

Embodied Consciousness and the Explanatory Gap (to appear in *Psyche*)

Charles Siewert
Department of Philosophy
University of California, Riverside
Riverside, CA, 92521
USA
© C. Siewert
siewert@ucr.edu

COMMENTARY ON: Evan Thompson, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Harvard, 2007, 543 pp., ISBN-13: 978-0-674-02511-0.

ABSTRACT: Thompson's *Mind in Life* aims to provide new resources for addressing the explanatory gap by recognizing a category of "comportment" or embodied consciousness that falls outside the usual schemes for conceptualizing consciousness and its place in nature. I consider how this distinguishes Thompson's view from more familiar approaches to the problem of the gap. I then raise a challenge for his approach, and his "phenomenological critique of zombies." I suggest a response to this challenge, by arguing that there are basic phenomenological concepts of experiential movement that cannot be factored into purely phenomenal and purely behavioral components—such as the concept of *looking at* something. Finally, I raise a question about how Thompson views the relationship between two kinds of "inwardness": the sort he says is a "precursor" to consciousness and can be found even in fairly primitive organisms, and the sort of "inwardness" that suffices for actual consciousness, and makes an entity a suitable focus of empathy.

1. The Problem of the Explanatory Gap

Evan Thompson's *Mind in Life* brings together a diverse array of themes and ideas drawn from classical phenomenology, theoretical biology, and dynamic systems theory. It is remarkable how Thompson synthesizes these into a coherent whole. I will not (and I can't) try to survey this whole, or match the breadth of his discussion in my comments. Instead I will focus on some pivotal aspects of his manner of meeting the challenge that frames the discussion—the challenge of the so-called "explanatory gap" of consciousness and nature.

Thompson does not purport to have *solved* the problem of the explanatory gap; there is no grand claim that consciousness has been explained. But he aims to make a substantial start on a new approach, to "provide new resources for addressing the explanatory gap," as he says in his Précis. I am sympathetic to the thrust of what I take him to be saying, and I want to make more explicit what

is involved in a view of the type Thompson wants to advance (or at least, to nudge him into being more explicit about this). For I want to see better how a view of this kind might be elaborated and defended.

What I am going to say about the problem of the gap clearly owes much to Joseph Levine's and David Chalmers' treatments of these issues—though, naturally, my exposition will be much rougher than theirs. I only hope it will be clear enough for my main purpose here: to help us to appreciate and begin to critically consider the kind of approach Thompson is suggesting.

First, let me give some orientation to what I will say by indicating briefly some of my assumptions. We are capable of “first-person reflection on experience”—that is, we have a way of forming thoughts in the first-person about our own experience that differs significantly in kind from that we have of forming thoughts about others' experience. *Experience* here is taken to include, at a *minimum*, certain kinds of sensory states, such as appearances of color, tastes, smells, and bodily feelings—but much richer conceptions of the range of experience, including conceptual thought and higher emotion, are by no means excluded. To have experience is to have *conscious* states—conscious in the “phenomenal” sense. We may call first-person reflection on experience “introspection”—as long as our use of that term leaves open the question of what exactly is distinctive about this “first-person perspective.” And an approach to inquiry into the general character of experience that relies in a fundamental way on such first-person reflection, I will call—and I don't think this is at odds with Thompson's usage—“phenomenological.”

Now let me sketch a way of thinking about the problem of explaining consciousness which, as Thompson remarks (pp.222-3), has become influential. First-person reflection on our own experience involves the use of certain concepts (which we may call “phenomenal” concepts)—concepts we have, for instance, of ways things look or feel to us, or more generally, of “ways of seeming.” And these concepts we have of our own experience—of forms of “phenomenal consciousness”—are distinct from the concepts we deploy in theories of our brain activities and the functional or causal roles these activities play. This is evident in that there seems to be no *entailment* leading us from application of the latter (physical/functional) concepts to application of the former (“phenomenal”) concepts.

One way of trying to make this failure of entailment particularly vivid is to invite us to conceive of a world type identical to our own at the level of fundamental physics, but from which the forms of experience we understand first-personally (ways of looking, feeling, tasting) have gone missing. Paraphrasing Chalmers: in such a conceivable world the same types of fundamental physical particles that are found in our world are distributed in fields of force in the same manner, conforming to the same physical laws as our own. Nonetheless, these particles are imagined not to make up *feeling beings* such as we ourselves are

(and such as we normally suppose our monkeys, dogs and kangaroos to be). We are to suppose there is nothing it's like to be our analogs in this imagined world, nothing it is like *for* them, any more than we usually suppose there actually is something it's like to be a coffee machine or a pile of hair. This is the idea of a world type identical to our own at the microphysical level, but totally bereft of feeling, and of phenomenal experience generally—a world full of movement, but *dead to consciousness*: a “zombie world.”

Use of the word ‘zombie’ in this context can invite an easy derision. But we shouldn't allow the underlying point to be dismissed. That our first-person understanding of experience employs concepts that permit such reflections points to a distinctive challenge facing efforts to construct a scientific theory of consciousness. For this may lead us to think that our usual strategy for the theoretical explanation of nature can't be applied here. Usually (so the story goes), we satisfy our scientific ambitions by conceiving of what is to be accounted for in terms of what it *does* (the role it plays, its function), and then seeking out and finding *physical mechanisms* that do that job, play that role, fulfill that function. Thus it is with boiling and freezing, raining and burning, respiration, digestion and reproduction. Once we see how—given the character of the underlying physical structures and their activity—the function in terms of which the explanandum is identified *can't help but get carried out*, we will have explained what we wanted to explain.

