Predicting the Body or Embodied Prediction?

New Directions in Embodied Predictive Processing

*Luke Kersten*

(word count: 4931)

In its broadest form, embodied cognition (EC) proposes that cognitive science is at a significant disadvantage when it fails to acknowledge the formative role played by the body in perception and cognition. When construed slightly more narrowly, it suggests that the real time, dynamic contributions of the body and its interactions constitute or deeply influence the character of cognitive and perceptual processes. For many, EC has proven a haven from the more reconstructionist and neurocentric tendencies in cognitive science (Varela et al. 1991; Chemero 2009; Shapiro 2011).

Recently, a new vision of cognition has begun to attract increasing attention within cognitive science. Emerging from converging work in cognitive and computational neuroscience, “predictive processing” (PP) suggests that the brain/mind does not have direct access to the world but, rather, must infer hidden causes from ongoing sensory input. In virtue of being inferentially secluded, cognitive systems, such as the brain, are constantly tied up in a process of trying to minimise the difference between incoming sensory inputs and prior expectations about the world, what is called ‘prediction error minimization’ (or PEM). For many, PP delivers a simple yet compelling story for a wide range of perceptual and cognitive processes and abilities, from vision and attention to consciousness and imagination (Clark 2013; Hohwy 2013).
Not surprisingly, several thinkers have begun to wonder how compatible these two influential approaches might be. Clark (2015, 2016, 2017), for instance, has suggested that PP and EC are not only compatible, but that they form a unifying vision of cognition. Hohwy (2016, 2018, 2019), on the other hand, has remained more sceptical, arguing that PP eschews stronger forms of embodiment in favour of a more inferential and representational picture of the mind. The task of sussing out the issue of compatibility is an important one. Not only have PP and EC been hailed as “paradigm shifts” and “revolutions” within cognitive science, but they have also stimulated a wealth of new research, including work on conceptual knowledge (Gallese and Lakoff 2005), emotion (Colombetti 2017), music cognition (Kersten and Wilson 2016) and consciousness (Deane 2021). Given its importance, this chapter looks to weigh in on the growing discussion surrounding PP-EC compatibility. I begin by providing a brief overview of recent discussion, and then turn to articulating two outstanding issues. I conclude by pointing in the direction of some fruitful lines of future development.

**Predicting the Body or Embodied Prediction?**

As mentioned, two authors have featured prominently in recent discussions of PP-EC compatibility: Andy Clark and Jakob Hohwy.1 Each has argued at considerable length for what they see as the correct vision of the relationship.

Hohwy’s (2016, 2018, 2019) view begins from the premise that the brain/central nervous system is constantly trying to provide its best ‘hypotheses’ for incoming sensory data. In line with the standard PP story, there are two general ways it can do this. One is to update its expectations about the world (change its predictive or generative model); the other is act on the world so as to bring it in line with its expectation (active inference). The former reduces the mismatch between prior expectations and incoming sensory signals via model updating, while the latter reduces mismatch via aligning the world with prior expectations. Both
strategies minimize the same sensory prediction error, albeit through different means. If Hohwy is right, then the brain is trying to constantly infer the shape and structure of a distal realm on the basis of partial and fragmentary information. The brain’s access to the world is limited by the flow of information from an individual’s sensory surfaces, whether that is visual, auditory or motor.

But importantly, PEM takes place behind what Hohwy calls an ‘evidentiary boundary’ (EE). The idea is that when some evidence becomes an indispensable part of the evidential basis for a hypothesis, an evidentiary boundary is formed, one which draws a circle around \( e \) and \( h \), on the one hand, and the hidden causes of \( e \), on the other. For example, following Hempel (1965), Hohwy offers the case of seeing footprints in the snow. While seeing footprints may give rise to the hypothesis that there is a potential burglar lurking nearby, the strength of the footprints acting as evidence for the burglar hypothesis rests on how well the burglar hypothesis, in turn, explains the evidence.

