

« 5 » Higher-order phenomena are constituted (emerge?) from lower-order causal relations. A disruption in lower-order intralevel causality, however, will correlate with changes in interlevel relations. What is important, Kirchhoff argues, is that we understand this in dynamical terms and specifically in terms of diachronic constitution.

« 6 » This indeed is the enactivist view. Embodied mental processes (i.e., processes of the embodied-enactive mind), distributed across different factors (neural, behavioral, environmental) at different time-scales,² are constituted in a temporally integrated dynamical system. The constituent elements may very well be in complex, reciprocal causal relations with each other, but just these reciprocal causal relations make the mental process what it is. On the diachronic view of constitution, causality and constitution are not independent. Thus, an intervention that changes the causal relations in a dynamic system will also change the system as a whole. In a gestalt (what Maurice Merleau-Ponty (1964) called a “form” or “structure”), the whole is said to add up to more than the sum of its parts. In a dynamical gestalt composed of processes that unfold over time, and characterized by recursive reciprocal causality relations, changes in any processual part will lead to changes in the whole, and changes in the whole will imply changes in the processual parts. In contrast to a synchronic, compositional notion of constitution, these kinds of causal relations are constitutive of the phenomenon. As Kirchhoff argues, the notion of a

2 | Kirchhoff’s description of extended cognition equally applies to enactive cognition: “cases of EC are dynamically embedded at multiple timescales: e.g., at the time-scales of lower-level processes and their components – the temporal dynamics of neural processes, the temporal dynamics of bodily manipulation, and the temporal dynamics of cultural practices within which the overall distributed process unfolds” (Kirchhoff 2015b: 322). Indeed, as part of his neurophenomenological program, Varela (1999) identified just these specific timescales: (1) the *elementary* scale (varying between 10 and 100 milliseconds); (2) the *integration* scale (varying from 0.5 to 3 seconds); (3) the *narrative* scale involving memory (above 3 seconds). Correlations run across these scales.

causal-constitution fallacy, where constitution is defined synchronically, simply does not apply to the type of diachronic processes described in dynamic patterns.

« 7 » Is it metaphysical dualism to distinguish between lower-order and higher-order? No, if by lower order we mean simply the materially causal parts that dynamically and diachronically constitute the higher-order whole, where a change in the dynamical relations of parts involves a change in the whole, and vice versa. If the material causal relations in this type of system constitute experiential phenomena, is this best characterized as an identity? I suggest, following Merleau-Ponty, it is better characterized as a dynamically articulated form, structure, or gestalt.

Shaun Gallagher is the Lillian and Morrie Moss Professor of Excellence in Philosophy at the University of Memphis, and Professorial Fellow at the Faculty of Law, Humanities and the Arts, University of Wollongong (AU). He is also Honorary Professor of Philosophy at Durham University (UK), and Honorary Professor of Health Sciences at Tromsø University (Norway).

RECEIVED: 12 FEBRUARY 2016

ACCEPTED: 19 FEBRUARY 2016

Crossing the Explanatory Gap by Legwork, not by Fiat

Michael Beaton

University of the Basque Country,
Spain • mjsbeaton/at/gmail.com

> **Upshot** • I strongly agree with Kirchhoff and Hutto that consciousness and embodied action are one and the same, but I disagree when they say this identity cannot be fully explained and must simply be posited. Here I attempt to sketch the outlines of just such an explanation.

Introduction

« 1 » The target article by Michael Kirchhoff and Daniel Hutto is a helpful contribution to the metaphysics of consciousness. It seeks to avoid the dualism that, the authors correctly argue, permanently sustains “the hard problem” of consciousness in the

form given by David Chalmers (1996). In a non-standard but largely convincing move, the authors claim to find a similar dualism, leading to a similar hard problem, within the non-reductionist neurophenomenology of Francisco Varela (1996). Their proposed resolution of the hard problem is “to deny, resolutely” (§36) that the phenomenal and the physical (specifically, the realm of embodied action) are distinct. Whilst they accept that belief in such an identity can be motivated, they deny that the identity can be explained. Indeed, they suggest that such inexplicable identities are quite normal (§43). The authors of the target article agree with both Chalmers and Varela that there is “no prospect” (§10) of the success of an alternative approach, in which one would attempt to give a non-reductive *explanation* of how phenomenal consciousness relates to some relevant set of physical properties or processes. The only explanation that the authors of the target article do offer (although they explore this issue all too briefly both here, and in Chapter 8 of Hutto and Myin 2013, to which they refer) is the explanation of a certain “illusion” that, they claim, exists in our thought: the illusion that their proposed identity does not hold.