The problem is, when it comes to consciousness, our understanding of what is to be explained seems to accord it no comfortable place in such a model. That is what the fantasy of a zombie world is supposed to make vivid: you have the structures and you have whatever functions can't help but get carried out when such structures are operative—but *must* you then have consciousness? There is no evident necessity that you must, since our subjective understanding of consciousness seems to make a consciousness-deprived physical analog of our world conceivable. Thus it seems we will inevitably be drawn up short in our efforts at explanation, at least as long as we hope to make our explanation of consciousness follow the same general pattern we find elsewhere in natural science. For we can neither see how justifiably to *identify* types of conscious experience with forms of activity peculiar to specific physical structures, nor can we honestly see how it *must* arise, as long as those structures are fulfilling certain functions in a larger picture.

The challenge of responding to considerations such as these is, I take it, the challenge of the explanatory gap. Certain general forms of response to the challenge, so understood, can be distinguished. Before considering Thompson's, let's review three main options with which it can, I think, be usefully contrasted.

Option (1). We have no concept we are warranted in applying to our own experience that would allow us to conceive of a “zombie” world in contrast with our own. For any concept we have a *right* to apply to our own experience really

marks nothing but a *purely functional* difference—and there would be no such difference between our world and a microphysical twin world. Since there is ultimately *no conceptual difference of the sort alleged*, neither is there an explanatory gap—and, once the confusions and illusions are cleared away, there are no obstacles to explaining consciousness in pretty much the same way science successfully explains other things. (In rather different ways, this is the sort of response found in Daniel Dennett, and implicit in conceptual functionalists like David Armstrong.)

Option (2). One can (by contrast) accept there is indeed a disparity in our (legitimately applied) concepts (perhaps made vivid by various thought experiments), but deny that it inevitably brings with it an explanatory gap of the form alleged. It would be a mistake to think that since a legitimate first-person understanding of our own experience does not conceive of it in terms of “structure” or “function,” there *is something more* than structure and function in reality for us to conceive of. (Rather different forms of *this* response can be seen in Michael Tye and Ned Block.)

Option (3). One can accept that consciousness should be seen differently than other targets of scientific explanation to which some had sought to assimilate it. To do this we needn’t abandon a broadly scientific, “naturalistic” view of the world. We just have to acknowledge that the psychophysical principles governing consciousness, while open to scientific discovery, are at bottom *contingent* laws relating the phenomenal and the physical, and their truth does not admit of deeper explanation by derivation from yet more fundamental principles. (A form of this view has been defended by Chalmers.)

This vision of the challenge of explaining consciousness (and of our options in response to it) can seem compelling, even inescapable. Part of what is interesting about Thompson’s book is that, while he accepts there is a deep and important scientific challenge in accounting for consciousness in its various forms, and he wants to make a start on meeting it, he rejects the now fairly familiar types of responses to it that I have just sketched—and wants to scout out another. The sort of view he is putting forward is not entirely unprecedented, since it is avowedly inspired in part by the philosophies of Husserl, Merleau-Ponty and others. But it seems to represent a distinctive option missing from the usual maps of this territory.

In my opinion the effort to articulate some such alternative deserves careful and sympathetic attention. For the other options seem to me to varying degrees problematic. First, as I have spelled out elsewhere, there are reasons I find convincing for rejecting option (1). We can—even without contemplating “zombie worlds”—see that our understanding of our own experience, and of how things appear to us, is not to be construed purely in terms of functional differences that are ultimately characterizable in ways that eliminate all talk of appearance. Or rather, if we do insist on such construals, we end up eliminating

phenomenal consciousness from our conception of reality altogether. (And that, I think, would be a false economy, to put it mildly.) Second, I can't see how option (2) ultimately manages to avoid trying to get deep, more than nomological (metaphysical) necessities "for free" so to speak. If there is nothing about our conception of the physical substratum of consciousness whence experience and its phenomenal character can be derived, and various functional notions fail to capture such contrasts as our phenomenal concepts allow, it's hard for me to see what entitles us to the relevant claims of strong necessitation here, or how justifiably to determine exactly what's minimally necessarily sufficient for it in the required, strong sense.

So, in the absence of palatable alternatives I would incline towards option (3). But (as I will try to bring out momentarily) I worry that something has gone wrong in this strict and exhaustive dualist division between a purely phenomenal realm (answering to our phenomenal concepts), and a purely physical realm (which is logically determined by the sorts of microstructure described in physics). What I fear this leaves out is what Thompson in his *Précis* (following Merleau-Ponty) calls "comportment," or what I will call here "embodied consciousness." I will devote the rest of this essay to exploring how recognition of this category might give Thompson a distinctive take on the explanatory gap.

2. Thompson's Response to the Gap: Embodied Consciousness

There may be aspects of Thompson's approach that are not yet sufficiently clear to me. However, I think it will be helpful to start by seeing him as advancing the following three ideas.

(a) Part of what we need to do to understand how consciousness arises, is to show why the idea of a "zombie world" that is *functionally just like* our own is a misconceived legacy of the Cartesian dissociation of mind and life (pp. 222-34).

