjective character as you have.

Although usually a philosophical device for raising epistemological questions about the relation between subjective experience and the world, the brain in a vat also functions as a methodological device in philosophical reflection on the neuroscience of consciousness. There it serves to illustrate the idea that the body can be "screened off" from the neural correlates of consciousness, which then get metaphysically conceptualized as the minimal constitutive supervenience base for subjective experience. We can see the brain in a vat put to this use in the following remarks by Thomas Metzinger:



[T]here is a minimally sufficient neural correlate for the content of consciousness at any given point in time. If all properties of this local neural correlate are fixed, the properties of subjective experience are fixed as well. Of course, the outside world could at the same time undergo considerable changes. For instance, a disembodied but appropriately stimulated brain in a vat could—*phenomenologically*—enjoy exactly the same kind of conscious experience you do right now while reading this book.<sup>20</sup>

What we want to do now is to examine precisely this idea—illustrated forcefully by the brain in a vat—that the body can be screened off from the neural substrates of consciousness. Our next step will be to reconsider the brain-in-a-vat thought experiment. Thinking carefully about this experiment will reveal that the brain and body are so deeply entangled, structurally and dynamically, that they are explanatorily inseparable. Dynamic entanglement implies that we can't understand consciousness—especially creature consciousness—by considering only the activity of neurons apart from the body, and hence we have good explanatory grounds for supposing that the minimal realizing system for consciousness includes the body and not just the brain. In this way, we intend to put the brain-in-a-vat thought experiment to a new use, one that supports the enactive view of experience.<sup>21</sup>

### **A Close Look at the Brain in a Vat**

What would it take to envat the brain so that it would function exactly like its embodied counterpart? To our knowledge, the only philosopher who has addressed this question is Dan Dennett. We're thinking not of his classic paper "Where Am I?,"<sup>22</sup> which dramatically portrayed the brain in a vat scenario for cognitive science, but rather of the "Prelude" to his book *Consciousness Explained*.<sup>23</sup> There he addresses the question from a technical, bioengineering perspective, not the philosophical perspective of what is conceivable or conceptually possible in principle. We're going to follow Dennett's lead, for as he notes, "sometimes an *impossibility in fact* is theoretically more interesting than a *possibility in principle*."<sup>24</sup> We'll address the mere possibility in principle of the brain in a vat later.

In thinking about how to envat the brain, we need to consider three main things—keeping the brain alive and up and running, the brain's self-generated activity in tight coupling with the body, and what it would take to mimic precisely the stimulation the nervous system normally receives from the body and the environment.

#### *Keeping the brain up and running*

Before we can appropriately stimulate the envatted brain, we need to keep it alive and functioning. This already is no mean feat.

First, we need some protective apparatus for the brain. This apparatus serves to replace the skull (and spine, if we choose to keep the spinal cord). To ensure the brain's flotation, the protective device will need to be filled with a liquid analogous to the cerebrospinal fluid. This liquid needs to be able remove

waste products of neuronal metabolism and so must be continually recycled.<sup>25</sup> One way to achieve this recycling would be to couple the protective fluid to the second thing we need—a circulatory system.

Almost everyone has experienced the intense dizziness and cognitive impairment that happens when you stand up fast. The unimpeded supply of blood to every part of the brain is critical for its functioning and by no means a trivial physiological accomplishment. To envat the brain, we must provide an adequate blood supply or a fluid with similar biochemical properties. For this task we could probably choose to keep the vascular system in place as a delivery structure. Alternatively, in the true spirit of the thought experiment, we can imagine replacing the entire cerebral vasculature with some synthetic device that shows similar properties of selective permeability and local and systemic responsiveness to the brain's ongoing demands. This kind of responsiveness is absolutely crucial. Without it there would be no way to compensate for even minimal departures from homeostasis due to neuronal activity, with fatal consequences for our experiment.

The tight coupling of blood flow and neuronal activity is a basic physiological fact known as functional hyperemia.<sup>26</sup> Neuroimaging techniques, such as fMRI and PET, rely on different aspects of this coupling.<sup>27</sup> Although the actual mechanisms underlying the coupling aren't fully understood, a variety of molecular and cellular factors are known to participate in the regulation of local blood flow in the brain.

Our life-sustaining system must be able to deal efficiently with these sorts of factors in order to meet the local needs arising from ongoing neuronal

activity. So our synthetic apparatus will probably have to be as sophisticated as an actual vascular system in its structural features and functional capacities.

