The Predictive Mind. Jakob Hohwy. Oxford University Press, 2013, 286pp, Hardcover, £65, ISBN 9780199682737.

In his book, The Predictive Mind, Jakob Hohwy builds a case for the increasingly popular idea that perception is a matter of prediction. You do not passively perceive the world around you. Instead you predict what you are going to see. For Hohwy, this translates into the brain having internal models or representations about the world, which the brain updates by generating predictions about the world and then modifying those predictions in the light of future input.

The unique selling point of this view of the brain, insists Hohwy, is both its simplicity (the brain is a hypothesis-tester, nothing more, nothing less) and its explanatory reach (it explains perception and action and everything in between). Yet while it is fairly uncontroversial to claim that perception can involve prediction - most are likely to agree that your previous expectations can influence what you currently see - it is much more controversial to claim, as Hohwy seemingly does, that minimizing prediction error is all the brain ever does or that this is an expression of how creatures like us self-organise, which is in turn to be understood via the information theoretic notion of minimizing surprisal. Hohwy's aim with his book is to convince us why we should take these further claims seriously.

Hohwy divides his book into three parts. In part 1, he details the sort of prediction error minimization (PEM) mechanism at the heart of his account of the brain. In part 2, he describes how this mechanism helps advance debates within cognitive science, such as that of the binding problem, cognitive penetrability and cognitive impenetrability, misperception and representation. In Part 3, Hohwy examines what PEM means for attention, perceptual unity and the fragility of perception. He also applies PEM to other aspects of mentality, like emotions, introspection, the privacy of consciousness, agency and the self.

Hohwy's book is a rich resource, ranging over many and varied issues. In this review, I will not attempt to cover all these issues. Instead, I will zero in on what Hohwy calls the "problem of perception" and how he thinks his version of PEM can solve this problem. Once done, I will detail some concerns I have with Hohwy's proposed solution and suggest how these concerns might impact on our understanding of the problem.

1: The problem of perception

Picture your brain tucked away inside your skull. Now consider that your brain is constantly on the receiving end of sensory input from the outside world (and produces output as a result). Given the brain's isolation inside the head, how does the brain make sense of all this input? More importantly, how does the brain *get things right*? This is the "problem of perception". Hohwy states:

> "The problem of perception is how the right hypothesis about the world is shaped and selected [by the brain]." (p15)

Solving this problem thus requires explaining, not only how the brain helps us perceive, but also how the brain helps us perceive in the *right* way.

Hohwy's solution is to argue that the brain is a Bayesian mechanism. Very briefly, Bayes' rule is a way of calculating the probability of a hypothesis given the current evidence. The rule does this by comparing: (1) the likelihood of a hypothesis given the current evidence and (2) the prior probability of that hypothesis, that is, how probable that hypothesis is independent of the current evidence. The result is the posterior probability for a given hypothesis.

According to Hohwy, the brain is a mechanism that realizes Bayes' rule. So, for example, the brain determines the posterior probability of a hypothesis through the use of a perceptual hierarchy. At the top of this hierarchy, the brain's internal prior constraints generate predictions, which then flow down through the hierarchy. At the bottom of the hierarchy, the brain receives sensory input, which in turn generates a prediction error signal, which is then sent back up through the hierarchy. This top-down, bottom-up process is then reiterated, such that the prediction error signal generated by the sensory input is minimized and hence the discrepancy between the brain's internal prediction and the received sensory input is reduced.

This ensures that the world acts as a supervisory signal for the hierarchy, enabling the brain to acquire feedback for its predictions and so allowing its own internal models or representations of the world to become better or more accurate models or representations of that world. The prediction error signal is thus understood as an “objective corrective caused in us by objects in the world” (p45).

Now, the world is often a highly uncertain and noisy place. It follows then that if the brain is to accurately model or represent the world, then it needs a way to accommodate this uncertainty and noise. It does so, claims Hohwy, by factoring in the *precision* of the prediction error signal. If the precision of the prediction error signal is taken to be low i.e. there is an expectation of a lot of noise, then less weight is accorded to that signal. By contrast, if the precision of the prediction error signal is taken to be high i.e. there is an expectation of little noise, then the prediction error signal is given greater weight. By such means the brain recapitulates the causal structure of the world and does so in a context sensitive manner.

