Perception and Disjunctive Belief: A New Problem for Ambitious Predictive Processing

Assaf Weksler

University of Haifa & The Open University of Israel

Penultimate draft

Forthcoming in Australasian Journal of Philosophy

Perception can't have disjunctive content. Whereas you can *think* that a box is blue or red, you can't *see* a box as being blue or red. Based on this fact, I develop a new problem for the ambitious predictive processing theory, on which the brain is a machine for minimizing prediction error, which approximately implements Bayesian inference. I describe a simple case of updating a disjunctive belief given perceptual experience of one of the disjuncts, in which Bayesian inference and predictive coding pull in opposite directions, with the former implying that one's confidence in the belief should increase, and the latter implying that it should decrease. Thus, predictive coding fails to approximately implement Bayesian inference across the interface between belief and perception.

Keywords: predictive coding, predictive processing, Bayesian brain, the content of perception, the perception-cognition interface

1   Introduction

Perception can't have disjunctive content. For example, whereas you can *believe* that a box is blue or red, you can't *see* a box as being blue or red. You can see it as blue and you can see it as red, but not as blue-or-red. Consider, also, the duck-rabbit ambiguous figure. Whereas you can think that an object is a rabbit or a duck, you cannot see it as a rabbit-or-duck. Instead, your perception fluctuates between seeing it as a duck and seeing it as a rabbit (these examples are adapted from Block forthcoming). Call this the thesis of non-disjunctive content (of perception), or ND.

Aside from its apparent intuitiveness, ND is, plausibly, also entailed by the influential claims that perception is iconic (Fodor 2008; Block forthcoming), or analogue (Beck 2019). Pautz (2020) holds that ND is among the 'laws of appearance', which govern experiences.

In this paper I argue that ND creates trouble for the predictive processing (PP) theory, understood as applicable to perception, cognition, and their interaction (Hohwy, Roepstorff, and Friston 2008; Hohwy 2013; Clark 2013, 2015, 2020; Lupyan 2015). Call this theory *ambitious PP*. The theory has two distinct components (Orlandi and Lee 2019): (1) A Bayesian model of the mind; and (2) a hierarchical predictive coding (PC) algorithm. Moreover, according to the theory, *the second component approximately implements the first*.

PC opposes the traditional view of the mind on which information about the world is transmitted mainly bottom-up, i.e., from low-level perceptual representations, to higher-level cognitive representations (e.g., beliefs), with top-down influences playing a minor role. Proponents of PC reverse this picture, claiming that information is mainly transmitted top-down in the form of predictions, with bottom-up signals serving merely as prediction error signals. The main function of the mind, on this view, is to minimize prediction error. Specifically, there is a hierarchy of hypotheses, where higher-level hypotheses, to which a certain probability is assigned, generate predictions, which are matched against lower-level hypotheses. A mismatch leads to a prediction error signal being sent upwards. As a result, the probability of the higher-level hypothesis (which generated the erroneous prediction) is lowered. As mentioned, according to PP, predictive coding approximately implements Bayesian inference, in which the probability of a given hypothesis 'A' is updated in light of sensory evidence 'B' in accordance with Bayes' rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$P(A|B)$ is the *posterior probability*, $P(B|A)$ is the *likelihood*, and $P(A)$ is the *prior probability* (or *prior*).

In this paper, I use ND to construct a new problem case for ambitious PP, concerning the interface between perception and belief, and specifically concerning the way disjunctive

belief is updated in light of perceptual experience of one of the disjuncts. In the example I develop, Bayesian inference and PC pull in opposite directions, with the former implying that one's confidence in the belief should increase, and the latter implying that it should decrease. Thus, PC fails to implement Bayesian inference in this case. This means that the two components of PP conflict with each other, in the case I consider, which implies that PP fails to apply to the interface between belief and perception, and hence ambitious PP seems to be false.

The basic line of argument is that if a disjunctive belief generates a *perceptual* prediction, the disjunctive content is lost in the process, because of ND. The loss of the disjunctive content leads—in the example I develop—to the mistaken formation of a prediction error signal, which, given the aim of prediction error minimization, leads to lowering the confidence the subject has in the disjunctive belief. This occurs even when the rules of Bayesian inference entail that confidence in the disjunctive belief should increase. Thus, PC entails a result that conflicts with the rules of Bayesian inference.

While my argument focuses on top-down predictions, from beliefs to perception, it does not concern cognitive penetrability. If it did, one could try to resist my argument by proposing that disjunctive beliefs can't penetrate perception. But even if cognitive penetration never occurs, the argument will still go through, since it focuses on *updating beliefs* in light of one's perceptual experience (via predictions and prediction error signals), and not on the way *perceptual experience changes* in light of one's beliefs (i.e., cognitive penetration).

Unlike Williams (2020), who argues that PP can't account for certain features of beliefs, and specifically for their generality and compositionality, my argument focuses only on the process of updating (disjunctive) beliefs in light of perceptual experience. So while both arguments focus on beliefs and their structure, they differ in the aspect of PP that they target.

Section 2 covers relevant background on ND and PP. In section 3 I present the argument from ND against ambitious PP. And in section 4 I consider objections.

2 Background

2.1 Clarifying ND

First, ND concerns both inclusive and exclusive disjunctions. The duck-rabbit case is relevant not only to exclusive disjunctions, but to inclusive ones too. If your perception could have an inclusive disjunctive content, you would be able (contrary to fact) to see a duck-rabbit figure as a duck-or-rabbit-or-both, *without fluctuations*. My main argument focuses on inclusive disjunction.

