Consciousness and Cognition 47 (2017) 6–16

Contents lists available at ScienceDirect

Consciousness and Cognition

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

The relationship between cognitive penetration and predictive coding ☆

Fiona Macpherson

Centre for the Study of Perceptual Experience, Department of Philosophy, University of Glasgow, Glasgow G12 8QQ, Scotland, United Kingdom

ARTICLE INFO

Article history:
Received 1 February 2016
Revised 29 March 2016
Accepted 4 April 2016
Available online 22 April 2016

Keywords:
Cognitive penetration
Predictive coding
Perception
Cognition
Perceptual experience
Early vision

ABSTRACT

If beliefs and desires affect perception—at least in certain specified ways—then cognitive penetration occurs. Whether it occurs is a matter of controversy. Recently, some proponents of the predictive coding account of perception have claimed that the account entails that cognitive penetrations occurs. I argue that the relationship between the predictive coding account and cognitive penetration is dependent on both the specific form of the predictive coding account and the specific form of cognitive penetration. In so doing, I spell out different forms of each and the relationship that holds between them. Thus, mere acceptance of the predictive coding approach to perception does not determine whether one should think that cognitive penetration exists. Moreover, given that there are such different conceptions of both predictive coding and cognitive penetration, researchers should cease talking of either without making clear which form they refer to, if they aspire to make true generalisations.

1. Introduction

Some advocates of the predictive coding account of perception have said that if the predictive coding account of perception is correct then cognitive penetration occurs. If true, this would provide a new route to arguing for the existence of cognitive penetration: argue for a predictive coding account of perception. In this paper, I investigate whether the predictive coding account of perception does entail that cognitive penetration occurs. I establish that there are different versions of cognitive penetration and different versions of predictive coding. I map out the relationship between these different versions. I show that some versions of predictive coding entail that there are some forms of cognitive penetration, some versions of predictive coding are compatible with, but do not entail, that there are some forms of cognitive penetration, and some forms of predictive coding entail that some forms of cognitive penetration do not exist.

Mapping out these relationships helps us to understand the predictive coding account of perception and cognitive penetration in depth. It also serves to warn us that we should be clearer which version of the predictive coding account of perception and which version of cognitive penetration we are referring to when we make claims about them. And it serves to show that an argument for cognitive penetration that appeals to the existence of predictive coding has to appeal to a specific version of predictive coding, the establishment of which may turn on just the same issues as the establishment of cognitive penetration itself.

☆ This work was supported by a grant from the Arts and Humanities Research Council (grant number AH/L007053/1). Thanks to audiences in Bochum and Tokyo who heard a version of this paper for their helpful comments, and to Clare Batty for reading a draft of the paper and providing very many helpful suggestions to improve it.

E-mail address: fiona.macpherson@glasgow.ac.uk

http://dx.doi.org/10.1016/j.concog.2016.04.001
1053-8100/

This paper is divided into seven further sections. Section 2 defines and examines the different forms of cognitive penetration. Section 3 outlines a minimal account of the predictive coding account of perception. In Section 4, I consider possible relationships between predictive coding and cognitive penetration and examine the claims that people have made about this relationship. In Section 5, I look at whether predictive coding is committed to cognitive states playing a role in the predictive coding account, and what this means for the relationship between various versions of predictive coding and cognitive penetration. In Section 6, I examine whether a predictive coding account that is committed to the idea that cognitive states play a role in predictive coding entails one version of cognitive penetration, namely the cognitive penetration of perceptual experience—that is a conscious, personal-level state. In Section 7, I consider whether a predictive coding account that is committed to the idea that cognitive states play a role in predictive coding entails another version of cognitive penetration, namely the cognitive penetration of early vision—that is, an initial level of visual processing that has been defined in different ways by psychologists and neuroscientists. I give three accounts of what early vision might be taken to be and I spell out what predictive coders should say about the penetration of early vision conceived of in each of them. Finally, in Section 8, I summarise what the various accounts of predictive coding and cognitive penetration are, and the relationships between them, in a table, and I comment on the broad picture that emerges of the relationships between them. Readers may wish to consult this table while they are reading the sections below.

2. What is cognitive penetration?

There are two related, but nonetheless distinct, alleged phenomena that go under the name 'cognitive penetration'. That they are clearly distinct emerges in the rest of this paper by consideration of their relation to predictive coding.

The first of these phenomena is the penetration of *early vision*. It is the phenomenon discussed in Pylyshyn's (1999) much cited paper on the topic. Early vision is defined functionally by Pylyshyn as the system that takes attentionally modulated signals from the eyes (and perhaps some information from other sensory modalities) as inputs, and produces shape, size and colour representations as output. These representations are then categorised and identified by the cognitive system making use of memory, knowledge and judgment. The question of whether cognitive penetration occurs, conceived of as the penetration of early vision, is the question of whether the function that the early visual system computes "is sensitive in a semantically coherent way, to the organisms goals and beliefs, that is, [whether] it can be altered in a way that bears some logical relation to what the person knows" (1999: 343). Pylyshyn explicitly denies that the output of early vision either is, or determines, one's visual experience, or equivalently, that the representational content of the output of early vision is the representational content of one's perceptual experience (1999: 362).