The evidentiary boundary is important Hohwy thinks because it can be used to demarcate the boundaries of cognition – that is, whether or not something qualifies as part of a cognitive process or state depends on what side of the evidentiary boundary it finds itself. This spells trouble for strong versions of EC (i.e., ones that view the body as constitutive of cognition), because it looks like only sensory states of the nervous system are appropriately placed to stand in for evidence. As Hohwy explains: “[W]hat is actually part of the EE-circle determined by our prediction error minimizing systems? The answer to this question is that it is brains and sensory states (i.e., the states of the sensory organs) that form the EE-circle. There is no good reason to include anything else in the EE-circle” (2016, p. 269). The trouble is that while we regularly interact with bodies such that we form hypotheses or models about them via predicting the sensory inputs they cause, we do not do so with our sensory organs. Given this, it seems that features of the body fall outside of the evidentiary boundary.
The implication is that EC can be accommodated by PP but only from within the strictures of the self-evidencing brain. As Hohwy (2016) makes the point:

[T]here is no fundamental difference between types of inference that rely on the body and types that don’t: they all consist just of inner representations of interacting, hidden causes and generation of expected sensory inference on the basis of these representations. Embodied cognition boils down to the fact that one of these modeled causes is the agent’s own body (p. 275).

It is inference and inference alone which secures evidence for the hypotheses the brain uses to guide action. It is inference which puts the brain in a position to match sensory inputs with expected outputs in the first place; a process which is wholly internal, and brain bound. So, while it is true to say that the body plays a role in cognition for Hohwy, it operates only as a useful parameter for PEM within the internal generative model of the brain.

In contrast, Clark (2015, 2016, 2017) envisages a more constructive relationship between PP and EC. Recall that on Clark’s view EC and PP are not only compatible, they form a unifying picture of cognition. Clark offers both negative and positive arguments for this position, though for reasons of simplicity and space I focus largely on the presentation provided in Clark (2017).

On the negative side, Clark attempts to show there is no direct link between PP and neurocentric visions of the mind. He suggests, for instance, that there is an ambiguity in how some authors use the concept of ‘inference’ within PP. One way to understand inference he thinks, in line with Hohwy’s take, is as a reconstructive process. On this interpretation, the brain is trying to piece together an internal model of the world based on impoverished incoming sensory data. This, Clark suggests, is what gives rise to the impression that the brain is secluded from the world, one has to infer the world from limited sensory input. However, another way
to interpret the concept is as a process that enables different strategies for co-ordinating behaviour. On this non-reconstructive interpretation, long term prediction error minimization is actually an action-involving process. The reason for this is that it is often simpler and more efficient to enact certain behaviours and shape the sensory flow than it is to try and create costly internal representations. As Clark (2017) puts it: “The task of PP systems is not to infer the best description of the world given the sensory evidence. The fundamental task, using prediction errors as the lever, is to find the neuronal activity patterns that most successfully accommodate (in action, and in readiness for action) current sensory states” (p. 734). For Clark, talk of inferring the right ‘hypotheses’ is unnecessarily reconstructivist. While initially couched in neurocentric terms, concepts such as inference are not inherently biased against EC. An equally viable interpretation, one compatible with the PP story, is that inference involves a way of structuring on-going interactions with the world.2

This idea of structuring on-going, fluid interactions with the world in the service of PEM leads Clark to a more positive proposal. Because agents engage in iterative, dynamic perception-action cycles, they often offload computational work onto their environment, which allows them to minimise prediction error more cost effectively. One illustrative example comes from the outfielder problem. The outfielder problem asks how a baseball player (an outfielder) is able to catch a fly ball while on the run. One suggestion, in line with the reconstructivist view, is that agents engage in form of internal replica building via inference, estimating and tracking the ball’s trajectory using rich, internal representations. However, an alternative, more cost-efficient method, is that agents keep the image of the ball stationary on their retina in order to keep the flow of sensory information within a certain range. On this second approach, behavioural success is achieved not by reconstructing costly inner replicas but simply by maintaining invariant relations between the agent and the world. For Clark, the key insight is that agents often use low-cost, action-driven strategies to minimise prediction error; they
engage in a form of ‘productive laziness’. In so doing, they create temporary neural-body-world ensembles to solve context-dependent problems, such as catching the fly ball. These temporary problem-solving ensembles emerge and dissolve out of local necessity and are driven by action-perception cycles.

There are two important ideas here. The first is that in order to sculpt the flow of sensory information for real time, on-going use, agents must often engage in bouts of action-perception cycles (see, e.g., Clark 2015). These are circular, casual interactions between the organism and its environment, ones which the computational workload from the brain to the non-neural body and world, what is sometimes called ‘cognitive scaffolding’. This process creates a ‘motor-informational weave’, as Clark calls it, which is necessary to sustain the creation of ‘transient extended cognitive systems’ (or TECs) (Clark 2008; Clark 2016, Clark 2017). The idea is that agents can sometimes get by with purely internal neural assemblies, and sometimes they need to call on external non-neuronal resources, such as the body, to form larger problem-solving wholes.

