Consciousness and Cognition xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Noise, uncertainty, and interest: Predictive coding and cognitive penetration ☆

Jona Vance [a], Dustin Stokes [b,*]

[a] Department of Philosophy, Northern Arizona University, USA
[b] Department of Philosophy, University of Utah, USA

ARTICLE INFO

Article history:
Received 20 February 2016
Revised 30 May 2016
Accepted 6 June 2016
Available online xxxx

Keywords:
Predictive coding
Cognitive penetration
Perception
Cognition
Directness
Noise
Uncertainty
Attention
Prediction error
Theoretical interest

ABSTRACT

This paper concerns how extant theorists of predictive coding conceptualize and explain possible instances of cognitive penetration. Section 1 offers brief clarification of the predictive coding framework and of cognitive penetration. Section 2 develops more precise ways that the predictive coding framework can explain genuine top-down causal effects on perceptual experience. Section 3 develops these insights further with an eye towards tracking one extant criterion for cognitive penetration, namely, that the relevant cognitive effects on perception must be sufficiently direct. In Section 4, we analyze and criticize a claim made by some theorists of predictive coding, namely, that (interesting) instances of cognitive penetration tend to occur in perceptual circumstances involving substantial noise or uncertainty. We argue that, when applied, the claim fails to explain (or perhaps even be consistent with) a large range of important and uncontroversially interesting possible cases of cognitive penetration. We conclude with a general speculation about how the recent work on the predictive mind may influence the current dialectic concerning top-down effects on perception.

0. Introduction

As the various contributions to this special volume will no doubt attest, the predictive coding model of the mind and brain is remarkably ambitious in scope of explanation. For one, its central proponents allege that it offers a single, unified mechanism for phenomena that are traditionally treated as distinct and, accordingly, explained in disparate ways. As Jakob Hohwy writes,

> Perception, action, and attention are but three different ways of doing the very same thing. All three ways must be balanced carefully with each other in order to get the world right. The unity of conscious perception, the nature of the self, and our knowledge of our private mental world is at heart grounded in our attempts to optimize predictions about our ongoing sensory input.
>
> [Hohwy, 2013: 2]

☆ This work was thoroughly collaborative and the paper thoroughly co-authored—the order of authors was chosen randomly.

* Corresponding author at: Department of Philosophy, University of Utah, Carolyn Tanner Irish Humanities Building, 215 S Central Campus Dr., 4th Floor, Salt Lake City, UT 84112, USA.
E-mail addresses: jona.vance@nau.edu (J. Vance), dustin.stokes@gmail.com (D. Stokes).

http://dx.doi.org/10.1016/j.concog.2016.06.007
1053-8100/

Please cite this article in press as: Vance, J., & Stokes, D. Noise, uncertainty, and interest: Predictive coding and cognitive penetration. *Consciousness and Cognition* (2016), http://dx.doi.org/10.1016/j.concog.2016.06.007

The mechanism that allegedly unifies these phenomena—prediction error minimization—is also supposed to promise unity across the cognitive sciences. As Andy Clark writes, this framework

> makes rich and illuminating contact with work in cognitive neuroscience while boasting a firm foundation in computational modeling and Bayesian theory. It thus offers what is arguably the first truly systematic bridge [endnote omitted] linking three of our most promising tools for understanding mind and reason: cognitive neuroscience, computational modelling, and probabilistic Bayesian approaches to dealing with evidence and uncertainty.
>
> [Clark, 2013, 190–1]

Ambition such as this begets excitement, but it should also encourage careful scrutiny. Our general interest is at the interface of two standardly distinguished mental kinds or processes—cognition and sense perception. Both of the above theorists, as well as others, embrace to some degree a radical consequence of the predictive coding framework, namely, that the cognition/perception distinction will have to be revised in some important ways, if not abandoned outright. An important way to test this feature of the framework, as well as use it to shed new insight on the cognition/perception distinction, is to consider how the models within the framework analyze and explain a specific, possible cognitive-perceptual relation, namely, the cognitive penetration of perception. That is the topic of this paper: how do extant theorists of predictive coding conceptualize and explain possible instances of cognitive penetration? Jakob Hohwy (2013) has offered the most comprehensive analysis on this set of points, and so the analysis offered here focuses primarily on his discussions.[1]