Suppose we've succeeded in setting up such an immensely complex system. We then need to move the fluid through the delivery structure. Here some kind of pump is needed, as well as some minimal and highly selective recycling system for replenishing the fluid's necessary components, including oxygen, glucose, and the numerous soluble ions, proteins, and other biomolecules that account for the fluid's osmotic, nutrient, and regulatory properties.

Because the pump and recycling system must be responsive to the brain's actual demands, they should be functionally coupled to the brain's activity. This coupling ensures the local availability of the soluble factors provided by the circulatory system and keeps the concentration of the circulating molecules and ions within a physiological range despite continuous demands from the neuronal tissue. The brain normally relies on a host of regulatory loops involving organs outside the brain to meet these needs.

Let's summarize things up to this point. Whatever life-sustaining system we produce, it will involve at least the capacity to keep up with the brain's energetic, ionic, osmotic, and recycling needs. It will therefore include some kind of circulatory system plus the necessary pumps, oxygenating devices, and additional subsystems for maintaining physiological levels in the circulating fluid.

The next point is less obvious. What the brain requires at any given instant depends on its own ongoing and self-generated activity, or what neuroscientists call "intrinsic activity." Our life-sustaining system must support

this intrinsic activity and respond to it locally and systemically at any given instant, independent of any outside evaluation of the brain's needs. Consequently, our life-sustaining system needs the kind of robustness and flexibility we see only in energetically open, self-sustaining, and self-regulating systems—that is, organisms.

This life-sustaining system is starting to look less like a vat and more like a surrogate living body or organism.

*Self-generating activity and the body-coupled brain*

When considering the requirements of keeping the brain up and running, we began by taking an external control perspective. From this perspective, the issue is how to control the brain from the outside so that it stays alive and functioning. Yet once we take into account the brain's endogenous workings, it becomes clear that our life-sustaining system must be intimately coupled to the nervous system's own intrinsic activity.

This basic requirement necessitates a radical shift in how we think about our vat. Whatever life-sustaining system we construct, the functioning of its every part, as well as its overall coordinated activity, must be kept within a certain range by the nervous system itself in order for the brain to work properly. Hence the external control perspective is not generally valid. Instead, our life-sustaining system and the brain must be seen as reciprocally coupled and mutually regulating systems.

According to a number of recent proposals, this kind of tight coupling between neural and nonneural factors—between brain and body in the normal embodied case—constitutes the organism as a functional unity and underwrites

the phenomenal feeling of self that permeates normal creature consciousness.<sup>28</sup> For example, according to Damasio, the nervous system continually maps the state of the body through a series of core neural structures that are crucial for both body regulation and the feeling of self.<sup>29</sup> In this theory, creature consciousness with a minimal phenomenal feeling of self arises as a feature of life-regulation processes effected by the nervous system in tight coupling with the body.

Damasio hasn't missed the implications of his proposal for the brain-in-a-vat thought experiment. In his book *Descartes' Error* he writes:

It might be argued that if it were possible to mimic, at the level of the dangling nerves, realistic configurations of inputs as if they were coming from the body, then the disembodied brain would have a normal mind. Well, that might be a nice and interesting experiment "to do" and I suspect the brain might indeed have *some* mind under those conditions. But what that more elaborate experiment would have done is create a body surrogate and thus confirm that "body-type inputs" are required for a normally minded brain after all. And what it would be unlikely to do is make the "body inputs" match in realistic fashion the variety of configurations which body states assume when those states are triggered by a brain engaged in making evaluations.<sup>30</sup>

Lawrence Shapiro, in his book *The Mind Incarnate*, relies on Damasio to defend what he calls the "embodied mind thesis," but he also thinks Damasio's

assessment of the brain-in-a-vat thought experiment is confused. Damasio seems to be saying that it would be possible to remove the brain without disrupting the type of inputs it receives, but that the brain would no longer function normally. Shapiro asks: "If the inputs to the brain in a vat were exactly what they are to a brain in a body, what explains why the envatted brain would have *some* mind, but not a normal mind?"<sup>31</sup> Furthermore, at the end of the passage, Damasio seems to be saying that it wouldn't be possible to mimic precisely the inputs that a brain in a body receives, so the thought experiment wouldn't be possible after all. But as Shapiro says: "Either it is possible to mimic the inputs to the envatted brain precisely or it is not. If it is, why wouldn't the brain perform normally? If it is not, then this is just to deny the coherence of the thought experiment."<sup>32</sup>