In contrast to this notion of cognitive penetration, which has mostly been the subject of study by psychologists and neuroscientists, another notion of cognitive penetration, that of the penetration of *perceptual experience*, has been a major subject of study by philosophers, although the division in what researchers study is far from exceptionless, and increasingly so. 'Perceptual experience' refers to the conscious state that we typically go into when we perceive the world—the conscious state of awareness of the world that has a distinctive phenomenal character compared to typical beliefs or judgments about the world. In my (2012) paper on cognitive penetration, I defined the cognitive penetration of experience as follows: in any case of perception, hold fixed what it is that is perceived (the objects properties and relations seen, heard, touched, and so on), the perceiving conditions (the level of light, shadow, mistiness, for example), the state of the sensory organ (perfect human vision, shortsighted human vision, for example) and the location of one's focus of attention. With those conditions fixed, if it is possible for two subjects (or one subject at different times) to have different perceptual experiences due to the differing content of the states of their cognitive systems, and moreover, there is a semantic or intelligible link between the content of the cognitive states and the content of the perceptual experience, then perceptual experiences are cognitively penetrable. States of the cognitive system include beliefs, judgments and desires, and should likely also be taken to include the concepts that we possess. For further discussion of the terms used in this definition see my (2012) paper.

There are several issues with both of these definitions of cognitive penetration. I address the two most pressing ones. The first concerns attention. The second concerns the semantic or intelligible link that is posited as necessary for cognitive penetration to occur.

Despite giving the above definition of cognitive penetration in my (2012) paper, later in the paper, I go on to question whether we should include attention in the conditions that we hold fixed. I noted that one rationale that someone might have for thinking that we should hold fixed attention is that a shift of attention is akin to a shift of the location to which one's eyes point. A shift in the direction that one's eyes look changes one's perceptual experience by changing which objects and properties are processed, but if such a shift were driven by one's cognitive states, such a shift should not count as cognitive penetration occurring, simply a shift in what is perceived. One might think of shifts of attention in a similar manner: a shift of spatial attention, driven by one's cognitive states, might affect how clearly, or in how much detail, certain objects and properties are experienced, but that is just a shift in what is perceived and shouldn't count as cognitive penetration. However, as I noted in my (2012) paper, there are forms of attention other than spatial attention, so even if one was persuaded that changes of spatial attention driven by cognition should not count as cognitive penetration, it is not clear that shifts in other forms of attention should also not count. One form of a non-spatial shift in attention is a shift in what properties one attends to. For example, one can shift from attending to the red things in one's visual field to attending to the yellow things, and that might change one's experience in certain ways. For example, it may lead one to have

an experience in which the salience of the yellow things is greater than that of the red, and so one might be aware of the number yellow things but not be aware the number of red things. If such a shift were driven by cognition in the right way, then it is tempting to consider such a change in experience a case of cognitive penetration because a kind of bias could creep into perception. For example, one might overestimate the proportion of yellow things to red things. A useful discussion of what factors, such as bias, should lead one to classify certain cases as cases of cognitive penetration can be found in Stokes (2015).

However, it is even unclear whether shifts in spatial attention caused by cognitive processes should be conceived of as failing to count as cases of cognitive perception. An argument for thinking that some such cases are instances of cognitive penetration has recently been put forward by Cecchi (2014), who discusses an experiment carried out by Schwartz, Maquet, and Frith (2002). In the experiment, subjects underwent a period of monocular training (subjects in one group used their right eye and subjects in another group used their left eye). They looked at a screen on which was displayed a homogenous background of horizontal bars. An "L" or "T" would appear in the centre of the screen randomly rotated for 16 ms. During the same short interval, three lines would appear randomly at some location within the upper-left quadrant in the periphery of their visual field. Subjects had to do two tasks: identify whether the object in the centre was an "L" or a "T", and identify the orientation of the three lines that appeared in the periphery of their visual field. Subjects' eye movements were restricted to fixation on the central letter. Detection of the orientation of the lines in the periphery of their visual field therefore required the allocation of voluntary covert spatial attention. The goal of the experiment was to improve subjects' performance in the detection of the orientation of the three peripheral lines.

In the first (training) phase of the experiment, subjects underwent a total of 1760 trials each, and received feedback about their accuracy after each one. During the first phase, performance of the trained and untrained eye was measured. Successful performance was one in which they accurately detected what the central letter was and what the orientation of the peripheral lines was. During the first phase, subjects improved their performance of detecting the orientation of the peripheral lines gradually in the eye that was being trained, but not in the one that was not. Progressive changes in neural behaviour elicited by the trained eye in the part of the V1 cortex (also known as the striate cortex) associated with seeing in the upper-left quadrant were detected using functional magnetic resonance imaging (fMRI). This was compared to activity in that area elicited by the untrained eye and activity in that area elicited by the same eye before training.

According to Cecchi (2014: 85), the voluntary covert attention required to detect the orientation of the lines required guidance by the cognitive states of the subjects to select orientation as the feature to be processed, and top-down signals to the cortex were observed (using fMRI), and discovered to be necessary, for the task to be successfully carried out during the first phase. The subjects were investigated again 24 hours later (during the "test" session), after they had slept the night, and this time, again, the part of the V1 cortex corresponding to seeing in the upper-left quadrant was observed to have increased activity when activated by the trained eye. This is in contrast to when it was activated during the same test session by the untrained eye, or the same eye before training. In both cases, no increased activity was observed. However, this time, fMRI indicated that the increased activity observed in the test session when using the trained eye did not require that top-down cognitive or attentional influences occur. The increased responses occurred before any top-down influences were observed. The same task performed by the untrained eye did recruit higher-level cognitive and attentional influences.[1] According to Schwartz (2002: 17140), training resulted in neural changes to the V1 area of the cortex, which had learned to detect the orientation of the peripheral lines when seen by the trained eye.