Within this hierarchical structure, there is a complex relationship between the different levels, such that when one level accords less weight to the prediction error signal, then that level or the level below will do more of the revision of that error signal. Alternatively, when a level accords greater weight to the prediction error signal, then higher levels will do more of the revision of that error signal. There is thus a delicate balancing act to be struck between each level.

Further, the prediction error signal can be minimized in more than one way. For example, the brain’s models or representations of the world can be updated by changing the predictions generated by the brain’s internal priors in such a way as to match the sensory input. This is termed “perceptual inference” and has a mind-to-world direction of fit, that is, the mind (the prediction) is adjusted to match the input from the world. However, these models or representations can also be updated by changing the input to match the predictions generated by those priors. This is termed “active inference” and has a world-to-mind direction of fit, that is, the world or the sensory input is adjusted, through action or active sampling of the world, to match the predictions generated by the brain’s internal priors.

Hohwy insists that the attractive feature of all this is that both perception and action can now be explained in terms of minimizing the prediction error signal (p88). Another allegedly attractive feature is that this understanding of the brain offers ways to tackle long-standing debates within cognitive science, such as that over the binding problem.

Suppose, for example, I see a red ball bouncing towards me. My visual percepts are seemingly ‘unbound’ (something red, something bouncing, something ball shaped etc). Yet if so, then why do I see a ‘bound’ visual percept, that is, a ‘red ball bouncing’? That is, how does my brain bind together the various non-bound percepts that arrive through my senses? This is known as the binding problem.

Hohwy’s solution is to invert this problem (p101). He claims it is not the case that the brain must somehow bind together what is otherwise an unbound sensory signal. Instead, the brain predicts that the sensory signal is bound and then looks for confirmation of this prediction in input. In other words, the brain assumes bound attributes and then queries the sensory input on the basis of this assumption. Binding is thus “essentially a statistically based inner fantasy, shaped by the brain’s generative model of the world” (p115).

However, if the variability of the input signal is sufficiently high, then this can result in illusory binding. According to Hohwy, this is what happens in the rubber hand illusion. This illusion occurs when an experimenter strokes a subject’s unseen hand, while simultaneously stroking a rubber hand that the subject is looking at. After a suitable period of time, subjects begin to report seeing the rubber hand as their own hand. Hohwy explains this illusion by saying that the brain assumes a common hidden cause when there is a strong spatiotemporal overlap in our sensations (p105-106). Yet due to the variability in the sensory input, predicting that the rubber hand is the subject’s own hand best explains away that input, that is, reduces prediction error.

Consideration of illusory binding might however lead one to wonder why Hohwy’s account of PEM doesn’t simply result in self-fulfilling prophecies. For if the brain is constantly seeking to minimize the prediction error signal, then why don’t we just spend our lives living in, say, darkened rooms, since doing so would enable us to predict perfect darkness and so effectively minimize any residual prediction error?

By way of reply, Hohwy acknowledges that the brain does engage in self-fulfilling prophecies but this is not a problem because underlying PEM is a drive to minimize surprisal, that is, to keep the organism within a certain range of expected states. The greater the discrepancy between the predictions generated and the error signal propagated up through the hierarchy, the greater the surprisal. Conversely, the less the discrepancy between the predictions and the error signal, the less the surprisal. Hohwy argues that this is linked to the phenotype of the organism. For the phenotype of the organism will ensure that certain states will be expected, that is, will have less surprisal.