« 2 » My main aim in this response will be to explore the key option that Kirchhoff and Hutto (along with Chalmers and Varela) rule out. That of giving a revealing explanation of just how the phenomenal relates to the physical, without this explanation entailing that the former collapses into the latter. However, it should become apparent that I agree with Kirchhoff and Hutto on a lot in this area. Indeed, I will use exactly what those authors use on the physical side of the mind-world equation (or identity): namely, embodied action. But I will argue that with just a little more material (specifically, with a careful account of the nature of self-knowledge), we can offer a convincing explanation of the fact that the relevant aspects of the physical and the phenomenal are one and the same thing, rather than just having to accept an identity laid down by fiat.

Naturalisation, not reduction

« 3 » Kirchhoff and Hutto state that there is “no prospect” of success (§10; see also §39) for a standard form of scientific explanation of the place of consciousness in

nature, along the lines of famous scientific “reductions” such as the explanation of the properties of water in terms of the properties of H₂O molecules, or the explanation of heredity in terms of DNA. In part, it seems to me, this is because they share a widespread misconception of how such scientific explanations work. Indeed “reduction” is a misnomer here, in that the former concepts do not “collapse into” the latter, even in these standard cases. The concepts of “boiling,” “flowing,” “freezing,” “running in rivers,” “being what the sea is made of,” etc. do not “collapse into” the concept of “the bulk behaviour of H₂O molecules,” because nothing in the former concepts entails the latter. On the other hand, there is a one-way entailment in the other direction: once we know the properties of water and the microphysics of H₂O molecules, we can make a clear argument that H₂O molecules must (as a “conceptual necessity,”¹ if you will) behave in bulk as water behaves.²

« 4 » There is an additional point about any kind of broadly materialist naturalisation of consciousness that is also not clearly enough and often enough emphasised. This issue seems to me to hint very strongly at the correct approach towards removing the hard problem of consciousness without denying the reality of our own conscious ex-

perience. The point is this: if we are trying to naturalise consciousness, we need more than just a third-person description of consciousness (for instance, in terms of embodied action) that can non-mysteriously exist in the physical world; we also (and uniquely, of course) need to make it convincing that the macrophysical story that we choose to give is identical to own, lived, first-person experience.

Legwork, not fiat

« 5 » It is at this point that Kirchhoff and Hutto resort to fiat (i.e., stipulation, as if by law), insisting that embodied action simply is the same thing as first-person experience. In contrast, I want to suggest that if we identify consciousness with embodied action (and with the occurrent potential for such action), as the target authors and I do,³ then it is a mistake simply to “deny, from the off” that the two classes of phenomena are distinct (§36). Rather, I believe we can and should explain how and why first-person phenomenal experience is the same thing as embodied action.

« 6 » Somewhat ironically (it seems to me), Kirchhoff and Hutto actually begin such an explanation themselves, trying to give enough detail of the embodied action involved in handling a book to make it plausible that it must be like something to engage in such activity (§40). I myself have argued that there is a strong case for identifying the third-person structure of our

(actual and potential) embodied action with the first-person structure of our phenomenal experience (Beaton 2013). Indeed, such arguments are quite typical in action-based accounts of consciousness (e.g., Noë 2004; O’Regan, Myin & Noë 2006). For these reasons, it seems to me that you can and must put in the required legwork here: you can and should make a convincing *argument* that the structure of embodied action is the very same thing as the structure of phenomenal experience.