Cecchi states:

> during the training session, the constant and synchronic stimulation of striate visual areas by cognitively guided attention resulted in physiological neural changes in the visual system. Therefore, structural modulations occurring in the visual system due to cognitively guided attention are the result of *synchronic architectural cognitive penetration*.[2]
>
> [2014: 90]

He also states,

> The successful detection of the peripheral object during the test session was possible thanks to *diachronic architectural cognitive penetration*. The consolidation of cognitively induced architectural modulations enabled the visual system to perform the peripheral-detection task without synchronic cognitive intervention. At that point, the cognitive influences that modified the architecture of the visual system were diachronically produced.
>
> [2014: 91]

Thus Cecchi argues that there is cognitive penetration brought about by means of attention. Cecchi argues that the penetration in question is both that of early vision and that of perceptual experience. With regard to the first, changes to the visual cortex occur in early vision, defined temporarily, as activity within the first 100 ms after stimulus presentation, because they occur around 60 ms after stimulus presentation. With regard to the second, there are changes to the content of visual experience, with experience coming to represent the orientation of the lines presented in the periphery of the visual field. In this way, he argues that there can be cases of cognitive penetration of the early visual system and of experience that involve attention.

[1] Cecchi (2014) discusses various follow-up replication studies which confirm and provide stringer evidence for the results of this study.

[2] Cecchi defines architectural cognitive penetration as "the process whereby the behaviour or the structure of the perceptual system is influenced by the cognitive system, which consequently may have an impact on the content of the perceptual experience" (2014: 63).

Cecchi certainly presents us with a tough case to adjudicate. An alternative interpretation of this case is that, in the first training phase of the experiment, the attention involved is simply selecting what to process—items in the upper-left quadrant and their property of orientation—and that this mere selection does not amount to cognitive penetration. As mentioned above, one might claim that this is just like selecting what to process as one might do by changing the direction in which one's eyes point—something that does not count as cognitive penetration. Moreover, one might claim that in the second test phase of the experiment, subjects had simply gained better eyesight, as one might gain from wearing glasses, which one has put on under the guidance of cognition (such as one's desire to see more clearly and one's belief that one's glasses will facilitate that), or as one might gain by voluntarily squinting one's eyes to see an object in the distance more clearly (caused by one's desire to see it and one's belief that squinting will accomplish that). Clearly such activities do not count as cognitive penetration.

It is not obvious to me how to adjudicate this case, but it is clear that the question of whether cognitive penetration—of either early vision or of perceptual experience—can involve attention is not a straightforward matter.

The second issue with the definition of cognitive penetration that I wish to discuss concerns why it specifies that, in order for cognitive penetration to occur, a semantic or intelligible link is necessary between the content of the cognitive state doing the penetrating and the content of the perceptual experience, or the content of early vision, that is penetrated. Is such a link necessary? For ease of exposition, I will discuss this case with respect to the cognitive penetration of experience, but readers will be able to reconstruct with ease a parallel discussion that could be made about the penetration of the content of early vision.

A semantic or intelligible link is posited as a necessary feature of cases of cognitive penetration because the lack of such a link seems to explain our intuitions about why some cases fail to be those of cognitive penetration.[3] In other words, there are some cases which, but for the lack of this link, would fall under the definition of cognitive penetration, but which we are rightly disinclined to think are cases of cognitive penetration. Here is one such case:

Migraine:	You believe that you have an exam tomorrow. The belief causes you to be stressed. The stress causes a migraine, which causes you to experience flashing lights in your visual field.

In *Migraine*, your belief affects your visual experience but, intuitively, this is not a case of cognitive penetration. It is merely an instance of a migraine brought on by belief-triggered stress. This is because there is no semantic or intelligible link between the belief that you have an exam tomorrow and the subsequent visual experience that you have as of flashing lights. If we hold that such a link is necessary for a case of cognitive penetration, its absence explains why the case is not one of cognitive penetration.

However, one might question whether the lack of such a link really does explain why such cases should not count as cases of cognitive penetration. For one can gerrymander the above type of case to yield one that is very similar, in which there does seem to be a semantic or intelligible link, and yet the case still does not appear to be one of cognitive penetration. For example, consider the following case:

Migraine's Revenge:	You believe that aliens might land on Earth. The belief causes you to be stressed. The stress causes a migraine which causes you to experience flashing lights in the sky.

In *Migraine's Revenge*, as in *Migraine*, your belief affects your subsequent visual experience; yet, intuitively, the case is not one of cognitive penetration, merely a case of migraine. However, in *Migraine's Revenge*, unlike in *Migraine*, there is a semantic or intelligible link between your belief that aliens might land on earth and your subsequent perceptual experience of flashing lights in the sky. (However improbable, it is entrenched in popular culture that aliens coming to Earth arrive in spaceships that have flashing lights.) So, it would seem that, in *Migraine's Revenge*, it is not the lack of a semantic link that explains why it is not a case of cognitive penetration. Given this, as well as its similarity to *Migraine*, one might conclude that the lack of such a link also does not explain why *Migraine* fails to be a case of cognitive penetration.