Second, according to ND, perception is non-disjunctive in all of the following senses: it does not represent a disjunction of propositions (e.g., there is a duck here or there is a rabbit here) and also not a disjunction of properties, e.g., of being a duck and of being a rabbit (so that it's veridical iff one of them is instantiated), and further it does not represent disjunctive properties, e.g., of being a-duck-or-a-rabbit (so that it's veridical iff this property is instantiated).

Third, for the purposes of my central argument, I don't need ND to apply to every possible case of perception. It's sufficient that it applies to the simple cases of Rawa and Adi that I consider below. Note also that below I discuss a case of experiencing a continuous range, which is equivalent to a disjunction of properties. Even if this is a plausible case of a disjunctive perception, it is not (as I argue below) of the right form, and so faces a different version of my argument.

2.2 Background on predictive processing

First, a prediction error signal does not always lead to lowering the probability of the higher-level hypothesis that generated the prediction. This occurs only when the prediction error signal is estimated to be sufficiently precise. If it is estimated to be insufficiently precise, as when, e.g., the viewing conditions are poor, it is inhibited (Hohwy 2013: 66–68).

Second, leading proponents of PP present it as a theory of the mind *as a whole*. That is, they accept ambitious PP. As Clark recently put it, PP accounts 'suggest that neurally realized predictions are the only fundamental "cognitive kind" needed to explain the full sweep of human behavior'. (2020: 1) and PP 'has been positioned as a new paradigm for

understanding perception, reason and action'. (2020: 2). According to Hohwy, PP 'is meant to explain perception and action and everything mental in between'. (2013: 1).

Third, ambitious PP is supposed to be applicable to conscious perception. On the PP framework, the *best* perceptual hypothesis—the one with the highest posterior probability—is selected for consciousness. Along these lines, Hohwy, Roepstorff, and Friston (2008) have developed an impressive PP account of binocular rivalry.

Fourth, Clark (2015) and Lupyan (2015) hold that beliefs penetrate perception whenever doing so lowers prediction error. They need to explain why perceptual illusions persist, i.e., are not cognitively penetrated by contrary beliefs, even though such penetration would reduce prediction error. To address this, they distinguish between local and global prediction error, and claim that the PP system can tolerate local prediction errors if doing so lowers global prediction error. On their view, perceptual illusions persist because, while allowing a belief to override (i.e., penetrate) them reduces local prediction error, it increases global prediction error. More specifically, they hold that perceptual illusions are computed based on certain assumptions that are globally important for perception, such as the assumption that light typically comes from above. To override an illusion, the belief must override the assumption in question. But if a belief is allowed to override such an important assumption in the case of illusion, they claim, it will also override the assumption in cases of veridical perception, which will lead to many prediction errors. So on a global level, prediction error is increased if an illusion is allowed to be overridden by belief.

Fifth, the PC framework is often described in terms of probability distributions, not single values (see Orlandi and Lee 2019; Williams 2020). A probability distribution describes the probabilities of different possibilities within a given range. For example, the probability of there being a circle ahead is (say) 0.5, of there being a triangle ahead is 0.3, of there being a square ahead is 0.2, etc. On typical formulations of the PC framework, a probability distribution of hypotheses at a given level is updated in light of a probability distribution of hypotheses at the level below. The case I develop in the next section and the discussion that follows do not concern distributions of hypotheses. I discuss a case in which a single belief

(a high-level hypothesis) is updated in light of a single experience (a lower-level hypothesis). But there is no problem here for my argument. For one thing, recall that in the case of conscious perception, a single hypothesis is selected from the distribution of posterior probabilities. Thus, probability distributions of beliefs are supposed to be updated in light of a *single* perceptual hypothesis. For another, it seems that if PC fails to update, in a Bayesian way, a single hypothesis (belief) in light of another single hypothesis (perceptual experience) at the level below, it will, *a fortiori*, fail to update a distribution of hypotheses in light of a single hypotheses at the level below.

Sixth, the PC hierarchy is ordered according to spatiotemporal scale of contents (Clark 2013; Hohwy 2013). Applied to the perception/cognition distinction (Hohwy 2013: 72), the idea is, roughly, that perceptual hypotheses concern relatively short time scales and small spatial regions. For example, perception of a triangle might involve a hypothesis about the precise orientation of its left edge at a temporal scale of milliseconds. In contrast, a belief that there is a triangle ahead concerns a larger spatiotemporal scale: it is committed to there being a triangle ahead, but not to specific fine-grained spatial details that change rapidly.

Seventh, the claim that PC *approximates* Bayesian inference does not mean that in some cases PC implements Bayesian inferences and in other cases it implements some other kind of inference. Instead, the claim is that in all cases (of bottom-up influence), PC implements Bayesian inference, but not perfectly. For example, if, according to Bayes' rule, the probability of a hypothesis H should increase by 0.7 given perceptual evidence E, we expect the PC algorithm to increase it by *roughly* that amount, say by 0.6 - 0.8 (the exact numbers do not matter).

3   The argument from ND against ambitious PP

Before moving to the main argument of this paper, let me illustrate how a simple perceptual belief is updated in light of a perceptual experience, given Bayes' rule. For concreteness, I assign specific probabilities to propositions here and throughout the paper.

The precise probabilities do not matter for the argument, and they can be replaced with verbal descriptions such as *high confidence*, *miniscule probability*, etc.