« 7 » Nevertheless, there is one further ingredient required to make such an argument compelling (or, indeed, to make any scientific or philosophical account of consciousness compelling), and that is an explicit and careful account of the nature of our first-person access to our own conscious states. This step is missing in most enactive and non-enactive philosophy of consciousness. I do understand that authors in the neurophenomenological and phenomenological traditions have said a lot about successful methods for accessing our own consciousness, but not so much about the nature of the first-personal access that underlies these methods. I have argued (Beaton 2009a) that such a philosophical account of self-knowledge or introspection must constrain our account of conscious properties, since, whatever these properties are, they are something that we can and do know from the first person.⁴ For myself, I endorse Shoemaker’s account of introspection (Shoemaker 1996).⁵ Furthermore, I have argued (Beaton 2009b) that it is possible to find personal, private (but only in a relatively modest sense of “private”) properties within the realm of embodied action that are close enough to “qualia” to count, and that can be known by Shoemakerian introspection.

« 8 » If we can make it convincing that certain properties of embodied action *are* the phenomenal properties with which we are directly, first-personally acquainted (as I believe we can, using this framework), then we have not just naturalised conscious-

1 | This is the so called “*a priori*” entailment to which Hutto and Myin (2013: 175) refer. I have argued (Beaton 2009a) that this philosopher’s term is misleading, for these explanations are not purely abstract and conceptual. Instead, they depend on the practical skills of engagement with the world that are as much required for knowledge of H₂O molecules as they are for knowledge of the day to day behaviour of water; and they further depend on the common sense ability (not formalisable, I have suggested) to see the self-evidence of the connection between the two realms of phenomena (once each realm is clearly understood).

2 | I simplify: 100% pure H₂O does not behave exactly as normal, impure water behaves (in terms of ease of super-cooling and super-heating, for instance), thus the properties of H₂O molecules alone are *not* sufficient to explain all the properties of normal, impure water; indeed, we still cannot predict all of the detailed bulk properties of water in complete, numerically exact detail, starting from the microphysics, and perhaps never will be able to.

3 | This approach is very different from that taken by most consciousness scientists. If pushed many consciousness scientists would accept that cognition is something done by whole animals (though many others might still locate cognition in the brain). Unfortunately, most consciousness scientists think consciousness itself is something that happens in some part or parts of the brain, under certain specified conditions. I disagree at the outset with this key assumption. For reasons given in my own target article in the current issue of this journal, I think that authors espousing action-based accounts of consciousness need not be tempted back inwards by variants of the Argument from Illusion; not even as regards the “minimal supervenience base” of consciousness, as Hutto and Myin (2013: 158–165) apparently are by their own variant, the “Argument from Shared Phenomenality.”

4 | For reasons that I explain in that paper, I do not believe that this rules out qualia in creatures that are too simple to introspect.

5 | Though I should clarify that I disagree with Shoemaker’s account of qualia.

ness by fiat, as Kirchhoff and Hutto seek to do. Instead, we have put in the required legwork. We have argued for a convincing, non-reductive naturalisation of consciousness.

« 9 » Importantly, I would disagree with the target authors that there need remain any persistent, albeit illusory, sense (§36) that consciousness as known from the first person is distinct from the relevant properties as known from the third person. When we think about the correct aspects of the macroscopic physical world – namely, engaged embodied action – and when we think about our own consciousness, and about the nature of our own knowledge of it, then far from there being a persistent illusion that first-person consciousness and third-person embodied action are different things, we can arrive at a compelling sense – achieved by convincing argumentation, not by stipulation – that these two are one and the same.