If the lack of a semantic link does not rule out both *Migraine* and *Migraine's Revenge* as cases of cognitive penetration, what other condition would? I believe that we can find a suitable condition, not by abandoning a semantic or intelligible link as a necessary condition for cognitive penetration, but by adding to, and thereby strengthening, the condition. I believe that what goes wrong in *Migraine's Revenge* is that the semantic or intelligible link is accidental. After all, I deliberately gerrymandered

[3] One might worry whether we do have intuitions about what counts as cognitive penetration and what doesn't. One might think that 'cognitive penetration' is a purely technical notion, the conception of which has been defined into existence, so how could there be intuitions about what it is? I think that such a view is wrong. I think that the notion of cognitive penetration is a notion that is intended to pick out a certain problematic form of interaction between cognitive states and perception—problematic because the form of interaction undermines the role of perception as a theory or cognition neutral independent source of evidence for belief. This is not to say that there can't be some instances of cognitive penetration that would make perception a better source of evidence than it might otherwise have been, but the form of that interaction also allows for there to be instances in which perception is a worse source of evidence than it would otherwise have been (and some cases in which it is neither better nor worse). Thus, I think that the definitions of cognitive penetration given in the literature can be criticized or praised to the extent that they accurately pick out such a form of interaction between cognition and perception.

Migraine to yield a case in which there just so happened to be such a link. I believe that what is required in cases of cognitive penetration is that there be a causal, semantic link between each of the steps in the chain that lead from the belief to the subsequent perceptual experience.[4] This does not exist in *Migraine's Revenge*. That case is one in which one's belief that aliens might land causes one to be stressed, which causes a migraine, which in turn causes one's perceptual experience of flashing lights. While there is a semantic link between the content of the belief and the content of the perceptual experience, the link is not a causal link. Moreover, there is not a causal, semantic link between the content of the belief and the subsequent state of stress. There is a causal link but not a semantic one. A state of stress is a physiological state that does not have content. Even if one thought that states of stress could have content (perhaps content about the source of the stress or the content that something bad was going to happen), the state of having a migraine is clearly not such a state. So there is no causal, semantic link between the state of one being stressed and the state of having a migraine. Moreover, there is no causal, semantic link between the state of having a migraine and the state of experiencing flashing lights in the sky.

Note that I proposed (2012) a mechanism by which cognitive penetration could occur. It was one in which a belief caused some imaginative state or imaginative processing to come into existence, which in turn causally interacts with one's experience yielding an experience that has content influenced both by perception and imagination. This mechanism could be a mechanism whereby cognitive penetration occurs because it allows for the existence of causal, semantic links at each stage of the process. I take it to be a virtue of the proposed mechanism that it could allow for a causal, semantic link between each of the steps, which I take to be necessary for cognitive penetration.

In this section, I have looked at what cognitive penetration is and how it should be defined. I noted that there are two distinct phenomena that go under its name: penetration of the early visual system and penetration of perceptual experience. I explained that there are hard questions to answer about whether there can be a role for attention in cases of cognitive penetration, and I explained why I believe that a necessary condition for cognitive penetration is that there be a causal, semantic link between each step in the process whereby a belief comes to affect perceptual experience.

In the next section, I consider what predictive coding accounts of perception are by outlining the minimal commitments of such an account.

3. Predictive coding accounts of perception

The minimal commitments of the predictive coding account of perception are as follows:

(i) Based on knowledge, assumptions, and expectations at different high-level processing stages, the brain produces top-down generative representations of the world that are predictions of how the world is.
(ii) Those representations are then modified bottom-up by incoming signals, which are propagated upwards only as a prediction error signal.
(iii) Those representations are then further modified by top-down high-level processing, and perhaps sideways by other sensory modalities.
(iv) The aim of these modifications are to reduce global prediction error.
(v) The processing is typically thought to be done according to Bayesian rules.[5]

I will refer to (i)–(v) as the 'minimal form of predictive coding'. As we will come to see in later sections, there are many, more specific, forms of predictive coding accounts of perception. We will see the ramifications of these further forms for the relationship between predictive coding and cognitive penetration.

4. Possible relationships between predictive coding and cognitive penetration

There are three possible relationships between predictive coding accounts of perception and cognitive penetration:

(a) predictive coding entails cognitive penetration;
(b) predictive coding is compatible with, but does not entail, cognitive penetration;
(c) predictive coding is incompatible with cognitive penetration.

In each case, 'cognitive penetration' might mean either penetration of the early visual system or penetration of perceptual experience.

One motivation for exploring which of (a), (b) or (c) is true (or, as we will come to see below, whether any *one* of these positions is true) is because they have been endorsed by different people.

[4] Note that semantic (or information) processing is, with all likelihood, realized or carried out by processes that can be described by a lower-level vocabulary involving only physical, non-semantic terms. For example, it could be that a process that carries information that *p* is a process that consists in a particular neuron firing at a certain rate. A causal, semantic process may, therefore, be describable in non-semantic terms. The existence of such a description would not undermine the process being one in which there is semantic processing. What is important is that there is a level of description of the process that is such that each step can correctly be described in causal, sematic terms.

[5] See Clark (2013), Hohwy (2013), and Lupyan (2015a).