Suppose you believe with confidence 0.6 that there is a triangle ahead. Suppose that your visual system is intact and that viewing conditions are optimal. You open your eyes and experience a triangle. Intuitively, your confidence that there is a triangle ahead should increase to a degree of complete confidence. Let's now see how this result is obtained using Bayes' rule. $P(Etri \mid tri)$, the likelihood, is the probability of having an experience as of a triangle (Etri) given that there is a triangle ahead (tri). This probability is very high, say 0.999. $P(tri)$ is the probability that there is a triangle ahead, regardless of the experience. It is the prior probability and (as assumed above) it equals to 0.6. $P(Etri)$ is the probability of having an experience as of a triangle. It equals to the probability of having an experience as of a triangle given that there is a triangle ahead (0.999, as I've assumed) times the probability of there being a triangle ahead (0.6) plus the probability of having an experience as of a triangle given that there is no triangle ahead (i.e., having an illusion of a triangle, under optima viewing conditions), which is miniscule (say 0.00001), times the probability of there being no triangle ahead (0.4). Thus, $P(Etri) = 0.999*0.6 + 0.00001*0.4 = 0.599404$. Putting all of this together, $P(tri \mid Etri)$, i.e., the probability that there is a triangle ahead given an experience as of a triangle—the posterior probability—is equal to

$$\frac{P(Etri|tri)P(tri)}{P(Etri)} = \frac{0.999 * 0.6}{0.599404} = {\sim}0.9999$$

which is very high, amounting to *complete confidence*.

I now turn to the main argument of this paper—the argument from ND against ambitious PP. Suppose Rawa stands in a room with her eyes closed, and believes with degree of confidence 0.8 that there is a circle ahead, and with degree of confidence 0.6 that there is a triangle ahead[1]. The existence of the circle is independent of the existence of the triangle, and vice versa, and Rawa is aware of this. She should believe with degree of confidence

---

[1] I am assuming that assigning probabilities to hypotheses in the PP framework is equivalent to assigning degrees of confidence to beliefs. But it is also possible to formulate the argument directly using hypotheses and not beliefs, see Section 4.6.

higher than 0.8 (specifically 0.92) that there is a circle *or* a triangle ahead, in accordance with the general disjunction rule $P(A \text{ or } B) = P(A) + P(B) - P(AB)$ [2]. Suppose Rawa, who has a normal visual system, opens her eyes and experiences a triangle, in optimal viewing conditions. Given Bayes rule, this perceptual evidence should raise Rawa's confidence in the proposition that there is a triangle ahead from 0.6 to the degree of complete confidence (0.9999 as I've explained in the previous paragraph). Moreover, given the general disjunction rule, after experiencing the triangle, Rawa should raise her confidence that there is a circle or a triangle ahead, from 0.92 to a degree of confidence higher than 0.9999 [3]. More generally, assuming a normally functioning visual system and optimal viewing conditions, if the prior confidence in the truth of the disjunction, and in the truth of each disjunct, is fair (neither very low nor extremely high), then having an experience of one of the disjuncts should—given Bayes' rule—raise the confidence that the disjunction is true to complete confidence. I will now argue that, given ND, PC delivers the wrong result concerning this simple case of Bayesian belief updating.

Given PC, Rawa's belief that there is a circle or a triangle ahead (with degree of confidence > 0.8) generates a disjunctive prediction that is transmitted top-down into the perceptual system through the hierarchy. When the prediction reaches perception, then, given ND, it must be *translated* into a non-disjunctive prediction. For, a disjunctive content, and *a fortiori* a disjunctive prediction, cannot exist at a perceptual level (below I discuss an objection to this claim). A natural suggestion is that the result is a 'circle' perceptual prediction and a 'triangle' perceptual prediction, without the logical connective 'or' (below I discuss alternative characterizations of the prediction). Recall that Rawa is looking at a triangle and experiences it as such. This means that the perceptual 'triangle' hypothesis is selected for consciousness (since it is the best one—with the highest posterior probability). While the aforementioned 'triangle' perceptual prediction matches this perceptual hypothesis, the aforementioned 'circle' perceptual prediction does not, leading to the formation of a prediction error signal. This prediction error signal is a byproduct—an

---

[2] Since A and B are mutually independent, one can use the restricted conjunction rule: $P(AB) = P(A)P(B)$. Thus, $P(A \text{ or } B) = P(A) + P(B) - P(AB) = P(A) + P(B) - P(A)P(B) = 0.8+0.6-0.8*0.6 = 0.92$.

[3] Assuming that the posterior probability of the proposition that there is a circle ahead drops to, say, 0.001, when looking and seeing only a triangle, $P(A \text{ or } B) = 0.001 + 0.9999 - 0.001*0.9999 = 0.9999001$.

artifact—of the mistranslation between the cognitive prediction and the perceptual one. It occurs just because the 'or' is lost in the translation. If the 'or' were retained in the translation, the perceptual prediction ('circle or triangle') would match the perceptual 'triangle' hypothesis, and so an error signal would not be formed[4]. Next, the prediction error signal is sent upwards to the level of the original disjunctive belief. Given the aim to minimize prediction error, the PC system will lower the confidence with which the belief is held in order to reduce prediction error (below I discuss a complication concerning local vs. global prediction error). But, as explained in the beginning of this section, this is the wrong result: in light of her experience, Rawa should raise her moderate confidence in the disjunctive proposition to the level of complete confidence, not lower it! Thus, PC entails a result that is incompatible with Bayesian inference.

Since PP holds that PC approximately implements Bayesian inference, it follows that PP fails to apply to the interface between belief and perception, and hence ambitious PP seems to be false. This is the argument from ND against ambitious PP. I will henceforth refer to it as *the argument from ND* for short.