On the explanation of identities

« 10 » I note that the target authors wish to make their stipulative approach more plausible by pointing out other example identities that they think clearly cannot be explained. For what it is worth, I do not agree that their examples *are* inexplicable, in any relevant sense.⁶ For instance it seems to me that, *contra* Kirchhoff and Hutto (§43), we can perfectly well explain why George Orwell is Eric Blair: George Orwell is the pen name of the main born Eric Blair; to know of him by his birth name is one valid way of referring to him; to know of him by his much better-known pen name is another; someone could perfectly well know of him by either mode of referring without knowing the other; there are clear explanations to be found of why authors choose pen names, in general, and of why this author chose this pen name, in particular; and so on. We can even, I would suggest, again *contra* Kirchhoff and Hutto (§43), explain why 1 is 1:

6] I do agree that George Orwell is Eric Blair, with nothing more to be said about it, *if* all parties understand what identity means, know that George Orwell and Eric Blair pick out the same referent, and then insist on thinking entirely about that unique referent and not about the differing means of picking it out.

we will need to talk about modes of referring; about what identity means; and so on. I would argue that to engage with someone seriously asking such a question is already to have enough in common with them to begin this discussion.

The end of the story?

« 11 » The account I have outlined here puts norms (purposes, goals) on the physical (or more accurately, the third-personal) side of the equation. However, since Kirchhoff and Hutto put embodied action, rather than mere norm-free movement or “behaviour,” on the physical side of their identity, they do this too. For my part, I strongly suspect that there is difficult work still to do in showing how norms can be a part of the physical world.⁷ I equally accept that many enactivists think that the problem of the naturalisation of norms is already more or less solved by the concept of autopoiesis (Maturana and Varela, 1980) and its more recent theoretical extensions (Di Paolo 2009). Be that as it may, what I hope that the approach argued for above shows, at the least, is that there are not two separate problems: the problem of norms and the problem of consciousness. Rather, I think we can convincingly *explain* why these are one and the same problem.

Conclusion

« 12 » The target article by Kirchhoff and Hutto contains much of value, and I strongly agree with their claim that conscious experience and embodied action are one and the same thing, known in different ways. However, the central purpose of their target article is to make a metaphysical claim with which I disagree. Their claim is that the best way to naturalise consciousness is to “deny, resolutely” (§36) that consciousness and embodied action are distinct, with no further explanation of the identity between the two given or possible. They contrast this identity-based approach with non-reductive approaches (including Varela’s neurophenomenology), which they say (and I agree) necessarily leave an insoluble “hard problem” of consciousness. They also, more briefly, contrast their approach with more

7] Indeed, for reasons which I cannot begin to go into here I am tempted towards a form of pan-proto-psychism about normativity.

traditional attempts to explain the place of consciousness in nature, in the manner in which we explain heredity in terms of DNA, for example. These, they state, have “no prospect” of success (§10).

« 13 » In this commentary, I have attempted to provide the form of explanation that Kirchhoff and Hutto state cannot be provided: I have attempted to explain why consciousness is the same thing as embodied action. Uniquely, for a scientific “reduction” (such as water to H₂O, or genes to DNA molecules), we need two explanations here, not just one. As is normal, we need a convincing explanation of how the higher level phenomenon (in our case, embodied action) can be present in a physical universe. But, uniquely in the case of consciousness, we also need a convincing explanation as to why the high level story we have given describes *us*; our very own phenomenal experience.

« 14 » I have argued that the correct philosophical analysis of the nature of our access to our own mental states and processes (I endorse Sydney Shoemaker’s account) can allow us to see why and how the realm of our embodied action is the very same thing as the field of our own consciousness experience. I have not just stipulated this, I have tried to explain it (briefly, in this commentary, and in more detail elsewhere). Thus, while I agree with Kirchhoff and Hutto on many things, I oppose the central claim of their target article and I argue that we can and should cross the explanatory gap by legwork, not by fiat.

Acknowledgment

I gratefully acknowledge support by the Spanish Government MINECO Project, Reference FFI2014-52173-P.

Michael Beaton has moved, academically, from physics via artificial intelligence to philosophy. Having realised (or so he thinks) that traditional artificial intelligence was getting something wrong, his central philosophical interest has been to find a scientifically acceptable naturalisation of consciousness that does not “miss the point,” as more traditional computationalist and representationalist approaches seem to.

RECEIVED: 12 FEBRUARY 2016

ACCEPTED: 18 FEBRUARY 2016