There's a way to understand Damasio, however, that might alleviate some of the confusion. The crucial point is that identical inputs don't suffice for identical states. It's not the case that an embodied brain and its envatted duplicate will remain qualitatively identical simply because they receive identical inputs throughout their lives. The brain isn't a reflex machine whose activity is externally controllable through input instructions. Rather, it's a highly nonlinear and self-organizing dynamical system whose activity exhibits an extreme sensitive dependence on initial conditions. Inputs perturb such complex systems, but don't specify particular outcomes. Furthermore, most inputs arise as a consequence of the system's own intrinsic activity. Hence to get the body-type inputs to match the normal inputs precisely would require getting them to match the bodily inputs to the brain that arise from the brain's nonlinear and unpredictable intrinsic activity.

So here's our gloss on Damasio: As a technical matter, it seems highly unlikely we could achieve such a matching. Indeed, in a universe with stochastic or noisy thermodynamic processes, it's probably impossible, not merely highly improbable. Nevertheless, whether some kind of creature consciousness could be achieved with something less than this kind of match—whether it could be achieved through whatever body-type inputs we could technically deliver—seems an open question.

With these thoughts we come to our third requirement—mimicking environmental stimulation.

#### *Mimicking environmental stimulation*

Here our concern isn't so much creature consciousness but phenomenal state consciousness. What would it take to produce conscious states with specific phenomenal contents like those of normal perception?

The minimal requirement is to deliver stimulation to the neuronal terminals that matches sufficiently well the stimulation the brain normally receives from the environment.

Let's not underestimate the complexity of the stimulating devices. Imagine an artificial device capable of stimulating every fiber of the optic nerve in perfect correlation with the light pattern of the scene to be recreated, guaranteeing all the dynamic receptive field relations found originally among retinal cells, maintaining perfect synchrony with the brain's exploratory motor efference signals as its sensory systems scan through the virtual image, and updating its activity so as to match precisely the sensory refference (the sensory feedback caused directly by the motor efference). As Dennett notes, this problem



will be computationally intractable on even the fastest computer and storing all the information is ruled out because of the combinatorial explosion of possibilities.<sup>33</sup>

Another crucial requirement is that we can't allow these stimulating devices to disrupt the life-sustaining system already established. Every stimulation we deliver produces a departure from homeostasis for which there must be immediate compensation or the whole system will crash. So whatever stimulating devices we construct, they must be integrated seamlessly into the "vat."

We now have two basic requirements. On the one hand, the stimulation must mimic the stimulation the embodied nervous system normally receives. On the other hand, the stimulating devices must not disrupt the life-regulation crucial for the functional unity of the system and for the stable realization of creature consciousness.

These two requirements imply that our artificial stimulating devices need to be controlled by the brain itself through sensorimotor loops. In other words, we need to equip our brain in a vat with real (synthetic, not virtual) peripheral sensorimotor systems.

A significant body of work in computational neuroscience supports this point. This work makes clear that adaptive behavior arises from the dynamical coupling of the nervous system and peripheral sensorimotor systems; it isn't programmed or commanded by the brain.<sup>34</sup> So the sensorimotor interface we give to our envatted brain must have sufficient flexibility and processing capability to establish this dynamical coupling for the brain. The best and

probably only way to establish this coupling is to give our brain real sensorimotor systems it can control.

Let's recall that in order to keep the brain alive and functioning, our so-called vat must already be self-regulating and self-sustaining like a living organism. We've just seen that we need to equip this system with synthetic sensorimotor devices that can operate with a certain degree of autonomy while tightly coupling with the brain's ongoing intrinsic activity. Our so-called brain in a vat now looks an autonomous sensorimotor agent. In trying to envat the brain, we have wound up with an embodied agent in the world.

### **The Null Hypothesis**

On the basis of the foregoing considerations, we can propose the following null hypothesis for the brain-in-a-vat thought experiment. *Any adequately functional "vat" will be a surrogate body.* We don't mean a body like ours in its material composition, but one sufficiently like ours in its functional organization. At this level, the term "body" means a self-sustaining system (life-regulation) that controls its own sensorimotor coupling with the outside world (sensorimotor agency). In short, the null hypothesis is that the so-called vat would be no vat at all, but rather an embodied agent in the world.

This line of thought suggests a way to state Enactive as an alternative empirical hypothesis to Brainbound. The enactive hypothesis is that *the minimal biological realizing system for creature consciousness is not the brain (or some neural subsystem) but an organism, understood as a self-sustaining system composed of some crucial set of dynamically entangled neuronal and extraneuronal subsystems.* This

hypothesis is the one that needs to be rejected in order to reject Enactive in favor of Brainbound.