I would like to emphasize that, in the example, PC does not even *approximate* Bayesian inference. It is not as if, say, Bayesian inference implies raising the posterior probability to 0.999 but PC raises it only to 0.94. Instead, PC updates the posterior probability in *the opposite direction* from the direction required by Bayesian inference. In other words, in the case in question, PC appears to implement some *other*, non-Bayesian kind of inference.

In claiming that, given PC, the formed prediction error signal should lead to lowering of the probability of the disjunctive belief, I have implicitly assumed that the prediction error signal (i.e., the sensory signal) is *estimated to be sufficiently precise* (otherwise the prediction error signal would be inhibited, see section 2.2). The assumption is justified

---

[4] To be more precise, *a* (circle-related) prediction error signal would still be formed, because Rawa believes that there is a circle ahead, and this leads to prediction error. The mere formation of a circle-related prediction error signal is not by itself a problem. The problem, as explained in the rest of the section, is that a (circle-related) prediction error signal leads to lowering the confidence in the disjunctive belief that produced it. This highlights the fact that my argument presupposes that disjunctive beliefs are tested (via PC) against perceptual evidence, *in addition* to the fact that logically simpler, non-disjunctive beliefs (such as the belief that there is a circle ahead) are tested against perceptual evidence. I defend this claim in Section 4.7.

since in the case described viewing conditions are assumed to be optimal and the visual system is assumed to function normally.

4 Objections and replies

4.1 Perceptual confidence

Some philosophers hold that perception represents probabilistically (Munton 2016; Morrison 2016). When seeing a red dress in a dim light, for example, perception assigns to the content 'scarlet' a certain probability and to the content 'crimson' a different probability. The defender of ambitious PP might suggest that Rawa's disjunctive belief should generate a 'circle' perceptual prediction and a 'triangle' perceptual prediction, without an 'or' connective, but with probability assignments, for example 0.8 and 0.6 (see section 3) or 0.46 and 0.46 (i.e., dividing the probability of the disjunction by 2). However, because the 'or' connective is gone, the original problem from Section 3 reappears. Suppose Rawa is looking at a triangle. Consequently, the 'circle' prediction fails to match the experience and an error signal is generated, which requires lowering the confidence that the disjunctive belief is true, in order to reduce prediction error. But this, again, is the wrong result. As explained in Section 3, seeing the triangle should raise Rawa's confidence that there is a triangle or a circle ahead.

4.2 Indeterminate content

According to Nanay (2020) and Raleigh & Vindrola (2021), perception can have indeterminate contents. A representation of an indeterminate content is in a certain sense equivalent to a disjunctive representation. You are looking at a tree from a distance, with your glasses off. You see it blurrily. Your experience, on the view in question, represents the tree as having a height in the continuous range of 2…2.5 meters, say. Now a height of 2-2.5 meters is equivalent to the disjunction '2 meters or 2.1 meters or 2.11 meters or…'. In this way, it seems that perception can represent disjunctively, contrary to ND.

Let us grant for the sake of argument that this objection is correct. To handle it, I introduce a weakened form of ND:

ND*

> Perception can disjunctively represent A and B only by representing the
> continuous range A...B.

So, suppose perception needs to represent 'a small circle (the size of a coin) *or* a large circle (the size of a Frisbee)', it must do so by representing a continuous range of circles, covering all the sizes that are intermediate between the coin and the Frisbee sizes, such as the size of a drink coaster (even though the target disjunction did not involve these intermediaries at all).

This leads to a problem that is structurally similar to the original problem from Section 3, only in reverse. Suppose Nandi (who has a normal visual system) believes that there is a coin-sized circle or a Frisbee-sized circle ahead (in her line of sight), with a degree of confidence of 0.6. Given PC, this belief should generate a prediction that is transmitted top-down into the perceptual system. When the prediction reaches perception, then, given ND*, it must be translated into a prediction of a continuous range of sizes of circles, including the intermediate size of a drink coaster. Suppose Nandi is looking at a drink-coaster-sized circle, under optimal viewing conditions, and experiences it as such. This implies that the hypothesis that is currently selected for consciousness (the best perceptual hypothesis) is that there is a drink-coaster-sized circle ahead (in Nandy's line of sight), and nothing else. The aforementioned range-of-circle-sizes perceptual prediction *matches* this hypothesis. Consequently, no prediction error signal is formed. Thus, given PC, Nandy's confidence that the belief is true will not decrease. But this result is wrong, from the point of view of Bayesian inference. Nandi is visually experiencing only a drink-coaster-sized circle ahead. Her belief is that there is a coin-sized circle or a Frisbee-sized circle ahead. Clearly, the experience provides evidence against her belief (her belief is that there is an A or a B ahead, and she experiences only a C ahead, which is different from both A and B). The probability of 'A or B is ahead' given an experience of 'C is ahead and nothing else' should be much

lower than the original 0.6 probability, close to zero in fact[5]. Thus, from a Bayesian point of view, Nandy should significantly lower her confidence in the truth of the disjunctive belief.