(b) However, we do this *not* by showing that a proper understanding of experience consists in purely functional notions, which will then leave us free to apply some general structure-and-function model for explaining consciousness. (See pp. 224-5.) Nor do we diminish the role of phenomenology. Rather, we show that *a proper phenomenological understanding of perceptual experience finds it to be so bound to embodied activity that, in a non-trivial sense, the same perceptual functions that occur in our world would not occur in a world in which there were no experience.* Even if it's not correct to say that phenomenal differences are "nothing but" functional differences, we can see that there is, after all, in a sense, no possible truly functionally *equivalent* world without consciousness. For these types of embodied activity must be missing, if consciousness is missing. This is the upshot of Thompson's "phenomenological critique of zombies," which draws on Husserl's (phenomenologically based) view that "...it is through one's movement and bodily self-experience in movement that

an object presents itself as a unified series of appearances.” (p.232) Thompson argues that if this is right, “...then bodily experience is constitutive of the perceptual function of individuating continuous objects in space through a manifold of sensory appearances. So any being that was capable of the same perceptual function would need to have experience of its own body and hence could not be a zombie.” (p. 233)

(c) Once we have banished the specter of a functionally equivalent zombie world, we have opened a way around the mind-body impasse. For we no longer have to regard a personal level, phenomenological understanding of experience as *radically discontinuous* with the forms of understanding proper to the natural sciences—particularly the life sciences. For we can now see that biological systems exhibit a kind of self-organizing, normatively intelligible activity that, at some level of sensorimotor engagement at least, deserves to be considered a form of (and literally to *embody*) full-blown phenomenal consciousness. (This is a view for which Thompson lays the groundwork mainly in Chapter 6.)

For ease of exposition, I need a label for the view characterized by (a)-(c). But while I do think this view can be attributed to Thompson, I hesitate to employ his own doctrinal labels here (such as “the enactive approach” or “embodied dynamicism”) since these include additional aspects of his position which, while important, I need to leave aside to make my discussion manageable. Let me then call the position roughed out in (a)-(c) the “Embodied Consciousness” View.

There are of course many questions we might raise about this view. But let’s start by noting how it diverges from perspectives on the explanatory gap earlier mentioned. One might at first mistake Embodied Consciousness for a version of option (1). For in a sense both would hold that a proper understanding of the concepts we rightfully apply to our own experience (even in “introspection”) will not leave room for us to conceive of a world truly functionally equivalent to our own, from which all experience has been expunged. However (as I read Thompson) there is an important difference. On his view, while we cannot properly conceive of an experience-free world where everything “functions” just the same, this is *not* because (as friends of (1) would maintain) our only legitimate concept of experiences would make them nothing but functional role fillers of some sort. Rather it is because *our forms of bodily engagement with the environment cannot be properly understood but as ways of experiencing oneself and things in one’s environment*. Neither our ways of experiencing, nor our forms of bodily interaction are rightly understood—even phenomenologically— independently of one another. Thus any world that had in it the *very same types of bodily engagement* as we have (and in *that* sense was “functionally” the same) could not help but be a world in which there was consciousness. So it couldn’t be a zombie world.

Now we can see that this view also differs from responses (2) and (3) to the alleged gap problem. Like (2), Embodied Consciousness maintains there is

no possible world functionally like ours but missing consciousness, even while allowing (unlike (1)) that we can legitimately understand what sorts of experiences we have in first-person reflection in ways that cannot be exhaustively “functionalized.” However, the Thompsonian view gives such a phenomenological understanding a more robust role in conceiving of the nature of experience and its necessary connection to bodily activity. According to (2), we cannot rule out, on *conceptual* grounds, a functional twin world without consciousness—the point is just that this fact lacks ontological import. By contrast, Embodied Consciousness, like (3), takes a phenomenological conception of experience more seriously as a guide to what experience *is*. But, unlike (3), it finds this perspective on experience does not support, but ultimately casts doubt on the possibility of a zombie world—at least *on a certain construal*.

That last qualification—“on a certain construal”—needs a little more attention. There is a sense in which the kind of connection Thompson would make between a phenomenological understanding of perceptual consciousness and embodied activity apparently would *not* preclude the conceivability (or even the possibility) of a zombie world. If a zombie world is simply taken to be a world the same as this (our consciousness-endowed) world with respect to the types of fundamental particles and their distribution—but just *minus all consciousness*—I don’t yet see how Thompson’s view would make it any less conceivable. All it would render inconceivable is the idea that vibrating swirls of particles in such a world would constitute *the very same types of embodied activity* as are to be found in our world. There would be in this experience-deprived world no “zombie twins” *doing just what we are doing*. While a zombie world could have in it the same types of *mere motion* as are found in our world (in virtue of having the same distribution of particles in fields of force), it would not—absent consciousness—contain (to speak with Merleau-Ponty) the same (or any) forms of “*comportment*.” Nothing in the zombie world would *have our being-in-the-world*. Though entities located in that world would *move around in space*, none would *have a world* at all.

Now that we start to situate a view like Thompson’s vis-à-vis options (1)-(3), we can start to see what difficulties need addressing. At this point friends of (3) may begin to suspect that no real alternative is being offered to their view after all. It is true that we could stipulatively define types of bodily activity in such a way that tied these, with conceptual necessity, to phenomenal experience. But that idle exercise ultimately does nothing to remove the brute contingency of the psychophysical relationship. Consider: we might, for example, define a category of “*painfultoestubbing*” so that it involved both a kind of movement (for example, jamming one’s toe into the bed post) and a kind of (painful) feeling—a form of phenomenal consciousness. And we may then say that *painfultoestubbing* is a form of behavior not to be found in the zombie world. But, we may then note: that is just for the *trivial* reason that *feelings of pain* would not be found there—the other part of this contrived category of “*painfultoestubbing*” (the collision of toe with bedpost) *would be* there in the zombie world. Thus ultimately, we are still left

with a brutally contingent causal relationship between movement and feeling. The lesson is, until we have explained why these forms of bodily engagement with the world mentioned in (b) above (comportment or what have you) cannot be factored into a “phenomenal, experiential bit” and a “bodily movement bit” with no more than a contingent causal link between them, then the Embodied Consciousness view has not successfully articulated a real alternative to option (3). Once more we are ultimately left with a picture of some (perhaps interestingly systematic, but entirely contingent) linkage between (on the one hand) phenomenal features and (on the other) features logically fixed by the distribution of physical particles—an ontological picture which, seen from a certain distance, is not so very different from Descartes’.