### **The Brain in a Vat as a Supervenience Thought Experiment**

The brain in a vat is an example of what Susan Hurley calls a *supervenience thought experiment*.<sup>35</sup> In this sort of controlled thought experiment, you divide candidate explanatory factors into an internal set and an external set relative to some boundary, such as the skull or skin. You then suppose the internal factors hold constant while the external factors vary. In the brain in a vat, neural states are supposed to hold constant across the embodied and envatted brains, while external factors obviously vary.

A crucial supposition of such supervenience thought experiments is that you can “unplug” internal factors from one set of external factors and “replug” them into another. Being unpluggable accordingly provides a critical condition of possibility for the thought experiment: If the internal factors can’t be unplugged from the external ones, if they vary together across the relevant situations, then they aren’t explanatorily separable and the supervenience thought experiment isn’t possible.

Our examination of the brain in a vat strongly suggests that neural factors aren’t unpluggable from bodily factors and thus that the two are explanatorily inseparable. In the range of possible situations relevant to the explanatory framework of the neuroscience of consciousness, the brain in a vat thought experiment, strictly speaking, doesn’t seem possible (because the envatted brain turns out to be an embodied brain after all).

We might wonder now whether the thought experiment is even coherent. If envatting the brain in the right way implies embodying the brain, then the thought experiment undermines itself. The whole point of the thought experiment is to remove the world and the body while preserving brain function. Yet to get the brain to be a brain in the right way we have to give it a body and a world.

### **Conceivability**

If you think the important issue is the mere conceivability or conceptual possibility in principle of a brain in a vat, then everything we've said so far will probably strike you as irrelevant. If all that matters is conceivability, then we can avail ourselves of whatever conceivable technical resources we need, regardless of whether such resources are remotely feasible or even possible in our world or in worlds with our laws of nature.

But such conceivability or possibility in principle tells us virtually nothing of interest with regard to what concerns us here, namely, the explanatory framework of the neuroscience of consciousness in relation to embodied cognitive science, or to be more specific, Brainbound versus Enactive as rival explanatory research programs.

Consider that even if our concern is mere conceivability, we still have to face the question of what the term "brain" designates, such that envatting *that* and *only that* would suffice to duplicate experience. Do we need to envat only the neurons and their synaptic connections? Or must we also envat the glial cells, which outnumber the neurons nine to one and are now believed to be critical to learning and the formation of memories?<sup>36</sup> Do we need to envat immune cells,

which together with glial cells form complex cellular and molecular communication networks that mediate mood, emotion, and pain?<sup>37</sup> Are the immune systems and endocrine systems mere boundary conditions on the neuronal determination of conscious experience or do they belong to the minimal realizing system for consciousness? We simply don't know the answers to these questions.

Moreover, if one accepts that there's an explanatory gap—that we presently have no understanding of how phenomenal consciousness could be explained in terms of neural activity—then we can hardly assume that if there were a successful biological explanation of consciousness, it would explain consciousness only in terms of strictly neural activity instead of appealing to some wider system comprising other nonneuronal biological factors.

One might respond to these considerations by saying that, given our present scientific knowledge, it's implausible to suppose that the realizing system for consciousness includes anything nonneuronal. Jesse Prinz expresses this thought in the following way:

We have never found any cells outside the brain that are candidates as correlates for experience. Such cells would have to co-vary with conscious states in content and time course. Every component of the body that we can experience is represented in the brain, and when the corresponding brain areas are damaged experience is lost. Conversely, bodily experience can continue after the body is damaged, as in the case of phantom limb pain.

There is, in short, little reason to think the correlates of experience extend beyond the cranium.<sup>38</sup>

But this assessment seems to miss the mark. It treats the problem of explaining consciousness simply as the problem of explaining phenomenal state consciousness in a given sensory modality and therefore assumes that the issue of embodiment is decided by the fact that we've found cortical neurons whose receptive field properties correspond in certain limited respects with aspects of phenomenal content,<sup>39</sup> whereas we haven't found any nonneuronal cells that correlate with conscious states in this way. Yet in every experiment that establishes these correlations, the subjects are already conscious and able to report the changing contents of their awareness. Thus, the neural correlates of these phenomenal changes aren't sufficient for consciousness because they presuppose that the subject is already conscious with some total field of background awareness.<sup>40</sup> The crucial problem is to explain this background consciousness. In other words, the crucial problem is to explain why the creature is conscious at all. Background states of consciousness, such as waking and dreaming, are global modulations of creature consciousness.<sup>41</sup> So the crucial matter is to account for creature consciousness. How can we be so confident that creature consciousness is explainable in strictly neural terms?