Arguably Lupyan (2015a, 2015b) endorses (a). He certainly comes very close to explicitly endorsing it. For example, he states:

> penetrability should be expected whenever constraining lower-level processes by higher level knowledge minimizes global prediction error,
>
> [2015a: 547]

and:

> My goal, however is not to simply argue that perception can be cognitively penetrated, but to present a theoretical framework on which cognitive penetrability is expected as a natural part of how perception works."
>
> [2015b: 588]

Moreover, he states:

> The primary goal of this paper is to show that the controversy surrounding cognitive penetrability of perception (CPP)—the idea that perceptual processes are influenced by "non-perceptual" states—vanishes when we view perception not as a passive process the goal of which is recreation of a veridical reality, but rather as a flexible and task-dependent process of creating representations for guiding behavior.
>
> [2015a: 548]

> My goal is to . . . argue that there is no in-principle limit on the extent to which a given perceptual process can be penetrated by knowledge, expectations, beliefs, etc.
>
> [2015a: 549]

Thus Lupyan thinks that if the predictive coding model of perception is correct then we should expect there to be cognitive penetration, and that the model entails that mechanisms exist that allow it to happen.

Clark (2013), prima facie at least (but I will explore other interpretations of him below), believes that predictive coding entails cognitive penetration too. He says that predictive coding:

> makes the lines between perception and cognition fuzzy . . . In place of any real distinction between perception and belief we now get variable differences in the mixture of top-down and bottom-up influence . . . To perceive the world just is to use what you know to explain away the sensory signal across multiple spatial and temporal scales. The process of perception is thus inseparable from rational (broadly Bayesian) processes of belief fixation, and context (top-down) effects are felt at every intermediate level of processing.
>
> [2013: 10]

In my (2015) paper, I argued for a version of (b): that the minimal form of predictive coding is consistent with, but does not entail, the cognitive penetration of perceptual experience.[6] Drayson (unpublished manuscript) claims this too. With respect to the penetration of early vision, no one has yet commented. I suspect that is because this issue is particularly thorny on account of what predictive coders can say about what early vision is—a topic that I address in Section 7.

I know of no one who explicitly endorses (c), however, I think that one reading of Clark (2013) would have him believe (c) with respect to the penetration of perceptual experience. I will explore this in Section 5. What to make of (c) with respect to early vision, is a question that requires detailed consideration that, as previously stated, I will provide in Section 7.

The most expeditious way to determine whether (a), (b), or (c) is true is not to address them each in in order but to consider, in turn, the following three questions:

I. Do cognitive states feature in the predictive coding account of perception?
II. If cognitive states do feature, is there reason to think that they penetrate perceptual experience?
III. If cognitive states do feature, is there reason to think that they penetrate early vision?

I do so, respectively, in Sections 5,6, and 7.

5. Do cognitive states feature in the predictive coding account of perception?

Should the high-level processes posited by predictive coders be taken to include cognitive states of subjects? By 'cognitive states' I mean doxastic states—states that are accessible to consciousness and inferentially integrated—such as beliefs and desires. Or should they be taken to be sub-doxastic information-carrying states of the brain that the subject does not in principle have access to?[7] The minimal account of predictive coding, outlined above, simply states that perceptual representations of the world are produced by high-level processes—high-level in relation to the level of perceptual experiences. The account leaves it open whether those high-level processes include cognitive states.

[6] I also gestured at some other forms of predictive coding and what relationship they might have to cognitive penetration, but not in the detail that I do in this paper.

[7] The doxastic/sub-doxastic distinction is introduced by Stich (1978) and it is in the sense that he spells out in that paper that I use it, as I have explained in the main text. Doxastic states are often called 'personal-level' states, and sub-doxastic states are often called 'sub-personal' level states, but there is reason to resist that terminology as Drayson (2012) clearly explains.

This shows that the minimal form of the predictive coding account does not *entail* cognitive penetration of either the early visual system or experience. Does it show that predictive coding is consistent with cognitive penetration sometimes or always occurring? The answer will depend on whether we hold that cognitive states sometimes or always affect early vision or perceptual experience—an issue that I will address shortly.

One could simply stipulate two different versions of the predictive coding account each of which would be a precisification of the minimal account. According to a first, the high-level states that make the predictions sometimes or always involve cognitive states; according to a second, the high-level states in question never involve cognitive states. Whether the first version entails or is consistent with cognitive penetration will depend on whether we think cognitive states sometimes or always affect early vision or perceptual experience—as was the case with the minimal account. The second account would entail that cognitive penetration of either early vision or perceptual experience never occurs, for cognitive states are not involved in the production of perceptual representations.

Which version of the theory do predictive coders actually put forward? Predictive coders often talk about knowledge, beliefs, information, expectations, priors, and so on feeding into the production of perceptual representations from high levels. But should we take some of this talk as literally talk of cognitive states?

Lupyan (2015a) is very clear that he thinks we should:

> there is no *in-principle* limit on the extent to which a given perceptual process can be penetrated by knowledge, expectations, beliefs, etc. The *actual* extent to which such penetrability happens can be understood in terms of whether it helps to lower system-wide (global) prediction error.
>
> [2015a: 549]

> In evolving to minimize prediction error neural systems naturally end up incorporating whatever sources of knowledge, at whatever level, to lower global prediction error.
>
> [2015a: 551]

And, in addition, it is clear that the specific examples of cognitive penetration that Lupyan (2015a) discusses are ones in which semantic doxastic states of knowledge and belief, as well as other cognitive states, affect perception. This view would entail that there is cognitive penetration if and only if those cognitive states can be seen to affect early vision or perceptual experience (as Lupyan argues that they do). Hohwy (2013) argues for the same position. In Sections 6 and 7, I will investigate this later claim for predictive coding accounts of perception in which cognitive states play a role. For now, however, in the rest of this section, I will explore whether other philosophers accept that cognitive states play a role in the predictive coding account of perception, and what follows from that about cognitive penetration.