Now if this constitutes a genuine challenge to Thompson’s view, I cannot yet find a clear answer to it in his “phenomenological critique of zombies.” Consider Thompson’s gloss on Husserl’s account of the “perceptual function” that would be missing in a zombie world: that of “individuating continuous objects in space through a manifold of sensory appearances” through “movement and bodily self-experience.” It seems that we are merely defining the function in a way that packs in phenomenal consciousness (the manifold of sensory appearance)—and so we risk trivializing the claim that the function could not be found in zombiedom. And while it may be that both movement and bodily self-experience play a role in executing this function, it is unclear why the linkages between experience and movement in this account need be regarded as anything but causal relations among separate existences. Why not suppose zombies to make *the very same types* of movements, just without their being hooked up to manifolds of sensory appearance and bodily self-experience?

The challenge that emerges is this. In order to show why his view really is an alternative to (3), and why it is preferable, Thompson needs to show why, from a phenomenological viewpoint, we distort the character of fundamental cognitive activities (such as object constancy in perception) if we analyze them into quite separate “mental” (or “phenomenal”) features and mere bodily movement types.

Maybe Thompson’s position is that the burden of proof lies on anyone who would claim they can be thus analyzed. But given how entrenched the quasi-Cartesian “factoring” view is in the theoretical imagination, it seems we should try to do more. I want now to suggest a way of meeting this need which, while it departs from direct engagement with the Husserlian (and Merleau-Pontyan) ideas on which Thompson draws, nonetheless concerns the same crucial phenomena, having to do with how the experienced (“lived”) body figures in our perception of constant objects through the “flux of appearance.” But I will try here to develop some basic ideas in this area without reference to their writings or terminology, and using (fairly) plain English. Of course, readers are free to try to map what I say onto Husserlian talk of “horizons,” “the indeterminately given,” “das Leib,” and “anticipation” as they will. But what I say is intended to stand on

its own. I will leave it to Thompson to explain whether he thinks my suggestions are helpful, mistaken, or just unnecessary.

3. A Defense of Embodied Consciousness

Let us consider two concepts, pervasively applicable in our lives, in which experience and movement are intertwined, but not in some apparently readily decomposable way (as in the artificial “painful to estubbing” example). What I have in mind are concepts of *looking* and *touching*.

We may readily enough understand talk of our looking at something or looking for something, so that this entails that something then *looks* somehow to us. (We may also then say: something visually appears somehow to us, we have visual experience of some kind, some kind of visual, phenomenal consciousness). And while it's true that we may say that *x* is touching *y*, even when *x* feels nothing (e.g., ‘the fork is touching the plate’), nonetheless, where the ‘touching’ is understood as *asymmetrical* (‘*x* touches *y*’ does not entail ‘*y* touches *x*’)—as in: ‘I touched the lamp’—the subject of the touching feels something somehow. At least, we may easily interpret ‘touch’ in that way. And looking and touching, so understood, are activities we engage in nearly constantly while awake, and through which we experience space. But while these are thought of not only as involving experience, but also bodily movement (e.g., of eyes, head, limb and hand)—it is not so clear the understanding I have of looking and touching from the first-person point of view, hence a phenomenological understanding, can be readily segregated into a *purely experiential* part, and a *mere movement* part, seen as only contingently, perhaps causally related. In fact, on examination, it seems that my understanding of changes in the character of the appearances I enjoy through looking and touching does not detach these from the activities involved. And likewise, my understanding of the types of movements I take to be involved in looking and touching does not tease them apart from the experiences they afford.

Let me elaborate on this idea, concentrating on the visual case. Again, I am looking at something or other most of the time I am awake, and what I am looking at and how much I am looking at it undergoes constant transformation. And in thinking of myself as looking at things, I think of myself as moving somehow—my eyes, my head, my entire body. Further, I can think of the variations in how things look to me by thinking of them as *those variations involved in differently looking at things*. The question is whether I can, in first-person reflection, factor out variations in the *experiential* part of “looking at” (the “how things look” part) and understand it quite separately from a *movement* part also essential to “looking at.” Can I completely capture in first-person thought how appearances differ for me depending on whether and *in what way I am looking at things*, but without thinking of this at all *in those terms*? The challenge here is to think of how the appearance of the area before me differs as I look at it

differently, but *without* thinking of myself as looking at things, and thus as *moving* somehow.

Let's consider how this might apply to an example. If I am looking at something—the plate on the table at which I sit—it looks somehow to me. But also, then, typically: the area *around* what I am looking at (e.g., where the fork and spoon, and the table surface are), and areas *within* what I am looking at (e.g., that of the designs around the edge of the plate), also look somehow to me—even while I am *not* looking *at* anything in those areas. And as what I am looking at changes—as it does nearly constantly—the way the area before me looks changes. As my gaze shifts from the plate to the fork then to some chipped bit of the plate's rim on the opposite side, appearances are changing. These changes are evident to first-person reflection. Now can I conceive of the area's appearance altering in just the way it does as I look at it, but without thinking of myself as *looking at* it, or at anything in that area?