According to current evidence, the biological basis for creature consciousness includes subcortical parts of the brain that are deeply involved in life-regulation and bodily homeostasis.<sup>42</sup> These subcortical structures don't command the body in a hierarchical, master-slave way. Instead, they modulate bodily processes by being densely interconnected to them on multiple cellular

and molecular levels. Given this dense interconnectivity, it seems that neuronal and nonneuronal factors aren't unpluggable, and hence aren't explanatorily separable. As Damasio especially has emphasized, it makes little biological sense to talk about "what the brain does" versus "what the body does."<sup>43</sup> If the physiological system that supports creature consciousness comprises densely coupled neural, endocrine, and immune processes, and if it comprises sensorimotor loops through the body and the environment, then the biological basis of consciousness isn't brainbound.

The upshot of these considerations for the mere conceivability of a brain in a vat is that since we don't actually know what the minimal biological requirements are for creature consciousness, we don't know what we're supposed to imagine when we imagine a brain in vat, so the mere conceptual possibility of a brain in a vat seems an empty scenario.

### **Brainbound Revisited**

Where do these considerations leave us with regard to Block's claim that the "minimal constitutive supervenience base" for experience is the brain or more precisely the core NCC?

Here we need to be careful. Neural supervenience for experience requires that when the neural factors are held constant across some range of cases, then so is the experience. As Hurley observes, this kind of *internal supervenience* can hold even though the corresponding *supervenience thought experiment* isn't possible because internal factors and external factors aren't unpluggable. So the minimal supervenience of the subjective character of experience on brain states is compatible with the explanatory inseparability of

brain and body states. As Hurley also observes, however, the mere truth of internal supervenience provides no support for internalist explanation, if the relevant supervenience thought experiment isn't possible, because internalist explanation requires explanatory separability.<sup>44</sup>

The main moral of our examination of the brain in a vat is that brain states can't be unplugged from body states. So internalist explanation isn't a good framework for the neuroscience of consciousness. We should prefer Enactive to Brainbound.

But what about Block's disembodied, freestanding brain that spontaneously arises from the chance fluctuations of microphysical particles? We would like to know more about this brain. Is it just the neurons and synaptic connections, somehow floating independently of its glial cells, cerebral vasculature, immune cells, and other nonneuronal, somatic partners? That seems physically impossible, not merely highly unlikely. How would such a system hold together long enough for us to suppose any experience could be instantiated?

Even if we were to allow for purposes of philosophical argument the conceivability or bare possibility of some sort of freestanding brain, completely decoupled from the body, we see no reason to think there would any experience present at all. Why should we think that this kind of brain is intelligible as a subject of experience or possessor of creature consciousness? Neuroscience and biology in general give us no reason. We see no good scientific motivation whatsoever for the idea. It strikes us as simply a holdover of the flight from Cartesian dualism.



Here, then, is the enactive response to Block. If creature consciousness is a life-regulation process of an organism, and if perceptual consciousness is a certain kind of interactive relationship between an organism and its environment, then a disembodied brain going through the same sequence of internal states as an embodied brain is like a disembodied stomach going through the same sequence of internal states as an embodied one. The disembodied stomach isn't digesting and the disembodied brain isn't experiencing, because the necessary contexts of the body and the environment are missing.

### **The Bandpass Argument**

There's one more argument for Brainbound we would like to consider—Andy Clark's "bandpass argument." Here's the argument in his words:

[I]t is plausible that speed (or fine temporal issues more generally) makes a crucial difference in the moment-by-moment construction of conscious experience itself. Thus suppose conscious experience requires cortical operations that involve extremely precise temporal resolutions, such as the synchronous activations of distinct neural populations where the required synchrony requires millisecond precision... In such cases the external environment may well *matter* insofar as it drives the neural systems, but the key effects that enable and explain the quality of felt experience may then be occurring at time-scales that are only possible within the neural apparatus itself. If this

were so, then everything that involves subsequent motor actions or bodily action (including active saccades around the scene) will be “screened off” (by the bodily “low-pass filter”) from the neural/CNS mechanisms that actually produce the conscious experience.<sup>45</sup>

The argument is that the extra-neural body acts as a low-pass filter for signals coming from the environment—in other words, the body admits only slow frequency signals to arrive at the brain—but the contents of conscious experience require fast time-scales on the order of milliseconds, and the only locus where such fast frequency processing can occur is inside the brain. So brain processes directly determine conscious experience, and the body contributes only causally but not constitutively to consciousness.