Consider Clark's (2013) view. Although I noted above in Section 4 that, prima facie, he holds that cognitive states penetrate perceptual states, there is quite a bit of room to doubt this. He often talks about sub-personal states of the brain having a top-down influence and, when he writes about "belief" having an influence, he often puts the word in quotation marks, signalling an unorthodox usage. Moreover, he never explicitly states that the highest level corresponds to doxastic level cognitive states. And while he clearly recognises that perception can influence doxastic belief, he doesn't explicitly mention the reverse case (2013: 17). Furthermore, when he mentions the debate concerning whether perception is theory-laden he says that it is "in at least one (rather specific) sense: What we perceive depends heavily upon the set of priors ... that the brain brings to bear in its best attempt to predict the current sensory signal" (2013: 7), thereby not explicitly mentioning cognitive states. All of this might suggest that Clark does not think that cognitive states are involved.

This is the reading of Clark that Drayson endorses. She says: "I suggest that Clark's 'assumptions' are like Churchland's assumptions in that they 'have nothing to do with beliefs, theories, or other doxastic commitments that we may have' ... (Stokes, in press)" (unpublished manuscript: 4). In consequence, she claims: "[Clark's] model of the brain is entirely consistent with the traditional view of perception and cognition. In particular, the debate over the cognitive penetrability of perception is left untouched" (unpublished manuscript: 1).

While this radical reading of Clark is not implausible, and there could be a version of predictive coding that stipulated cognitive states did not play a role, it seems to me that there is an even more radical reading of Clark that is at least as good a reading of Clark as Drayson's. On this super radical reading, Clark thinks that there is no distinction, or at least no sharp distinction, between perception and cognition. He says:

> These accounts thus appear to dissolve, at the level of the implementing neural machinery, the superficially clean distinction between perception and knowledge/belief. To perceive the world just is to use what you know to explain away the sensory signal across multiple spatial and temporal scales. The process of perception is thus inseparable from rational (broadly Bayesian) processes of belief fixation, and context (top-down) effects are felt at every intermediate level of processing. As thought, sensing, and movement here unfold, we discover no stable or well-specified interface or interfaces between cognition and perception. Believing and perceiving, although conceptually distinct, emerge as deeply mechanically intertwined.
>
> [2013: 10]

On this interpretation of Clark, one should be eliminativist about the normal categories such as perception and cognition that traditional views of the mind ascribe to. If that is right, then according to such a view, there can be no cognitive penetration for cognitive penetration requires such a distinction.

When I discussed (2015: 582) this view as a possible version of predictive coding Lupyan replied that, "I am aware that in the process of challenging these distinctions", such as,

> the presence of a clear boundary between (modal) perception and (amodal) cognition/semantics . . . I am supporting a collapse of perception and cognition that makes the whole question of the penetrability of one by the other, ill-posed. But I would be thrilled if I my arguments contribute to the eventual demise of this question".
>
> [2015b: 58]

Here Lupyan is more cautious with respect to his previous (2015a) commitment to the role of cognitive states in predictive coding and their role in cognitive penetration. Clearly, this is a view of the mind that predictive coders may consider adopting. This version of the predictive coding account of perception entails that there is no cognitive penetration of perception (either of early vision or of perceptual experience), for there is no cognition and there is no perception.

A slightly less radical reading of Clark would be that cognition and perception lie on either ends of a spectrum and that cognitive penetration happens when there is the right kind of influence by states that lie on the more cognitive end of the spectrum on states that lie on the more perceptual end. Indeed, Clark sometimes writes as if this were the case, saying that the higher-level something is the "more cognitive" and the lower is the "more perceptual" (2013: 10). One might surmise that the highest levels correspond to doxastic beliefs and desires and the lowest levels to perceptual states. On this account of predictive coding, it entails that there is cognitive penetration.

Still, the picture that Clark paints is unclear between the four options just outlined. And Lupyan seems as yet undecided about his final view. Thus, predictive coders should, I think, give further consideration to these questions, specify more precisely which version of predictive coding they wish to endorse, and in that light, reconsider the claims that they make about the relationship between predictive cognitive penetration.

To summarise this section, the minimal version of predictive coding leaves open whether cognitive states are some of the high-level states that generate perceptual representations. Given this, the minimal version does not entail, but is consistent with, cognitive penetration. Lupyan's (2015a) and Hohwy's (2013) version of predictive coding stipulates that cognitive states are some of the high-level states. If those cognitive states can be seen to affect early vision or perceptual experience, this version of the theory entails cognitive penetration. (I will explore this issue in detail below). Clark is not particularly clear about whether his version of predictive coding stipulates that cognitive states are some of the high-level states. A super radical reading of him would suggest that there is no distinction between cognition and perception and so no cognitive penetration. Interestingly, we saw Lupyan (2015b) indicate that he would consider such a position. A less radical reading of Clark would suggest that perception and cognition lie on a continuum and that cognitive penetration occurs when states that lie far on the cognitive side affect states that lie far on the perceptual side (in the right way).