The second is that many of the same basic rules and principles that govern inner neural coalitions, such as efficacy and efficiency, also govern TECs – Clark points to ‘Optical Acceleration Cancellation’ (the strategy of keeping the ball’s image on the retina) as an example of how systems trade off efficacy and efficiency. There is no real difference, according to Clark, in how coalitions of internal neural assemblies and body-spanning ones select actions to solve specific problems. Crucial to both is the ability to make task-specific information available for fast, fluid use. This means that PP actually fleshes out the computational story behind why TECs are formed: they are one strategy amongst others for minimizing prediction error.

The takeaway is that PP is not only opposed to more “reconstructive” interpretations, but it also accommodates key ideas from EC, such as cognitive scaffolding and productive laziness.
This more “radical” reading offers a systematic way of combining deep, model-based flexibility of PP with the frugal, environmentally exploitative actions of EC.

**Two Outstanding Issues**

There is much to be said for Hohwy and Clark’s proposals. Both authors offer instructive and nuanced discussions of various aspects of the PP-EC relationship, such as the relationship between active inference and perception-action. However, despite their informative character, there are still several outstanding issues that need to be addressed if discussion of PP-EC compatibility is to progress.³

One issue concerns the conceptual status of PP and EC. The trouble is that it is often unclear what type of unit of analysis PP and EC are supposed to denote within philosophic and scientific theorising. For instance, while some have suggested that PP is a unifying “theory” (Hohwy 2013, 2020; Litwin and Milkowski 2021), others have contended it is a research “framework” or “paradigm” (Sprevak 2021; Michel 2022). Similarly, while some have suggested that EC is a “thesis” or “hypothesis” (Meteyard et al. 2012; Wilson & Golonka 2013; Mahon 2015), others have claimed it is a “research programme” or “research tradition” (Shapiro 2007; Shapiro and Spaulding 2021; Milkowski and Nowakowski 2021). A wide, and largely distinct, array of terms has been applied to discussions of PP and EC, and not always in consistent ways.

This “status” problem is important because different theoretical units carry with them different standards of comparison and modes of relation. For example, while behaviourism and cognitivism are arguably incompatible when cast at the level of theory, each offering rival attempts to explain specific phenomenon such as language acquisition, the same cannot be said when viewed at the level of research programme. While theories can be in direct tension with one another, it is less clear the same is true of theories and research programmes (Moore 1997).
Because different theoretical units can vary with respect to level of abstractness and function, the character of compatibility between PP and EC can change depending on the type of conceptual unit being compared. If PP denotes, for instance, a “theory-like” structure while EC picks out a “research programme-like” structure, then the question of compatibility may arise in a different form than it would if they referred to the same structure. Without first getting clear on the theoretical status of PP and EC, any discussion of compatibility may prove premature.

A second issue concerns the theoretical commitments of PP and EC. The trouble this time is that it is unclear which theoretical commitments are supposed to be central to PP and EC. For instance, in various places Clark (2015, 2017) points to “cognitive scaffolding” and “productive laziness” as key elements of EC, whereas Hohwy (2018) singles out “tight agent-environment coupling” and “fast and frugal processing”. A similar point holds in the case of PP. For instance, Michel (2022) suggests that the central tenet of PP is that the “mind entertains a probabilistic, hierarchical generative model”, while Sprevak (2021) rejects such a characterisation in favour of a cluster of claims about the computational, algorithmic and implementational details of cognition (see Sprevak (2021) for details). There is little agreement, even at a very general level, about the core ontological and methodological commitments of PP and EC. As Milkowski (2019) ably puts the point in the context of EC: “While most surveys, defences, and critiques of embodied cognition proceed by treating it as a neatly delineated claim, such an approach soon becomes problematic due to the inherent plurality of this perspective on cognition” (p. 221).