In his book, Hohwy provides no further details about what exactly those expected states are. Nonetheless, the idea here seems to be that we don't spend our lives, say, living in darkened rooms, since, unlike other kinds of organisms, doing so would not keep us in the sorts of expected states that we can reasonably take to define us as the organisms we are. I shall leave it to others to decide how satisfactory an idea this is. Still, it reveals the central role that minimizing surprisal plays within Hohwy's account. Indeed, Hohwy goes so far as to remark:

> "the only reason for minimizing the error through perception and action is that this implicitly minimizes surprisal." (p87)

So, to recap: the "problem of perception" was to explain, not only how the brain helps us perceive, but also how the brain helps us perceive in the *right* way. For Hohwy, the mechanism that solves this problem is prediction error minimization (PEM). PEM is what enables creatures like us to hone and refine our internal models or representations about the world such that we can then acquire *right* hypotheses about the world.

2: Does PEM get things right?

I have outlined Hohwy's proposed solution to the problem of perception. I turn now to some concerns I have with Hohwy's solution and how I think these concerns might impact on the problem itself.

At one point in his book, Hohwy appeals to a probability distribution, that is, to the claim that when a number of visual percepts are averaged across time, this provides a standard against which future percepts can be judged 'true' or 'false'. Hohwy then uses this probability distribution to explain misperception. To borrow one of Hohwy's examples, I see a sheep but because of funny lighting I think I see a dog. Why is my percept of a dog false? A probability distribution ensures that my visual percepts of sheep will likely carry more information about sheep than about dogs. Hence, I can expect to see sheep-as-sheep and not sheep-as-dogs. If I then see a sheep-as-a-dog, then my percept is false because it misaligns with that probability distribution, that is, it "pair[s] percepts with environmental causes that are not on average best paired with" (p174). Misperceptions are thus "inferences that undermine average, long-term prediction error minimization" (p176). Note that, as we saw with Hohwy's account of binding, this then ties the brain's models or representations of the world directly to statistical physics (p180).

All of this raises a serious worry, however. Consider that most philosophers now accept that not all information is created equal. For example, some worldly states of affairs can co-vary with other worldly states of affairs such that the co-variances between such states of affairs can be used to reveal information about those states of affairs (think of how the rings of a tree can be used to determine the age of a tree). This is called "informational covariance". However, other worldly states, like my stated desire to go to Greece on holiday, can have properties like true or false (I may in fact desire to go to Turkey on holiday and so my stated desire is false). This is called "informational content". Crucially, properties like true or false ensure that informational content is logically distinct from informational covariance (Hutto and Myin, 2013).

The distinction between covariance and content is not controversial. What is controversial however is determining if or how covariance can constitute content. This controversy impacts on Hohwy's account of misperception. For a probability distribution simply is a set of covariances between, on the one hand, states of the PEM hierarchy in the brain, and on the other, external worldly states as revealed through sensory input. Yet if content is logically distinct from covariance, then it is not clear how Hohwy's proposed probability distribution can ensure that, say, my visual percept of a sheep-as-a-dog has the contentful property of being false. Or, to put the same point another way, if my percept is to be false, then Hohwy needs to show how the sorts of covariances as displayed by a probability distribution can in fact rise to the level of content. This will be no easy task (indeed, Hutto and Myin label this the "Hard Problem of Content"). Importantly, Hohwy does not tackle it in his book. Yet until he does so, all his account succeeds in showing is that my percept of a sheep-as-a-dog is statistically unusual, since statistically I can expect to see sheep-as-sheep and not sheep-as-dogs. It does not also show that my percept has the property of being false.

Now, you might think: so much the worse for Hohwy's account of misperception. But the problem identified here generalizes. For a related issue emerges when Hohwy claims that the brain "mirrors" the world (p228). This mind-world relation, insists Hohwy, is both direct and indirect. It is direct in the sense that the brain's models or representations are, as we have seen, honed and refined by the prediction error signal. Yet it is also indirect in the sense that PEM is always a fragile, non-robust process. This is because the sorts of fine-tuning needed to maintain a balance between perceptual inference and active inference can easily break down (or be broken down, as in settings like the rubber hand experiment).

However, mirrors simply co-vary with what they mirror. If covariance and content are logically distinct, then it is not obvious that mirrors do in fact have any contentful properties. As such, even if Hohwy is right to claim that the brain is "like a mirror of the causal structure of the world" (p228), this still leaves entirely unexplained how or why the brain is also a "truth-tracker" (p229).