4.3   Disjunctive predictions, non-disjunctive hypotheses

On the PP framework, we can consciously experience only the contents of *hypotheses*, not of the *predictions* (of lower-level hypotheses) that they generate, or of the *prediction error signals* that are sent upwards. The predictions' role is to update hypotheses (via the mechanism of prediction error minimization), not to directly contribute to consciousness (see Hohwy, Roepstorff, and Friston 2008's account of binocular rivalry). ND is a thesis about conscious perception, and so one might object that ND is a constraint on the hypotheses selected for consciousness, rather than on the computational (and predictive— according to PC) mechanisms that update these hypotheses. In short, ND concerns only perceptual *hypotheses*, not perceptual *predictions*. On this proposal, although a perceptual hypothesis cannot have a 'circle or triangle' content, a perceptual prediction can have such a content. This move appears to block the argument from ND against ambitious PP. In presenting the argument, I have assumed that because of ND, a disjunctive cognitive prediction generated by a disjunctive belief ('there is a circle or a triangle ahead') is translated into a non-disjunctive perceptual prediction. On the present objection this is not true: the perceptual prediction is disjunctive too.

To respond to this worry, we need to consider neuroscientific accounts of the computation of prediction errors. These accounts share the idea that a prediction P (coming from a higher-level hypothesis) amounts to a reconstruction, within 'error neurons', of the expected activity pattern of the population of neurons underlying the hypothesis at the level below. The error neurons subtract this reconstructed activity from the actual one, or vice versa, and the remaining activity functions as an error signal (for a recent account of

---

[5] Here is an illustration: Assume that the prior probability that C is ahead is 0.3. The probability of having an experience as of C (*Ec*) given that C is ahead is very high, say 0.999. The probability of having Ec when there is no C ahead—i.e., an illusion as of C—is very low (the visual system is functioning normally in normal conditions), say 0.00001. So the prior probability P(Ec) is equal to 0.999*0.3 + 0.00001*0.7 =~0.3. Plugged into Bayes's rule:

$$P(A \text{ or } B|Ec) = \frac{P(Ec|A \text{ or } B)*P(A \text{ or } B)}{P(Ec)} = \frac{0.00001*0.6}{\sim0.3} = \sim0.00002.$$

this sort, see Keller and Mrsic-Flogel 2018). So, a perceptual prediction is a reconstruction of the neural activity underlying a possible perceptual hypothesis. In short, a perceptual prediction is a reconstructed perceptual hypothesis. Thus, since perceptual hypotheses can't be disjunctive, perceptual predictions likewise can't be disjunctive.

Let me close this subsection by briefly mentioning a different argument, which seems to be *prima facie* compelling, for the conclusion that perceptual predictions can't be disjunctive, given ND and ambitious PP[6].

1. If perceptual predictions can be disjunctive, then perceptual Bayesian inference can be disjunctive (given ambitious PP).
2. If a perceptual Bayesian inference (a perceptual hypothesis H_1 is updated in light of a lower-level perceptual hypothesis H_2) can be disjunctive, then at least one of the *perceptual hypotheses* it operates with (H_1 or H_2) can be disjunctive.
3. But perceptual hypotheses can't be disjunctive (given ND).
4. Thus, perceptual predictions can't be disjunctive.

4.4   Double negation

The disjunctive thought that there is a triangle or a circle ahead is logically equivalent to the thought, 'it is not the case that there is no triangle and no circle ahead'. Call this *the double negation formula*. One might propose that, though a disjunctive belief cannot be translated into a disjunctive perceptual prediction, it can be translated into a perceptual prediction with a double negation formula content. If this were right, the argument from ND against ambitious PP would be blocked. However, we cannot imagine what an experience as of a not (not-circle and not-triangle) ahead would be like. It thus seems that perceptual experience can't have a double negation formula content.

In other words, we can't imagine having an experience that would be veridical iff it's not the case that there is neither a circle nor a triangle ahead. Consider an apparent counterexample. Imagine a rapid antigen test that is valid iff a circle or a triangle appear in the test window. Equivalently, the test is valid iff it's not the case that neither a circle nor a

---

[6] I would like to thank an anonymous reviewer for suggesting this argument.

triangle appear in the window. We can imagine that, when one looks at a test in which *only a circle appears*, one can see the test as containing a circle in the test window and also as valid (assuming for the sake of discussion that we can see high-level properties in general, and specifically the property of validity). It seems that the content of the imagined experience is: 'there is a circle and a not (not-circle and not-triangle) in the test window'. It might seem that this experience has a double negation formula content, but this is not true: the experience is veridical only if there is a circle in the test window, yet an experience with the content, 'there is a not (not-circle and not-triangle) in the test window', is veridical even if there a triangle, not a circle, in the test window. The point is that we can't imagine seeing a test as having a not (not-circle and not-triangle) in the test window without also imagining seeing a specific shape—a triangle, a circle, or both—in the test window. Consequently, the accuracy conditions of the imagined perceptual experience are not those of the double negation formula.

Moreover, the duck-rabbit consideration mentioned at the beginning of the paper applies equally to the double negation formula: if you could see the duck-rabbit figure as a not (not-rabbit and not-duck), your perception would not fluctuate between seeing it as a duck and seeing it as a rabbit.

4.5   Local vs. global prediction error

I have argued that the PC algorithm (as applied to cognition, perception and their interface) implies that Rawa's confidence in the disjunctive belief should be lowered in order to reduce prediction error, but from a Bayesian perspective this lowering of confidence is wrong. We have seen that Lupyan and Clark explain the occurrence of perceptual illusion by invoking a distinction between local and global prediction errors, arguing that illusions are not penetrated by beliefs because doing so increases global prediction error, i.e., increases prediction error in other cases. Lupyan and Clark focus in their discussion on cognitive penetrability alone, i.e., on top-down influence of belief on perception (updating perception in light of belief). A defender of ambitious PP might try to extend Lupyan and Clark's approach so that it would cover bottom-up effects as well (updating belief in light of perception). On this proposal, a prediction error signal causes lowering of confidence in the

belief that generated the prediction only if doing so does not increase global prediction error, i.e., does not increase prediction errors in other cases. So if lowering Rawa's confidence in the disjunctive belief in light of the bottom-up local prediction error increases prediction error in other (non-disjunctive) cases, the PC system will block this bottom-up influence: Rawa's confidence will not be lowered, thereby avoiding the conflict between PC and Bayesian inference.