This “commitment” problem is again important because a lack of consensus could spell trouble for the scope of any account of compatibility. If the commitments are spelled out too narrowly, then compatibility might be achieved but only at the cost of sacrificing wider insight. For example, if, as Clark suggests, PP and EC are compatible with respect to cognitive
scaffolding, but not, as Hohwy suggests, constitutive embodiment, then Clark’s account, while illuminating, would only have limited scope. It would only demonstrate compatibility for specific formulations of EC, such as those pertaining to cognitive scaffolding, but it would remain silent on how PP relates to other (potential) core commitments of EC. If we are not clear about which commitments are central to EC and PP, then a given account could exclude other relevant commitments. Conversely, if the commitments are defined too broadly, then compatibility could become trivial. For example, if, as is sometimes suggested, EC is simply the commitment to a “crucial role for the body” (Meteyard et al. 2012) and PP is simply the idea that brain is a “prediction machine” (Venter 2021), then the two views may be compatible but this compatibility is not particularly informative. What we need is a way of making sense of the diversity of commitments found within PP and EC, but such that it does not sacrifice the informativeness of the account of compatibility.

So, to summarise, there are two outstanding issues impeding progress on thinking about PP-EC compatibility. The first concerns the theoretical status of PP and EC, while the second concerns how to specify the theoretical commitments of PP and EC. To be clear, I raise these issues not as specific criticisms of existing proposals but, rather, as invitations for further development.

**New Directions in Embodied Predictive Processing**

In this final section, I want to say a bit about how to tackle the two outstanding issues. The proposal, in short, is that discussions of compatibility will benefit from more explicit engagement with the resources of philosophy of science.

To begin, consider that there is a general distinction within philosophy of science between two types of units of analysis or scientific representations (Lakatos 1970; Laudan 1977). On the one hand, there are the more detailed, concrete units, ones which attempt to explain and
predict specific phenomena, such as optical occlusion in visual perception or the phenology of hibernation. These units usually march under the banner of terms like ‘theory’ or ‘model’, with examples including Maxwell’s theory of electromagnetism or Baddeley’s model of working memory. On the other hand, there are broader, more abstract units, ones which operate above the level of individual theory or model and often guide and constrain theory or model formation; examples include Hutton’s uniformitarianist theory in geology or Gibson’s ecological approach in psychology. These units are sometimes referred to as ‘research programmes’, ‘paradigms’, or ‘frameworks’.

Now the question of how these two types of units or representations relate is a subtle and complex one. Fortunately, for present purposes, we can sidestep this issue and focus instead on how PP and EC have been more generally characterized in discussions. This is interesting because many authors see PP and EC as falling into either one or the other category. For example, Shapiro (2007) maintains that: “the point of labelling EC a research programme, rather than a theory, is to indicate that the commitments and subject matters of EC remain fairly nebulous” (my emphasis). In a similar vein, Sprevak (2021) suggests of PP that: “it is more accurate to think of predictive coding [PP] as a research programme rather than as a mature theory (my emphasis).” For both authors, PP and EC are best thought of as either concrete, testable units or abstract, research guiding units, but not both. What I want to suggest, though, is that we should also favour the abstract interpretation of PP and EC, and that interesting possibilities begin to open up when we do so. There are two main reasons for this.

The first is the characteristic function of PP and EC within scientific practice and theorizing. In a broad sense, research programmes, paradigms, and frameworks all aim to motivate and guide theory or model development. They do so by providing high level schematic proposals for how to implement and interpret various principles or concepts. This is in keeping with how PP and EC generally function within scientific and philosophic discourse (see, e.g., Shapiro
2007, 2011; Milkowski 2019; Sprevak 2021; Michel 2022). For example, while many descriptions of PP express a commitment to the idea that precision weighting is important for optimising prediction error minimization (e.g., Clark, 2013; Hohwy, 2013), the question of how precision weighting is actually implemented in the dynamics of neurotransmitters is often left unspecified.\footnote{This makes sense if PP operates as a broader, more abstract unit of analysis. The implementational details are left open because PP is not trying to provide specific, testable claims. Rather, PP aims to guide and constrain specific theory and model construction through the articulation of general principles and concepts. The same is also true of EC. It is better to talk of PP and EC as unifying frameworks or research programmes, rather than unifying theories.}