We see several problems with this argument.

First, there’s a basic empirical problem. The time it takes for visual stimulation to pass through the lens of the moving eye and reach the first stages of visual processing is a fraction of the time it takes for neural systems to build up any correlated activity (from retina to early visual areas, and then to recurrent loops with higher visual areas and long-range coordination with parietal and frontal regions). So treating the body as a low-pass filter in relation to the brain doesn’t work.

Second, Clark seems to be working with a conception of perceptual experience that enactive theorists reject. He seems to conceive of a perceptual experience as an internal state—or the content of a perceptual state as the content of an internal representation—whereas enactive theorists argue that to perceive

is to be in an interactive relationship with the world. From the enactive perspective, all the bandpass argument shows is that fast frequency brain processes are a *necessary* part of the biological basis of the interactive relationship that constitutes perception, but not that these processes are minimally *sufficient* for conscious perception (or directly determine or constitute conscious perception). Indeed, it's hard to see how the argument could show that fast frequency synchronous oscillations directly determine perception unless one has already assumed that perception is an internal episode rather than an interactive relationship with the world.<sup>46</sup>

Finally, Clark focuses entirely on phenomenal state consciousness and neglects creature consciousness. Although modality-specific perceptual contents come and go on a fast time-scale, they do so against the much more slowly changing and domain-general background state of waking creature consciousness. The crucial concern of the enactive view is to account for creature consciousness. Clark, however, says nothing about creature consciousness and seems not to recognize that it's the crucial issue. In any case, the bandpass argument says nothing against the enactive view that the body belongs to the minimal biological requirements for creature consciousness.

## **Conclusion**

The main conclusion we would like to draw from this paper is that we shouldn't saddle the neuroscience of consciousness with Brainbound or what Block presents as the metaphysics of the orthodox view. Research on the neural correlates of consciousness provides a wealth of interesting and important findings about the brain. It's a mistake, however, to think that this research

amounts to a search for a minimal constitutive supervenience base for experience in the brain. On the contrary, this kind of heavy philosophical interpretation goes way beyond anything neuroscience gives us reason to believe. More importantly, it also hinders research because it prematurely shuts down a whole range of important and wide-open questions about brain-body relations in the biological realization of consciousness.

One of these questions—the one that’s concerned us here—is whether the brain alone suffices for creature consciousness or whether the body is also required. We’ve given reasons to think that the body and brain are so dynamically entangled in the causation and realization of consciousness as to be explanatorily inseparable. If these reasons are sound, then we should prefer Enactive to Brainbound. In other words, we shouldn’t expect there to be a purely neural explanation of consciousness. Instead, what we should expect—or at least aim for—is a much richer biological account of consciousness as a life-regulation process of the whole organism dynamically engaged in its world.

---

## Notes

<sup>1</sup> For helpful comments and criticism, we are indebted to the Philosophy and Neuroscience Reading Group at the University of Toronto, especially Stephen Biggs, Nigel Francis, Matthew Fulkerson, Mike Lachelt, and Adrienne Prettyman; thanks also to Tim Bayne, Dorothee Legrand, Thomas Metzinger, and Mog Stapleton.

<sup>2</sup> See Evan Thompson, *Mind in Life: Biology, Phenomenology and the Sciences of Mind* (Cambridge, MA: Harvard University Press, 2007). See also Diego Cosmelli and

---

Evan Thompson, "Embodiment or Envatment? Reflections on the Bodily Basis of Consciousness," in John Stewart, Olivier Gapenne, and Ezequiel Di Paolo, eds., *Enaction: Toward a New Paradigm for Cognitive Science* (Cambridge, MA: The MIT Press, 2010), pp. 361-385.

<sup>3</sup> We borrow the term "brainbound" from Andy Clark, *Supersizing the Mind: Embodiment, Action, and Cognitive Extension* (Oxford: Oxford University Press, 2008). Clark uses it as a name for the thesis that the material vehicles of cognitive processes reside exclusively inside the skull.

<sup>4</sup> Ned Block, "Review of Alva Noë, *Action in Perception*," *Journal of Philosophy* CII (2005): 259-272.

<sup>5</sup> Alva Noë, *Action in Perception* (Cambridge, MA: The MIT Press, 2004).