Even if this strategy can avoid the conflict between PC and Bayes' rule in Rawa's case, it seems to generate a similar conflict elsewhere[7]. For, the idea that a prediction error signal lowers confidence in a belief only if doing so does not increase global prediction error, implies that the PC algorithm updates beliefs in accordance with Bayes' rule only when doing so does not increase global prediction error. Thus, the strategy in question implies that sometimes prediction error reduction implements some other, non-Bayesian, inference, contrary to ambitious PP.

The original proposal by Lupyan and clark, which concerns cognitive penetrability, does not have this problematic implication, since it does not concern the updating of beliefs (or higher-level hypotheses) in light of perceptual evidence (or lower-level hypotheses) at all, and Bayesian inference is restricted to this sort of updating (Bayes' rule tells us how to update beliefs in light of evidence, using priors and likelihoods).

In any case, the proposed strategy seems to fail to accommodate Rawa's case. The proposal is that allowing the prediction error signal to lower Rawa's confidence in the disjunctive belief would lead to prediction error in other cases. But, on its face, allowing a prediction error signal to lower Rawa's confidence in the disjunctive belief seems to have no influence whatever on the way other kinds of beliefs (e.g., conjunctive) are updated in light of perceptual evidence.

4.6 Eliminating beliefs, denying the perception-cognition distinction, denying representations

---

[7] I would like to thank an anonymous reviewer for pointing this out.

Ambitious PP is a revisionist view. It is therefore open for ambitious PP supporters to adopt eliminativism about beliefs, holding that the hypotheses in their PP hierarchy are not beliefs (for a claim roughly along these lines see Dewhurst 2017). If there are no beliefs then there are no disjunctive beliefs and so the issue of the predictions they generate does not arise. However, even if there are no disjunctive beliefs, the argument from ND will go through (*mutatis mutandis*) if there are disjunctive *hypotheses.* Moreover, as I explain next, it is difficult to deny, given an PP framework, that we (or our brain) have disjunctive hypotheses.

Suppose you are playing a game of dice, and if you roll '5' in the next round you lose. You know that the chances of getting '5' when rolling one die is low, and so you are relatively confident that you won't get '5' in one roll. But you also know that the chance of getting '5' when rolling six dice—i.e., in the first die or in the second die or … in the sixth—is high. You act on this hypothesis in the game, choosing to roll one die and not six. Given that on the ambitious PP framework, you represent the world via hypotheses to which probabilities are assigned, it seems to follow that you have a *disjunctive hypothesis* that the first die or the second die or…the sixth die will land on '5', and to this hypothesis high probability is assigned. It is difficult to see how this could be denied, from within a PP framework.

A different kind of example involves knowing the probability that a disjunction is true without having any information about the probabilities of the individual disjuncts. Suppose Kathrine, who is reliable, tells you that Jones owns a Ford or Brown is in Barcelona, but she doesn't tell you anything about the disjuncts themselves. So as far as you know, the probability that Brown is in Barcelona could be very low. Suppose also that the stakes are high: a lot hangs on whether or not Brown is in Barcelona. In this case, it seems, you should form a disjunctive hypothesis on the basis of Kathrine's testimony. That is, you should hypothesize with high confidence, that Jones owns a Ford or Brown is in Barcelona. Such a hypothesis will allow you, later on, to infer, e.g., that Brown is probably in Barcelona, from evidence that Jones does not own a ford. So, again, it is difficult to deny that we are able to perform inferences of this sort, and that this requires—on the PP framework—disjunctive hypotheses.

The idea of denying the existence of beliefs relates to a somewhat different revisionist strategy: Andy Clark has suggested that PP

> makes the lines between perception and cognition fuzzy, perhaps even vanishing. 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, and differences of temporal and spatial scale in the internal models that are making the predictions. (2013: 190)

When Clark says that the distinction between belief and perception is 'fuzzy, perhaps even vanishing', he does not explicitly consider the claim that whereas beliefs can be disjunctive, perceptions (at least in simple cases such as Rawa's) can't. But it seems very costly to deny it. That perception in the simple case of Rawa cannot be disjunctive seems to be an obvious fact about conscious, subjective experience, and proponents of PP want to explain facts about subjective experience, not deny them, as Hohwy, Roepstorff, and Friston's (2008) influential account of binocular rivalry demonstrates. Furthermore, that we can have disjunctive beliefs, or at least disjunctive hypotheses, is also hard to deny, for the reasons described above.

Finally, Downey (2018) proposes that the cognitive states within the PC framework are not representational. Some of them are merely covariational, others are merely biases: they fail to meet the metaphysical requirements for being genuine representational states. This, however, does not matter for the argument from ND: the problem I raise for ambitious PP concerns the PC framework, with its hierarchy of hypotheses, predictions, and prediction errors, regardless of whether these state are truly representational, metaphysically speaking.