What I need here is some finely discriminating grasp of how things look to me, quite detachable from my understanding of myself as a motor agent. I need, it seems, some way of thinking of *looking at* as two quite separable things: a movement, and a series of representations that this movement causes (which series then in turn may perhaps be used to control the movement), where the manner of visual appearance corresponds to the representations. Perhaps I can do this by finding some verbally expressible content in vision, which content possession I can understand independently of how it happens to be generated by motor activity. Or perhaps how things look to me can be made out as some visual "picture" or "image" of the scene, which I can think of as the product of my movements, but whose character I can adequately conceptualize in isolation from them.

Perhaps. Let's examine this more closely. What grasp do I really have conceptually of these changes in appearance? First, let me try to characterize them by *attributions of distinct verbally expressible content* to my experience over time, using common nouns, adjectives, demonstratives of ordinary language. (It looks to me as if there is a fork; that looks to me green; that (design) looks to me *this* shape). But this does not *exhaust* the alterations in the character of my experience of which I am aware as I look around in ordinary circumstances. In that still apparent region where the fork lies waiting for my gaze to light upon it, it either does not yet appear to me "as a fork" until I get a better look (and then, when I do, what is there appears to me differently than it did), or else if I have got a good enough look at it for it to appear to me as a fork, there are still many subjectively discernibly *different* ways a fork can appear to me "as a fork" depending on how I am looking at it. The point is that the variations in appearance either resist such verbalized content attributions, or exceed what is thereby differentiated. I would extend the point to whatever general terms one might find suitable for reporting how things look. For whatever lies in a visually apparent region, and whatever general 'F' one cares to consider, either it is not

(yet) correct to report any manner of appearance I enjoy by saying ‘...looks to me...F’ (for I haven’t looked at something well enough), or if it is correct, I will be able to find discernibly *different* appearances that may *both* be so reported, because of differences in how much or in what way I am looking at something.

Here it may be suggested that I can always capture this “fine grained” alteration in visual appearance by beefing up my linguistic resources with resort to *demonstrative* predicates (it looks to me “that shade,” “that shape”). So, every change in visual appearance—as I look from the plate’s edge to the fork, from its tines to its handle, to the fork as a whole, to the interior of the plate—can be marked: ‘*that* looks *this* shape to me, *that* looks *this* shape to me, *that* looks *this* shape to me’ where the different ‘that’s’ and ‘thises’ pick out distinct localizations of distinct shapes apparent to me and this tracks in a conceptually independent way exactly all the on-going variations in appearance of which I am aware in reflection.

But now: can you really succeed in capturing all the differences in appearance of which you are aware by means of such demonstrative iterations? That is, might you really manage to identify, with a suitably populous crowd of ‘thises’ or ‘that’s’, a set of distinct specific shapes that appear to modify a specific area before you in a way that completely coincides with every change in how the scene appears to you? There is no reason to think so, if you do not have even a fleeting capacity to recognize and distinguish the various “that’s” and “thises” allegedly discriminated. Do you?

Consider any reasonably complex and variegated visual field—such as a page of ordinary text seen with normal acuity. As your gaze moves across a line of text and down the page, how will you capture the continual variation in appearance, by saying (even demonstratively) *which* shapes appear to you *when*, both in the area where you are looking and where you are not? Singling out a set of specific shapes and locales individually that will take in the continual dynamic transformation of the entire visual field seems obviously hopeless. Maybe you will say, *in one grand gesture* encompassing an extended episode of looking: “*these* (various items) look to me *these* (various) shapes at *these* times.” But you must not pretend that mouthing such words somehow guarantees you know what you are talking about here—specifically *which* various shapes, or items, or times are in the class of “these.” Certainly here it will be clear that you have no capacity to recognize and distinguish all the varying shapes you would allegedly be picking out with this characterization. This “grand gesture” turns out to be an empty one. And from the fact that the way the shapes of what is before me appear varies over time it doesn’t follow that there simply *must* be some specific set of distinct shapes that maps onto this variation such that just these shapes appear to me at just these times in specific places.

At this point one might turn back to the idea that seeing is, after all, more like getting a “picture in one’s head” than having words in it. So perhaps

specifically *what spatial details are apparent when* can be expressed and identified by means of an *image*, in a way that detaches my understanding of the changing appearances from any notion of myself as engaged as looking *at* this or that. But any *image* also seems to fail at capturing the manner of appearance, in a way similar to that we just discovered. Either the candidate image will have more represented in it of the apparent region than was actually apparent to me, or the same image would also be offered to capture how it appeared to me, even where the manner of appearance was *not* the same (since, again, I am *looking at* the scene before me somehow differently). Once more, either way one has not found a manner of conceptualizing the appearance that is commensurate with its variation, and which detaches one's understanding of that variation from one's understanding of oneself as a *looker*.

So, in these cases I *can* judge that changes in appearance are the sort that occur when I am looking at something (or not looking at it, or looking at it *more*, or looking at it *less*). I can tell from a first-person point of view what I am looking at, whether I am looking at it more or less, and I can understand these changes in how something looks to me relative to this. But I think of looking at something as involving movement. What I still do not have then is an understanding of this variation in appearance that is entirely purified of any association with the activity of looking, and thus detached from my understanding of myself as motor agent. When I concern myself in first-person reflection with the ordinary changes that occur in the manner of visual appearance as my gaze shifts around and explores the area before me, what I find is not something whose character I can conceive of quite independently of the activity that affords it—a “picture in my head” whose contents and details alter as a result of movements in the visual apparatus. When I try to isolate a purely experiential component of “looking at”—a “receptivity” quite distinct from the “spontaneity” of visuomotor activity—I do not succeed. For phenomenologically I do not succeed in thinking of the variations in appearance in all their fleeting and fluctuating subtlety, while holding quite apart any notion of myself as looking at things.