The paper proceeds as follows. Section 1 offers brief clarification of the predictive coding framework and relevant mechanisms, and a brief characterization of cognitive penetration and some challenges that come with defining it. Section 2 develops more precise ways that the predictive coding framework can explain, and of course thereby allow for, genuine top-down causal effects on perceptual experience, of the kind discussed in the context of cognitive penetration. Section 3 develops these insights further with an eye towards tracking one extant criterion for cognitive penetration, namely, that the relevant cognitive effects on perception must be sufficiently direct. Throughout these discussions, we extend the analyses of the predictive coding models, as we know them. So one open question that surfaces is how much of the extended analyses are genuinely just part of the predictive coding models, or something that must be added to them in order to generate these additional explanatory benefits. In Section 4, we analyze and criticize a claim made by some theorists of predictive coding, namely, that (interesting) instances of cognitive penetration tend to occur in perceptual circumstances involving substantial noise or uncertainty. It is here that our analysis is most critical. We recognize, and indeed take pains to carefully identify, why the theorist of predictive coding is motivated to make this claim. However, we argue that, when applied, the claim fails to explain (or perhaps even be consistent with) a large range of important and uncontroversially *interesting* possible cases of cognitive penetration. We conclude with a general speculation about how the recent work on the predictive mind may influence the current dialectic concerning top-down effects on perception.

1. Predictive coding and cognitive penetrability

Broadly, predictive coding theories embrace an inferential approach to perception, which can be traced back at least to Helmholtz (1867). And Clark (2013) traces the idea of giving a central role for error signals in all brain processes to the influential cybernetician W. Ross Ashby, who claimed that "the whole function of the brain is summed up in: error correction" (Ashby, 1947). Clark interprets this provocative claim with a provocative claim of his own: "brains are essentially prediction machines" (Clark, 2013).

More precisely, predictive coding can be thought of as a framework within which one can develop specific, testable models of mental processing. The predictive coding framework is primarily distinguished from other frameworks by the following core commitment: a single mechanism of prediction error minimization plays a central role in all or most mental processes. Accordingly, let's refer to the overall framework as 'PCF' (predictive coding framework) and the central mechanism as 'PEM' (prediction error minimization).

Within the PCF, the most prominent extant proposals for how to develop models make further commitments about the processing that carries out prediction error minimization. These commitments entail that the processing is probabilistic, hierarchical, and dynamic. We illustrate these further commitments with respect to perceptual processing. Begin with the standard assumption that the objects, events, and features in the environment stimulate sensory receptors of a given perceiver. According to standard developments of PCF, perceptual processing draws on stored representations of the distal environment approximating a probability density function ("priors") and predictions concerning which sensory inputs are likely given the priors ("likelihoods"). So for any given hypothesis h and some bottom-up sensory information e, there are two probabilities of initial relevance: the prior probability of h independent of the current evidence, $P(h)$, and the likelihood, which is the probability of the occurrence of e given h, $P(e|h)$. Given Bayes' theorem, these two probabilities can be combined to yield the probability of h given e, $P(h|e)$, the posterior probability.

[1] We will, when relevant, attempt to connect our analysis with other examples of predictive coding theorizing. But again, beyond Hohwy (2013), there are few if any extended discussions of predictive coding *and* cognitive penetration. (And this we take it to be one of the virtues of this special issue: there will hopefully be further analysis of this interface.)

Please cite this article in press as: Vance, J., & Stokes, D. Noise, uncertainty, and interest: Predictive coding and cognitive penetration. *Consciousness and Cognition* (2016), http://dx.doi.org/10.1016/j.concog.2016.06.007