4.7   A two-step process for updating disjunctive beliefs

Consider Rawa's case again. An objector could argue as follows:

> PC can unproblematically be used to calculate the probability of 'circle' given an experience of a triangle alone (say 0.00001), and the probability of 'triangle' given an experience of a triangle alone (say 0.999), and from this

the probability of the disjunction 'circle or triangle' can be calculated, using the general disjunction rule (it is higher than 0.999). There is thus no need to us PC to calculate the probability of the disjunction given the sensory evidence. On this story, the role of PC is only to calculate the probability of each disjunct given the sensory evidence—thus implementing Bayes' rule— but not to calculate the probability of the disjunction.

In short, according to the objection, updating Rawa's disjunctive belief consists of two steps, the first applies Bayesian inference to each disjunct, and the second applies the general disjunction rule the outputs of the first step. Only the first step is implemented by PC, and this is how it should be: after all, PC is supposed to be an algorithm of Bayesian inference and not for every rule of probability theory, such as the general disjunction rule. Thus, on this picture, *disjunctive beliefs do not generate predictions at all (let alone perceptual predictions)*, and consequently the argument from ND appears to be blocked.

In response, consider Adi, who believes with high confidence (say 0.92) that there is a triangle or a circle ahead, in her line of sight (with eyes closed). Further, assume that she has no idea what the probability of there being a triangle ahead, and of there being a circle ahead, are. She opens her eyes and sees a square (and only a square). Assuming that her visual system operates normally, and that the viewing conditions are optimal, Bayes rule dictates that her confidence that there is a triangle or a circle ahead should significantly drop[8]. But in this case the two-step updating process is not available. Bayes rule cannot be used to update the probability of each disjunct. For, to calculate the posterior probability of, say, 'circle', one needs to know what the prior probability of 'circle' is, but (*ex hypothesi*) Adi does not know what it is. In this case, Bayes rule must therefore be applied *directly* to the disjunction. So, if Bayes' rule is implemented by PC then disjunctive beliefs (and not only beliefs about individual disjuncts) must generate perceptual predictions after all.

An anonymous reviewer has proposed a different objection, but with a somewhat similar upshot: Perhaps hypotheses at levels closest to the sensory periphery never encode disjunctive contents: disjunctive beliefs are located higher (in the hierarchy) than logically

---

[8] Roughly: $P(\text{tri or cir} \mid \text{Esqr}) = \dfrac{P(\text{Esqr} \mid \text{tri or cir}) * P(\text{tri or cir})}{P(\text{Esqr})} = \dfrac{0.00001 * 0.92}{0.00001 * 0.92 + 0.999 * 0.08} = 0.00015.$

simple ones. In other words, perhaps disjunctive beliefs are never directly tested against sensory evidence, and hence they never generate perceptual predictions, but only doxastic ones, which undermines the argument from ND. This, however, conflicts with the claim that the PC hierarchy is ordered according to the spatiotemporal scale of their contents (see Section 2.2). For, the content 'there is a circle or a triangle here, now' has the same spatiotemporal scale as does the content 'there is a circle here, now'. Moreover, Adi's case shows—contrary to the objection in question—that disjunctive beliefs must generate perceptual predictions: Adi's disjunctive belief must be directly tested against sensory input.

4.8   Disjunctive formation of an error signal

A defender of ambitious PP might propose that, although there are no disjunctive perceptual hypotheses/predictions, the *formation* of prediction error signals is disjunctive. That is, in Rawa's case, where a disjunctive higher-level hypothesis leads to two distinct perceptual predictions (i.e., of a triangle and of a circle), which are each independently tested against a lower-level perceptual hypotheses (i.e., a triangle perceptual hypothesis), the formation of a prediction error signal is disjunctive in the following sense: a prediction error signal is produced (and sent upwards) if and only if both of the two separate predictions don't match the perceptual hypothesis. Since the 'triangle' perceptual prediction matches the triangle perceptual hypothesis, no prediction error signal is formed, which blocks the argument from ND[9].

However, on the PC framework, a higher-level hypothesis can send downwards predictions of lower-level hypotheses, or predictions (estimations) of the precision of the sensory signal. A higher-level hypothesis can't send any additional information downwards. Consequently, the higher-level hypothesis can't 'tell' the error-formation mechanism (the mechanism responsible for sending upwards a prediction error signal) that the two separate predictions (of a triangle and of a circle) should be treated in a disjunctive way

---

[9] I would like to thank an anonymous reviewer for raising this objection.

(rather than in a conjunctive way). Thus, the present proposal appears to be incompatible with the PC framework.

4.9   A fluctuation between two different predictions

I have assumed that, given the PP framework, when a subject holds a disjunctive belief, two perceptual predictions are sent downstream *at the same time*, each corresponding to one of the disjuncts. A friend of PP might deny this assumption, proposing instead that the two perceptual predictions are sent downstream one at a time. For example, it might be that the perceptual system first produces a prediction based on one disjunct (circle), gets a prediction error, thus lowering the confidence in the disjunction. Next the system tests the second disjunct (triangle), gets no prediction error, thus raising the confidence in the disjunctive belief. But then the first disjunct is again tested against the input, and so on. In this way, the confidence in the disjunction wobbles but remains stable[10].

The problem with this proposal is that adds a new layer of complexity to the standard PC algorithm, since it posits a mechanism that activates a fluctuation between two predictions (i.e., it inhibits one prediction for a few moments, then the other, and so on), when it is needed. This proposal is *ad hoc*: a fluctuation between predictions is not a part of standard PC algorithms, and adding it might make the algorithm overly complex, and hence less likely to be implemented in the visual cortex.