But perhaps here it will be pointed out that, even if this shows us that we have no phenomenological grasp of the subtle dynamics of appearance independent of exercise of visuomotor skill that comprises “getting a look at” something—its shape, location, texture, size—this does not yet yield the result that our understanding of the type of *motor activity* in question cannot be pulled apart from our understanding of the kind of experience it makes available. Perhaps when we think of what we're doing as *getting a look* at the fork, or the design on the plate, we do have a concept of *that type* of motor activity, which would allow us to attribute the very same type of activity to a zombie or a robot, to whom nothing looks any way at all. It's just that we think of the experience we have in terms of its being (in us at least) caused by movement of *that type*.

But will this work? Note the question is not whether there is some way, in principle, of conceptualizing the history of movements of your eyes, head and so

on, that imputes no experience to you, and which could therefore correctly be attributed to some imaginary “zombie twin,” or perhaps some robot imitation of you, to whom nothing looked any way at all. The question is whether the way you have of conceptualizing your activity when you think of yourself as looking at something involves such concepts. Well, consider the movements you have made in reading this page of text. Do you have some grasp of just what movements you made, independently of understanding them as the movements required to look at (and thereby read) the words found on the page, from left to right, top to bottom one line at a time? If you had some such grasp, it seems you should be able to use it to follow instructions like these: close your eyes and repeat the same sequence of eye movements you just made reading this page. But this seems quite hopeless. On the other hand, you can, in a sense, make “the same movements” by *re-reading* the page if you’re allowed to *look* at it. This seems to indicate that your understanding of what you’re doing when looking depends on things looking to you as they do when you look at them, and in the absence of this you have no conception of just what you are doing when you are looking. You conceive of the movements as, irreducibly, *those required to read this text by looking at it*.

Maybe someone will be tempted by the thought: “Of course I have a conception of that way a region differently appears to me, as my gaze is drawn to it, as I look at it, quite detached from any notion of the direction of my gaze or my activity of looking. For *there*, I just expressed it—by saying ‘that way a region differently appears to me.’” But nothing about this form of expression shows that the claimed independence of conception is achieved. This deployment of a demonstrative device does not in and of itself indicate the presence of a totally separate conceptual purchase on what is spoken of. (Consider the anaphoric use of ‘this’ at the beginning of the previous sentence.) And notice that the demonstrative phrase at issue was first introduced with essential support from a conceptualization in terms of the *direction of the gaze*: ‘that way it appears differently as my gaze is drawn to it.’

Another objection may arise here. “Haven’t you read the *Meditations*? Descartes shows how I can conceive of myself having experience of the very same character as I am and have been having, even as I suppose myself to have no body at all—and to be a demon’s plaything. Or, leaving Descartes aside, equally apropos for current purposes: I may suppose myself to be a *brain in a vat*, not moving at all—though ‘from the inside’ the experience is the same.” But these sorts of considerations do not seem to me to show that we can, in first-person reflection, conceive of the variation in our visual appearance, without thinking of ourselves as looking at things, thus as moving. What they suggest to me instead is that, if I were (for example) a brain in a vat of the sort I am asked to envisage, I would be defective not only in my thoughts about my surroundings and my body, but also—at least to a significant extent—in my thoughts about the character of my own experience. And perhaps it reveals that when I consider the matter more closely I see that I do not succeed in thinking *of my own current and*

actual situation as that of a disembodied brain, unless I could stop thinking of myself as looking at (and touching) things. But I cannot do that without losing a grip on the character of my own experience. It should perhaps be no great surprise that, when it comes down to it, I cannot really think of my current situation as that of the brain in the philosophical fable.

But even if I cannot make out the character of *my own* experience, while consistently conceiving of myself as a disembodied being, may I not conceive of *some* disembodied being (brain in vat or Cartesian ego) having experiences of the same character as mine? I can consider *what it's like for me to look* at things before me, and entertain the notion of a disembodied being having experience such that *what it's like for it* to have its experience is none other than *what it's like for me* to have mine. And then: may I not in this fashion imagine that I have all the phenomenal features this entity has? And isn't this to say I have conceived of my phenomenal features as *purely* phenomenal, in isolation from my comportment? Well, perhaps I can conceive of many of my phenomenal features in this manner, purified of movement. But I have done this in a way that is *parasitic* on conceiving of my phenomenal life that is *not* free of a notion of myself as an embodied mover—a way that presupposes what we might now call a “comportmental” phenomenal concept: “looking at.” And we still have not analyzed *that* concept into a purely phenomenal and purely motor factors.

So the point remains that *looking at* things, understood *phenomenologically*, is not to be decomposed into a purely sensory or purely “mental” part, and a purely “behavioral, physical” part. Looking at things is understood as a *form of behavior*. However, this is not “behavior” as Descartes, Skinner or Quine would see it, but essentially, *experiential* (“lived”) behavior, Merleau-Pontyan *comportment*. Paraphrasing Merleau-Ponty, I might say: looking at things is, for me, *indissolubly both consciousness and movement*. Something similar emerges, I believe, when we shift from the case of *looking* to examine the case of *touch and feeling*. Much as more of what you see appears to you than what you're *looking at*, so also you *feel* more of what you *touch* than the area with which your body makes contact. But we have no conceptual, phenomenological grasp of this “what one feels” and its variation as one moves that thoroughly prises this apart from one's tactile engagement with one's surroundings—that detaches it from one's awareness of how one touches the world.