<sup>6</sup> *Ibid.*, p. 25.

<sup>7</sup> Block, *op. cit.*, p. 264.

<sup>8</sup> See David J. Chalmers, "What is a Neural Correlate of Consciousness?," in Thomas Metzinger, ed., *Neural Correlates of Consciousness* (Cambridge, MA: The MIT Press, 2000), pp. 18-39.

<sup>9</sup> Christof Koch, *The Quest for Consciousness: A Neurobiological Approach* (Englewood, CO: Roberts and Company, 2004), p. 16.

<sup>10</sup> See Tim Bayne, "Conscious States and Conscious Creatures: Explanation in the Scientific Study of Consciousness," *Philosophical Perspectives* 21 (2007): 1-22.

<sup>11</sup> Here we follow Tim Bayne in not defining creature consciousness as wakefulness, but rather as being a phenomenally conscious creature. On the one hand, phenomenal consciousness can occur during sleep, notably while

---

dreaming. On the other hand, patients in a vegetative state can be awake but apparently lack phenomenal awareness.

<sup>12</sup> See Bayne, *op. cit.*, pp. 13-14.

<sup>13</sup> See G. Rees, G. Krieman, and C. Koch, "Neural Correlates of Consciousness in Humans," *Nature Reviews Neuroscience* 3 (2002): 261-270.

<sup>14</sup> See J. Parvizi and A. Damasio, "Consciousness and the Brainstem," *Cognition* 79 (2001): 135-159

<sup>15</sup> See Chalmers, *op. cit.*, and Ned Block, "Two Neural Correlates of Consciousness," *Trends in Cognitive Sciences* 9 (2005): 46-52.

<sup>16</sup> See Sidney Shoemaker, "Some Varieties of Functionalism," *Philosophical Topics* 12 (1981): 93-119, Robert A. Wilson, "Two Views of Realization," *Philosophical Studies* 104 (2001): 1-30.

<sup>17</sup> Block, *op. cit.*, p. 47.

<sup>18</sup> See Robert A. Wilson, *op. cit.*

<sup>19</sup> Ned Block, "Review of Alva Noë, *Action in Perception*," p. 265.

<sup>20</sup> Thomas Metzinger, *Being No One: The Self-Model Theory of Subjectivity* (Cambridge, MA: The MIT Press, 2003), p. 547.

<sup>21</sup> Here we draw from Cosmelli and Thompson, *op cit.*

<sup>22</sup> Daniel C. Dennett, "Where Am I?" in *Brainstorms: Philosophical Essays on Mind and Psychology* (Cambridge, MA: MIT Press/A Bradford Book., 1978), pp. 310-322.

<sup>23</sup> Daniel C. Dennett, *Consciousness Explained* (Boston: Little Brown, 1991).

<sup>24</sup> *Ibid.*, p. 4.

---

<sup>25</sup> See P.D. Brown, S.L. Davies, T. Speake, and I.D. Millar, I. D. "Molecular Mechanisms of Cerebrospinal Fluid Production," *Neuroscience* 129 (2004):957–970; H. Davson and M.B. Segal, "Secretion and Drainage of the Cerebrospinal Fluid," *Acta Neurologica Latinoamericana* 1 (1971): Suppl 1: 99-118; M.B. Segal, "Extracellular and Cerebrospinal Fluids," *Journal of Inherited Metabolic Disease* 16 (1993):617-638.

<sup>26</sup> See F. Hyder, R.G. Shulman, and D.L. Rothman, "Regulation of Cerebral Oxygen Delivery," *Advances in Experimental Medicine and Biology* 471 (1999): 99-110; M.E. Raichle and H.L. Stone, "Cerebral Blood Flow Autoregulation and Graded Hypercapnia," *European Neurology* 6 (1971): 1-5; R.G. Shulman, F. Hyder, F., and D.L. Rothman, "Biophysical Basis of Brain Activity: Implications for Neuroimaging," *Quarterly Review of Biophysics* 35 (2002): 287-325.

<sup>27</sup> See N.K. Logothetis and J. Pfeuffer, "On the Nature of the BOLD fMRI Contrast Mechanism," *Magnetic Resonance Imaging* 22 (2004): 1517-1531.