On the other hand, Hohwy seems to simply assume that the brain minimizes prediction error and as a consequence gains truth. But one could endorse the idea that (1) the brain is a mechanism for the manipulation of causal regularities as revealed through sensory input, while nonetheless not endorsing the further idea that (2) there is any informational content inside the brain. [1] Indeed, given the sorts of heavy lifting that would need to be done to show that a probability distribution can have any contentful properties, then there may be good reason not to endorse (2).

Hohwy might respond to all this by claiming that we could not successfully engage with our environment unless our senses told us something true about the world around us. And this successful engagement can be cashed out in terms of maintaining integrity and avoiding entropy. Yet this just pushes the problem back a stage. For what exactly is the link between minimizing surprisal and an organism having internal states that are true or false, right or wrong? It seems whichever way you cut it, this question encounters a Hard Problem, one that needs to be dealt with before Hohwy can claim that PEM solves the "problem of perception". We can still ask: how does minimizing prediction error enable the brain to *get things right*?

However, perhaps what is really at fault here is the problem itself. Recall how we initially set up the problem: we have a picture of an isolated brain within a skull trying to make sense of raw sensory input. Maybe what is problematic is this picture. If so, then rather than trying to make sense of this picture, we might get further by asking ourselves, not, *how does the brain get things right*, but instead, *how do we get things right*? And answering that question will require extending our explanatory focus way beyond the narrow confines of the skull.

As stated earlier, all of this is to skip over plenty of the detail within Hohwy's book. And it would remiss of me if I ended without pointing out that Hohwy does successfully demonstrate the sheer scope of PEM, that is, how PEM does seem to have the resources to tackle lots of the issues currently confronting cognitive science. Nonetheless, an important lesson Hohwy's version of PEM might show is that maybe

---

[1] Clark (2015) seems to come close to endorsing this view. For example, he claims that states of the brain are action-oriented. That is, the brain's internal model is geared towards "delivering a grip on the *patterns that matter* for the *interactions that matter*" (p5, italics in original) As a result, even high-level states of a PEM mechanism do not describe the world, that is, have anything like recognizably contentful properties. Yet Clark also sees this as entirely compatible with contentful talk. However, how or why this should be so is, in my opinion, much less clear.

not all the answers to the questions that interest us are going to be found by looking inside our heads. [2]

Bibliography

Clark. A. (2013). Whatever next? Predictive brains, situated agents and the future of cognitive science. *Behavioural and Brain Sciences*, 36 (3), 181-204.

Clark. A. (2015). Predicting Peace: The End of the Representation Wars. A reply to Michael Madray. In T.Metzinger & J.M Windt (Eds). *Open Mind: 7 (R). Frankfurt am Main: MIND group.*

Hohwy, J. (2013). *The Predictive Mind*. Oxford University Press.

Hutto, D. & Myin, E. (2013). *Radicalising Enactivism: Basic Minds without Content.* Cambridge, MA: MIT Press.

[2] It is worth pointing out that PEM is compatible with both externalist and internalist views about mentality. For example, Clark sees PEM as revealing those "key aspects of neural functioning that makes structuring our worlds genuinely continuous with structuring our brains and sculpting our actions" (2013, p194). Hence, PEM supports the situatedness of mentality, that is, it supports the view that the environment plays more than simply a causal role in underscoring human mentality. This is an externalist view. Hohwy, on the other hand, acknowledges that mind and world are "genuinely continuous", but understands this as revealing the fragility of our perceptual inferences, and so as evidence of our attempts to compensate for such fragility. That is, we structure our environments and/or our engagements with those environments so as to optimize the incoming sensory signal i.e. make it more precise and so improve our internal predictions. This ensures that the situatedness of mentality is, pace Clark, confirmation of the brain's seclusion from the hidden causes of the world. As such, the sensory boundary between brain and world is not malleable but rather "principled, indispensable, and epistemically critical" (Hohwy, 2013, p239). In other words, PEM supports an internalist view.