4. Does this Help with the Gap?

The activities with which we have just been concerned are *pervasive and fundamental aspects of cognition*. For we have been speaking here of that through which *object constancy in perception* arises. Or (to revert to Kantian language) we are concerned with that through which *objects are given to the understanding*. For it is by looking at things in such a way as to constitute, though looking, over time, coherent appearances of their shape, size and

position, that there are (visible) objects “for us.” Thus adequately developing these points about *looking* and *touching* could be consequential for our understanding of mind. A thorough phenomenological treatment of this, including an examination of the complex parallels and differences between vision and touch would be an involved affair. All I want to do here is to show in a preliminary fashion how we might go about filling the lacuna I find in Thompson’s argument, by dwelling on the phenomenology, while remaining true to the sources of his inspiration, and his general orientation. I also want to bring out how this offers an importantly different way of breaking the Cartesian spell that radically estranges mind and body than is to be found in the behaviorist-materialist tradition that has dominated Anglophone philosophy of mind.

But this last is not yet satisfactorily clear. To make it clearer, let’s return to what prompted these reflections about “looking”—the challenge of showing that the Thompsonian picture as characterized by (a)-(c) above really can be elaborated and defended so as to distinguish itself not only from options (1) and (2), but also from (3). We saw that there was a difficulty in making good the claim that a proper phenomenological understanding of perceptual experience finds it so bound to embodied activity that, in a non-trivial sense, the same perceptual functions that occur in our world would not occur in a putative “zombie world.” The worry was that phenomenologically there would be none *but* a trivial (because merely stipulative) sense in which the character of experience was necessarily bound to embodied activity. Any consciousness-involving form of motor activity we found in our own case, which could not be projected into the consciousness-free zombie world, could really just be seen as a *conjunction* of distinctly conceivable experience types and movement types, no more than contingently causally related. So the view on offer would provide no real alternative to (3). Once again we would ultimately come back to the proposal that we think of the mind-body relationship in terms of fundamental psychophysical principles joining consciousness (as phenomenologically conceived) to such physical properties as would be strictly guaranteed by the microstructure, and could obtain in a world without consciousness. But I hope to have given support to the prospect of a genuine alternative here, by suggesting a way to defend Thompson’s view. Understood phenomenologically, looking (and, I would argue, touching) involve movement whose unity of kind is conceived of by reference to the manner of experience it affords, which experience in turn cannot be conceived of entirely in detachment from that kind of movement. Basic phenomenal concepts of experience are *comportmental* phenomenal concepts. So, from a phenomenological perspective, the relation between experience and motor activity cannot be correctly represented as a contingent connection between pure experience and mere movement. In other words, a phenomenological understanding of the relation of experience and movement does not leave us with brute psychophysical contingencies, as does option (3)—for it does not divide the world into the purely phenomenal and the purely physical.

But now, even if all this is granted, it still is not exactly clear where this puts us with respect to the challenge of the explanatory gap. Perhaps we cannot explain why any world with the very same microstructure as ours must have consciousness in it. If that is required to close the explanatory gap, no prospect of closure has been offered. Nor is there any proposal in the offing as to why purely mental and purely physical domains should be linked in a particular manner, and not some other. But still we may discern some significant explanatory aim that requires neither of these. The basic goal should *instead* be to get a precise and general understanding of how an organism's embodied nervous system, in interaction with the environment, can generate the sort of movements that constitute the appearance to it of things in that environment—and thus how comportment comes to be.

But once again: how is this supposed to help? It might seem that, even if Embodied Consciousness is distinguishable from (3), it is really not all *that* different. Whereas Descartes and his descendants leave us with a brute contingency linking the domain of physical properties with that of mental properties, the phenomenologist leaves us with a brute contingency between the physical universe (what's fixed by fundamental physics) and the "life-world" of our comportment, in which we look at and touch things. In either case we find no evident *necessity* joining the universe of natural science with the world of experience. In the end, aren't we still simply left with this naked contingency, and a question about our intellectual comfort with this failure of necessitation?

Thompson is sensitive to this problem. It is what he addresses when he concedes that he might be taken to have simply replaced the old mind-body problem with a new "body-body" problem. (pp. 236-7) (For now instead of the problem of how to relate mind to body, we have the problem of how to relate the *experiential body* to the *physiological body*.) His response is that such a replacement is progress, since the second problem is in principle more tractable, as now there is at least clearly a "common term" viewed under different aspects to deal with. Let me try to say why it seems to me that the Thompsonian alternative does indeed give us an interestingly different approach to the issue of the explanatory gap. Perhaps then Thompson can explain how or whether he agrees.

Maybe this will help. (Or again, maybe Thompson can explain why it won't, or isn't needed.) First, we may note that the activity of the brain makes the organism move—in saying this, as yet we reveal no deep "in principle" gap, even if there are enormous theoretical challenges in understanding how it works. But now the question is, how should we conceptualize that *movement*, if we are to make sense of it—as we *should* if we want a realistic theory of cognition and behavior—as a creature's *interaction with an environment*? Well, in principle we could try to conceptualize it in the Cartesian/Behaviorist fashion as "bare movement." But we could also follow our natural inclinations to think of movement from the outset in terms (like *looking* and *touching*) that are, from a

phenomenological point of view, tied conceptually both to consciousness and to movement, *inseparably*. We could conceptualize movement as comportment. And if this second framework helps us find intelligible patterns in our own and other animals' interaction with their surroundings we have no reason to abandon it. (Anyway, it is doubtful that we really ever manage to sustain for long the intellectually induced autism of behaviorist methodology in a natural setting.) The burden of proof then would fall on someone who claimed we should attribute only some zombified analogue of looking and touching to ourselves and other animals.