The second reason is that PP and EC possess a wide array of ontological and methodological commitments. For example, as mentioned, PP regularly employs talk of generative models, probabilistic inference, efficient neural coding, prediction error minimization, and top-down effects; while EC invokes talk of sensorimotor contingencies, action-perception cycles, exploitative representations, and temporally extended dynamic systems. The easiest way to accommodate this range of concepts and principles is to conceptualise them as organising or guiding commitments for research. Research programmes, paradigms, and frameworks all specify, in very general terms, the basic types of fundamental entities that exist within a domain, and the methods used for conducting inquiry. Or, to frame the point in more Marrian terms, PP and EC provide sets of concepts and principles that guide and constrain further algorithmic and implementational level accounts. While individual theories and models also exhibit various ontological and methodological commitments, it is the diversity and generality of the commitments found in PP and EC that speak in favour of their interpretation as more abstract units of analysis.
If it is right to say that PP and EC are best understood as broader, more abstract units, then this opens up interesting possibilities when it comes to addressing the theoretical commitment issue. This is because there is already a well-established literature dealing with the character and function of abstract units within philosophy of science (see, e.g., Lakatos, 1970; Laudan, 1977). One point often made in these discussions is that the same general commitments can be implemented in a variety of ways with particular theories or models – for example, to develop a theory of the motion of a compass when near a current carrying wire a researcher must go beyond the general commitments of Newtonian physics, such as the claim that all non-rectilinear motions should be treated as cases of centrally directed forces.

This provides one potential explanation for why it has proven so difficult to clarify the commitments of PP and EC: one cannot read off the commitments of abstract units from their associated theories or models, nor vice-versa. If one were to survey embodied theories of semantics, for instance, one would be forgiven for thinking that a notion of “representation” is common to EC; talk of the sensory and motor information in cognitive representations is a constant theme in discussions of embodied semantics (see, e.g., Lakoff 2012; Dove 2022). However, as is well known, a number of distinct EC theories of vision, such as Noë (2009), explicitly eschew talk of representations in favour of explanations involving the action-dynamics of an agent. Because different theories or models have different ways of implementing the more general commitments of abstract units, attempting to infer the core commitments from a survey of associated theories or models is not only unlikely to succeed but also potentially misleading. One implication is that researchers should avoid making inferences about core commitments solely on the basis of an abstract unit’s associated theories or models. Instead, they must provide a rich, detailed study of the unit’s history.

Whether or not this particular point about the relationship between abstract and concrete units is correct is less important than the example it provides. For it shows how the resources
of philosophy of science can fruitfully be brought to bear in discussions of compatibility. The toolkit provided by philosophy of science better positions researchers to identify the core commitments of PP and EC, and thereby make progress on the question of compatibility.

So, taking stock, this chapter has attempted to weigh in on the growing discussion surrounding PP-EC compatibility. It did so by, first, surveying two prominent proposals from Clark and Hohwy and, second, articulating two outstanding issues. The positive proposal was that both outstanding issues could be fruitfully addressed by drawing on the resources of philosophy of science. The significance of the analysis, as I see it, lies not only in the specific proposals it makes, but in the general direction it sets for discussion. It raises a number of interesting further questions. Questions such as: What specific types of units of analysis are PP and EC? Are they research programmes, research traditions, or frameworks? Are their important differences depending on which we choose? Is there particular model of science that we should adopt when thinking about compatibility? Lakatos (1970) or Laudan (1977)? And under what conditions is integration possible? The answers to some (or all) of these questions could notably reshape the form of compatibility the PP-EC relationship takes going forward.

Acknowledgements

Some passages in this chapter have been adapted with revision from Kersten (2022, 2023).

Notes

1. For other interesting discussions of PP-EC compatibility, see also Anderson (2017), Kirchhoff (2018), and Venter (2021).

2. While I cannot go into detail for reasons of space, it is worth noting that Hohwy (2019) does respond to Clark’s (2017) critique, arguing in turn that Clark’s embodied-friendly
interpretation of PEM as “accommodation” is equally problematic, and that only an inferential interpretation coheres with PP.

3. To qualify slightly, while I am placing scientific units broadly into two categories, I do not think these are the only categories possibility for scientific units. I acknowledge, for example, that there is an important difference between ‘models’ and ‘theories’. Nevertheless, I think this omission is largely justified at present because the taxonomy marks at least one important distinction with respect to function and abstractness of different units.

4. For a more detailed discussion, see Kersten (2022, 2023).

5. Precision weighting describes the ‘gain’ as it applied to components of the prediction error. Errors that have a high precision weighting are prioritised during a task; errors that have a low precision weighting are given a lower priority or even partly discounted.

Notes on Contributors

Luke Kersten is currently a Killam Postdoctoral Fellow in the Department of Philosophy at the University of Alberta. His research focuses primarily on issues in the philosophy of mind and cognition. He received his PhD in Philosophy from the University of Edinburgh.

References