<sup>28</sup> See Parvizi and Damasio, *op. cit.*; A.D. Craig, "How Do You Feel? Interoception: The Sense of the Physiological Condition of the Body" *Nature Reviews Neuroscience* 3 (2002): 655-666; Antonio R. Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness* (New York: Harcourt, Inc., 1999); J. Panksepp, "The Periconscious Substrates of Consciousness: Affective States and the Evolutionary Origins of Self," *Journal of Consciousness Studies* 5 (1998): 566-582; C.B. Saper, "The Central Autonomic Nervous System: Conscious Visceral Perception and Autonomic Pattern Generation," *Annual Review of Neuroscience* 25 (2002):433-469.

<sup>29</sup> Parvizi and Damasio, *op. cit.*

---

<sup>30</sup> Antonio Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: Harper Perennial, 1995), p. 228.

<sup>31</sup> Lawrence A. Shapiro, *The Mind Incarnate* (Cambridge, MA: The MIT Press, 2004), p. 218.

<sup>32</sup> *Ibid.*

<sup>33</sup> Dennett, *Consciousness Explained*, *op. cit.*, p. 5.

<sup>34</sup> See H. Chiel and R. Beer, "The Brain Has a Body: Adaptive Behavior Emerges from Interactions of Nervous System, Body and Environment," *Trends in Neurosciences* 20 (1997): 553-557; V. Brezina, C.C. Horn, and K.R. Weiss, "Modeling Neuromuscular Modulation in *Aplysia*. III. Interaction of Central Motor commands and Peripheral Modulatory State for Optimal Behavior," *Journal of Neurophysiology* 93 (2005): 1523-1556.

<sup>35</sup> Susan Hurley, "The Varieties of Externalism," in Richard Menary, ed., *The Extended Mind* (Cambridge, MA: The MIT Press, 2010), pp. 101-154.

<sup>36</sup> See R.D. Fields, "The Other Half of the Brain," *Scientific American* 290(4) (April): 54-61.

<sup>37</sup> See S.F. Maie and L.R. Watkins, L.R., "Cytokines for Psychologists: Implications of Bi-Directional Immune-to-Brain Communication for Understanding Behavior, Mood, and Cognition. *Psychological Review* 105 (1998): 83-107; J. Scholz and C.J. Woolf, "The Neuropathic Pain triad: Neurons, Immune Cells, and Glia," *Nature Neuroscience* 10( 2007): 1361-8.

<sup>38</sup> Jesse Prinz, "Is Consciousness Embodied?" in Philip Robbins and Murat Ayde, *The Cambridge Handbook of Situated Cognition* (Cambridge: Cambridge University Press, 2009), pp. 419-436.



---

<sup>39</sup> For critical discussion of these findings, see Alva Noë and Evan Thompson, "Are There Neural Correlates of Consciousness?," *Journal of Consciousness Studies* 11: 3-28, and "Sorting out the Neural Basis of Consciousness. Authors' Reply to Commentators," *Journal of Consciousness Studies* 11: 87-98.

<sup>40</sup> John R. Searle, "Consciousness," *Annual Review of Neuroscience* 23: 557-578.

<sup>41</sup> See Bayne, *op. cit.*

<sup>42</sup> See Parvizi and Damasio, *op. cit.*

<sup>43</sup> Damasio, *The Feeling of What Happens*, *op. cit.* See also Shapiro, *op. cit.*, p. 217.

<sup>44</sup> Hurley, *op. cit.*, pp. 7-8. Hurley and Noë deny internal supervenience. By appealing to cases of neural plasticity, they argue that experiences with the same phenomenal quality can have variable neural correlates. Fixing the neural states does not fix the experience, for the phenomenal quality depends on the dynamic sensorimotor pattern in which the neural states participate. For example, the neural correlates of seeing an object as white differ before and after adaptation to colored goggles that distort the wavelengths reaching the eye as a function of eye movement or object movement across the midline. Although we agree that dynamic sensorimotor accounts of perception provide better *explanations* of such phenomena than purely neural accounts, we are not persuaded that such phenomena count against internal supervenience as such. What look like variable neural correlates could still share some higher-order neural invariance, which might then qualify as a supervenience base for the experience. We cannot see that the arguments from plasticity as developed so far rule out this possibility.

---

<sup>45</sup> Andy Clark, "Spreading the Joy: Why the Machinery of Consciousness Is (Probably) Still in the Head," *Mind* 118 (2009): 963-993, at pp. 984-985.

<sup>46</sup> For a different view of the role of synchronous neural oscillations in perception, see Evan Thompson and Francisco J. Varela, "Radical Embodiment: Neural Dynamics and Consciousness," *Trends in Cognitive Sciences* 5 (2001): 418-425.