But now, from this perspective, we do not have the usual problem of how *to put consciousness back* into a world of matter in motion once it has been *exiled* from our conception of that world. For the conception of motion that matters, for the purpose of making good biological sense of ourselves and other creatures in the actual world, is the phenomenological one, the comportmental one, from which consciousness is *not to be exiled in the first place*—since it is inextricable from that conception. If, nevertheless, there remains a sense in which the zombie world is conceivable (and even in a broad sense (metaphysically) possible), because we cannot conceptually derive comportment from microphysics, this should not trouble us. For whatever way of conceptualizing movement and mid-sized entities' interaction with their surroundings would be suitable to some world from which we conceive all the consciousness (sadly) drained away, this should serve as no guide for how we make sense of this, *our own world*—inhabited by billions of conscious beings, looking at and touching things. And if there remains a sense in which a *disembodied* consciousness is conceivable, this doesn't affect the fact that the basic conception of visual consciousness, in terms of which we understand ourselves, is one in which movement and consciousness are indissolubly fused, and that concept is the concept of comportment in terms of which we should couch our explanatory project.

Now it does seem that even if all this is embraced, it would not put to rest all worries that we really don't know just how widely to extent *comportment*. Do salamanders have comportment in a sense that incorporates consciousness? Do dung beetles? What about the artifacts that roboticists dream of constructing? Perhaps these and similar questions are containable epistemic problems that do not really threaten our prospects of explaining the comportmental consciousness to be found in "central," allegedly less problematic cases (to be found in humans, and most (all?) mammals). But this is a challenge that needs to be faced.

In any case, what's different (and, I think, refreshing) about the Thompsonian picture is that, while it can accept the modern idea that there is a way of conceiving of physical reality as a law-governed self-contained whole without thinking of it as inhabited by experiencing beings, unlike other approaches, it refuses then to see the challenge of explaining consciousness as that of how (either causally or epistemically) to *reintroduce* consciousness into

the universe thus conceived. For if we're trying to make sense of our interaction with our environment, the terms in which we do this should not be stripped of consciousness, as subjectively understood, *to begin with*. And the task of explaining consciousness is, most basically, the task of explaining that interaction so conceived. We find we can oppose the Cartesian estrangement of mind and body not (like a behaviorist) by spurning or belittling the reflective turn to consciousness, but by engaging in it more deeply. The Cartesian, we might say, turns to consciousness in a way that alienates us from our bodies, while the behaviorist (and many modern materialists) lock us into a way of thinking about our bodies that alienates us from consciousness. We can overcome both forms of alienation by discovering the unity of body and consciousness in sensorimotor life. What Descartes (and others) have put asunder we can rejoin—from *the preferred Cartesian standpoint*, that of consciousness, i.e., phenomenologically, “from the inside” if you like—which after all will be the only truly satisfying way to heal the split.

This little exhortation I hope is faithful to the spirit of Thompson's view. But perhaps I have strayed too far from this. Even if I have, then maybe setting me right will afford him a good opportunity to make his views yet clearer.

Perhaps I should say something about my fears of having strayed. In spite of my generally supportive response to Thompson's vision, I do have some misgivings that are perhaps not trivial, relating to aspects of his account that I have so far neglected, and which may ultimately show that my attempts to bolster the phenomenological side to the story are really not fully in line with his approach. A key aspect of Thompson's account has to do with how he finds a kind of “inwardness” constituted by a living thing's activity of self-organization—a kind of “immanent purposiveness” relative to which features of its environment have “valence” for it, and with respect to which its activity is normatively assessable. This kind of inwardness, he says, already goes beyond what can be captured in an “external” conception of “structure and function”—and is a “precursor” to phenomenal consciousness. (pp. 128-162, 222-5) One thing that is not clear to me is how this kind of “inwardness” is supposed to relate to the sort of subjectivity one might think integral to phenomenal consciousness itself (and not merely its “precursors”).

Thompson connects consciousness with the notion of *empathy*. (See pp. 162-5.) Perhaps the idea here is that phenomenal consciousness is that which by its very nature somehow demands a kind of empathetic understanding, an understanding “from the inside”—in the sense that forms of consciousness need to be understood by *sharing in them oneself*, if not actually, then in imagination. Real consciousness is present just where that kind of empathetic regard is called for. This seems to me on the right track—though of course the devil is in the details. But now I worry that there is some gap not fully acknowledged here that threatens to become sharper and wider as we are clearer about these two ways in which something may have “inwardness”—namely: (1) Being such as to

maintain (either relatively well or badly) a boundary between itself (inside) and others (outside); (2) Being such as to serve as a suitable focus of empathetic regard (something that can be understood “from the inside”). I get the idea that the first is somehow a necessary but not-quite-sufficient condition of the second, but the details are hazy for me. Does Thompson mean to show us just how (1) can be built up to get us to (2)? Or is *that* perhaps a gap he does not aim to close? Do we *need* to close *that* gap to explain consciousness?¹

NOTES

¹ I wish to thank Evan Thompson for discussions of his views, and Jeff Yoshimi for much helpful feedback on an early draft of this essay.