A defender of this proposal might claim that while it looks *ad hoc*, it in fact resembles the influential PC analysis of binocular rivalry (Hohwy, Roepstorff, and Friston 2008), which involves a fluctuation between two different predictions in the visual cortex. However, this resemblance is misleading. The analysis of binocular rivalry does *not* posit a special mechanism that generates the fluctuation. Instead, this fluctuation occurs as a natural consequence of the standard PC algorithm operating in the special circumstances in which different images are projected to each eye. Indeed, the strength of this analysis is precisely in the fact that it shows how the standard PC algorithm can explain binocular rivalry *without* positing any additional mechanism (Clark 2015: 36).

---

[10] I would like to thank an anonymous reviewer for raising this objection.

5   Conclusion

If ambitious PP were true, disjunctive beliefs (or at least disjunctive hypotheses) would generate perceptual predictions, but given ND (or the weaker ND*), this would lead to cases of belief updating that conflicts with Bayes' rule. Ambitious PP thus appears to be false: its second component conflicts with its first: predictive coding can't implement Bayesian inference across the belief-perception interface, not even approximately.

In response, proponents of ambitious PP can try to weaken their view somewhat while retaining many of its core benefits. On one weakened view, the mind as a whole is still Bayesian, and PC implements Bayesian inference within the doxastic system and within the perceptual system, *but not across the interface between the two*. A different algorithm implements Bayesian inference across the belief-perception border. Aitchison and Lengyel (2017) review non-PC algorithms that implement Bayesian inference. In these algorithms, posterior probabilities are computed without involving prediction and errors at all. Such algorithms may be unaffected by the argument from ND.

On a different weakened view, the mind as a whole is governed by PC, and PC implements Bayesian inference within the doxastic system and within the perceptual system, but not across the belief-perception border. Across that border, PC implements some other function. For example, when Rawa sees the triangle, the *practical* benefit of the belief that there is a circle or a triangle ahead diminishes. Maybe, then, the prediction error signal does not lower the posterior probability of the disjuncive belief, but instead lowers something else, namely its (subjective) practical value.

Acknowledgements

ORCID

Assaf Weksler's ORCID iD: 0000-0002-6902-1155

References

Aitchison, Laurence and Máté Lengyel (2017) 'With or Without You: Predictive Coding and Bayesian Inference in the Brain', *Current Opinion in Neurobiology* **46**: 219–27. doi:10.1016/j.conb.2017.08.010

Beck, Jacob (2019) 'Perception is Analog: The Argument from Weber's Law', *The Journal of Philosophy* **116**: 319–49. doi:10.5840/jphil2019116621

Block, Ned (forthcoming) *The Border between Seeing and Thinking*. MIT Press.

Clark, Andy (2013) 'Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science', *Behavioral and Brain Sciences* **36**, 181–204. doi:10.1017/S0140525X12000477

Clark, Andy (2015) *Surfing uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.

Clark, Andy (2020) 'Beyond Desire? Agency, Choice, and the Predictive Mind', *Australasian Journal of Philosophy* **98**: 1–15. doi:10.1080/00048402.2019.1602661

Dewhurst, Joe (2017) 'Folk Psychology and the Bayesian Brain'. In Thomas Metzinger and Wanja Wiese, eds., *Philosophy and Predictive Processing*, ch. 9. *MIND Group*. doi:10.15502/9783958573109

Downey, Adrian (2018) 'Predictive Processing and the Representation Wars: A Victory for the Eliminativist (via Fictionalism)', *Synthese* **195**: 5115–39. Doi:10.1007/s11229-017-1442-8

Fodor, Jerry (2008) *LOT 2: The Language of Thought Revisited*. Oxford University Press.

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

Hohwy, Jakob, Andreas Roepstorff, and Karl Friston (2008) 'Predictive Coding Explains Binocular Rivalry: An Epistemological Review', *Cognition* **108**: 687–701. doi:10.1016/j.cognition.2008.05.010

Keller, Georg B and Thomas D Mrsic-Flogel (2018) 'Predictive Processing: A Canonical Cortical Computation', *Neuron* **100**: 424–35. doi:10.1016/j.neuron.2018.10.003

Lupyan, Gary (2015) 'Cognitive Penetrability of Perception in the Age of Prediction: Predictive Systems are Penetrable Systems', *Review of Philosophy and Psychology* **6**: 547–69. doi:10.1007/s13164-015-0253-4

Orlandi, Nico and Geoff Lee (2019) 'How Radical is Predictive Processing?', In Mateo Colombo, Elizabeth Irvine, and Mog Stapleton, eds., *Andy Clark and His Critics*: 206–21. Oxford University Press.

Morrison, John (2016) 'Perceptual Confidence', *Analytic Philosophy* **57**: 15–48. doi: 10.1111/phib.12077

Munton, Jessi (2016) 'Visual Confidences and Direct Perceptual Justification', *Philosophical Topics* **44**, 301–26. doi:10.5840/philtopics201644225

Nanay, Bence (2020) 'Perceiving Indeterminately', *Thought: A Journal of Philosophy* **9**: 160–66. doi:10.1002/tht3.454

Pautz, Adam (2020) 'The Puzzle of the Laws of Appearance', *Philosophical Issues* **30**: 257–72. doi:10.1111/phis.12184

Raleigh, Thomas and Vindrola, Fillipo (2021) 'Perceptual Experience and Degrees of Belief', *The Philosophical Quarterly* **71**: 378–406. doi:10.1093/pq/pqaa047

Williams, Daniel (2020) 'Predictive Coding and Thought', *Synthese* **197**: 1749–75. doi:10.1007/s11229-018-1768-x