6. If predictive coding involves cognitive states, is there reason to think that they penetrate perceptual experience?

Recall that in Section 2, I outlined two forms of cognitive penetration. One was the penetration of perceptual experience; the other was the penetration of early vision. In this section, I will focus on the former. Recall that perceptual experience is a conscious state that has a distinctive perceptual phenomenal character. It is the state that we go into when we consciously see, hear, touch, and so on, but also a state that we go into when we consciously hallucinate. The phenomenal character of a perceptual experience of a red rose in front of one (or better as of a red rose in front of one, to include the case of hallucination) is distinct from the phenomenal character of the corresponding belief that there is a red rose in front of one—a belief that one often has when one experiences a red rose (but not always, for one sometimes believes, rightly or wrongly, that one may be suffering from an illusion or hallucination).

If cognitive states are involved in the predictive coding that produces perception, is there reason to think that those cognitive states penetrate perceptual experience? In order to answer this question, we must ask: what exactly does the predictive coding account of perception say about perceptual experience?

Clark (2013) explains what he and Hohwy, Roepstorff, and Friston (2008) hold is perceptual experience, according to the predictive coding account:

> [it] is determined by a process of prediction operating across many levels of a (bidirectional) processing hierarchy . . . and their interactive equilibrium ultimately selects a best overall (multiscale) hypothesis concerning the state of the visually presented world. This is the hypothesis that 'makes the best predictions and that, taking priors into consideration, is consequently assigned the highest posterior probability' (Hohwy et al., 2008, p. 690). Other overall hypotheses, at that moment, are simply crowded out: they are effectively inhibited, having lost the competition to best account for the driving signal."
>
> [Clark, 2013: 185]

In short, perceptual experience is identified with the best—i.e. surviving—prediction or representation of how the world is. And that is generated by high-level and low-level processing.

If one holds that the high-level states are cognitive states (or cognitive enough), then this will entail that perceptual experience is affected by cognition according to predictive coding, and that there is cognitive penetration. Hohwy (2013) holds this position. (Of course, as we have seen, one might desist from this opinion and hold a version of predictive coding that is silent about whether high-level states are cognitive, that denies that cognitive states are involved, or that denies the

cognition/perception distinction. These versions will not entail that there is cognitive penetration, with the last two entailing that there is no cognitive penetration.)

So far, things have been reasonably straightforward. However, the matter becomes more complex when we consider what predictive coding says about the penetration of early vision, as I explain in the next section.

7. If predictive coding involves cognitive states, is there reason to think that they penetrate early vision?

Recall from Section 2, that early vision was defined functionally as the system that takes in information from the eyes (perhaps via attention) and produces shape, size and colour representations. How might predictive coders understand early vision?

One option for predictive coders would be to think of early vision, neutrally, as the system—whatever it is—that produces shape, size and colour representations. According to the predictive coding account, high-level systems produce those representations, causally affected by information from the eyes. The information from the eyes therefore feeds into the high-level system that produces those representations; therefore, the whole system, including high-level states will count as early vision. That's a very odd result—and a very strange conception of early vision! Suppose, for the sake of argument, that we accepted this conception. (I realise that one might object to it.) In that case, a version of the predictive coding account that said that cognitive states were involved would entail that cognitive penetration of early vision occurred. (And if a version of predictive coding was non-committal about whether cognitive states were involved (such as the minimal version) then it would be consistent with but not entail cognitive penetration of early vision. If a version denied that cognitive states were involved, then the view would entail that there was no cognitive penetration of early vision.)

As I indicated above, one might object to this account of early vision on the grounds that it is at odds with the conception of it conceived of by Pylyshyn (1999). Therefore, one might reconsider what predictive coders should take early vision to be. Pylyshyn thought that his functional definition would be realized by signals coming from the eyes, perhaps after being filtered by attention, and then a limited portion of bottom-up processing thereafter. Given this, one approach predictive coders might adopt would be to follow Raftopolous and espouse a temporal conception. Raftopoulos (2014: 603) states: "Early vision . . . includes a feed forward sweep (FFS) in which signals are transmitted bottom-up and which lasts, in visual areas, for about 100 ms."

If we take the first 100 ms of processing to be early vision, what does predictive coding say about that? Does the high-level predictive signal reach down as low as this, or does the first 100 ms contain only an incoming signal that will, in due course, be treated as an error signal? To my knowledge, the current literature on predictive coding does not address this question. Standard evidence concerning whether there is any top-down influence on the first 100 ms of processing would be what we would look to.[8] Whether there is penetration of early vision, considered as the first 100 ms of processing, will depend on such evidence (and, of course, as we have seen previously on whether such high-level processes include cognitive states of subjects).

There is another conception of early vision that a predictive coder might appeal to. Recall the functional definition given by Pylyshyn: it is the system that takes in information from the eyes (perhaps via attention) and produces shape, size and colour representations. A predictive coder might claim, contra Pylyshyn, that there is no system that takes in information from the eyes (perhaps via attention) and produces shape, size and colour representations. Instead, they could claim that such representations are produced top-down. This would amount to the denial that there is such a thing as early vision! If that is right, then predictive coding is incompatible with the penetration of early vision. Predictive coding entails that there is no cognitive penetration of early vision as there is no such thing as early vision.

To summarise this section, I have outlined that there are three conceptions of early vision that someone who endorses the predictive coding account of perception might adopt:

(1) that it includes all processing (including top-down processing) that produces shape, size and colour representations;
(2) that it is that which occurs during the first 100 ms of processing;
(3) that it doesn't exist.

Predictive coding entails penetration of early vision conceived of as in (1) if and only if the high-levels contain cognitive states. Predictive coding entails, or is compatible with, penetration of early vision conceived of as in (2) if and only if the empirical evidence turns out the right way, and the high-levels contain cognitive states. Predictive coding in any form is not compatible with penetration of early vision conceived as in (3).

8. Conclusion

The results of the above discussion of various forms of predictive coding accounts of perception and various different forms of cognitive penetration are summarized in Table 1.

[8] Evidence of the kind presented by Schwartz et al. (2002) that was discussed in section two.

Table 1
The relationship between different forms of both predictive coding and cognitive penetration.

	Penetration of experience	Penetration of early vision (functional definition)	Penetration of early vision (100 ms definition)	Penetration of early vision (denial it exists)
Minimal predictive coding account	Consistent with penetration	Consistent with penetration	Either consistent with penetration or no penetration (depending on empirical matters)	No penetration
Predictive coding specifying involvement of cognitive states	Entails penetration	Entails penetration	Either entails penetration or no penetration (depending on empirical matters)	No penetration
Predictive coding denying involvement of cognitive states	No penetration	No penetration	No penetration	No penetration
Predictive coding denying a perception/cognition distinction	No penetration	No penetration	No penetration	No penetration
Predictive coding specifying a continuum along which lies more and less cognitive and more and less perceptual states	Entails penetration	Entails penetration	Either entails penetration or no penetration (depending on empirical matters)	No penetration

As discussed previously, some proponents of the predictive coding account of perception have claimed that it entails that cognitive penetrations occurs. I have argued that the claim that the predictive coding account entails cognitive penetration is too simplistic. This is because the relationship between the predictive coding account of perception and cognitive penetration is dependent on both the specific form of the predictive coding account and the specific form of cognitive penetration. I have spelled out different forms of each and the relationships that hold between them. While some forms of predictive coding entail some forms of cognitive penetration, some are consistent with but do not entail some accounts of cognitive penetration, and some entail that there is no cognitive penetration. However, many of those forms of predictive coding that entail that cognitive penetration does not occur do so because they give such a radically different account of the mind compared to the standard conception of it. That is, they rule out the existence of cognitive penetration on the grounds that the components of the mind required to play a role in cognitive penetration do not exist. As we have seen, some accounts deny that there is a distinction to be made between perception and cognition and some deny that there is any such thing as early vision. Such models allow for a great deal of top-down processing of the sort that those who have heretofore argued that there is no cognitive penetration would not endorse.

In consequence, mere acceptance of the predictive coding approach to perception does not determine whether one should think that cognitive penetration exists. Moreover, as is demonstrated by the preceding sections, establishing as true the version of predictive coding that does entail cognitive penetration occurs will turn on determining many of the same issues as are required for directly establishing cognitive penetration occurs, such as determining that there are there top-down cognitive effects by cognitive states on perceptual experience or early vision. Thus, arguing for cognitive penetration via establishing that predictive coding is the right account of perception does not actually provide a novel means of establishing that cognitive penetration occurs. Nonetheless, it is clear that many forms of predictive coding models of perception either entail that cognitive penetration occurs or they entail that top-down processing of a radical kind occurs—of such a radical kind that they entail a picture of the mind greatly at odds with that accepted by those who believe in cognitive penetration.

Finally, a general lesson of this paper is that, given that there are such different conceptions of both predictive coding and cognitive penetration, researchers should cease talking of either without making clear which form they refer to, if they aspire to make true generalisations.

References

Cecchi, A. (2014). Cognitive penetration, perceptual learning, and neural plasticity. *Dialectica, 68*(1), 63–95.

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

Drayson, Z. (2012). The uses and abuses of the personal/subpersonal distinction. *Philosophical Perspectives, 26*(1), 1–18.

Drayson, Z. (unpublished manuscript). *Perception, prediction, and penetration.*

Hohwy, J. (2013). *The predictive mind.*New York: Oxford University Press.

Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition, 108*(3), 687–701.

Lupyan (2015a). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology, 6*(4), 547–569.

Lupyan (2015b). Reply to Macpherson: Further illustrations of the cognitive penetrability of perception. *Review of Philosophy and Psychology, 6*(4), 585–589.

Macpherson, F. (2012). Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research, 84*(1), 24–62.

Macpherson (2015). Cognitive penetration and predictive coding: A commentary on Lupyan. *Review of Philosophy and Psychology, 6*(4), 571–584.

Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*, 341–423.

Raftopoulos, A. (2014). The cognitive impenetrability of the content of early vision is a necessary and sufficient condition for purely nonconceptual content. *Philosophical Psychology, 27*(5), 601–620.

Schwartz, S., Maquet, P., & Frith, C. (2002). Neural correlates of perceptual learning: A functional MRI study of visual texture discrimination. *Proceedings of the National Academy of Sciences, 99*, 17137–17142.
Stich, S. P. (1978). Beliefs and subdoxastic states. *Philosophy of Science, 45*, 499–518.
Stokes, D. (in press). Cognitive penetrability of perception. *Philosophy Compass* (in press).
Stokes, D. (2015). Towards a consequentialist understanding of cognitive penetration. In A. Raftopoulos & J. Ziembekis (Eds.), *Cognitive effects on perception: New philosophical perspectives*. Oxford: Oxford University Press.