

---

# Folk Psychology and the Bayesian Brain

Joe Dewhurst

---

Whilst much has been said about the implications of predictive processing for our scientific understanding of cognition, there has been comparatively little discussion of how this new paradigm fits with our everyday understanding of the mind, i.e. folk psychology. This paper aims to assess the relationship between folk psychology and predictive processing, which will first require making a distinction between two ways of understanding folk psychology: as propositional attitude psychology and as a broader folk psychological discourse. It will be argued that folk psychology in this broader sense is compatible with predictive processing, despite the fact that there is an apparent incompatibility between predictive processing and a literalist interpretation of propositional attitude psychology. The distinction between these two kinds of folk psychology allows us to accept that our scientific usage of folk concepts requires revision, whilst rejecting the suggestion that we should eliminate folk psychology entirely.

In section 1 I will introduce predictive processing, giving a quick summary of the framework that focuses on the details most relevant for my comparison with folk psychology. I will also introduce folk psychology and define the distinction between propositional attitude psychology and folk psychological discourse. In section 2 I will consider the relationship between predictive processing and propositional attitude psychology, and in section 3 I will consider the relationship between predictive processing and folk psychological discourse. Finally, in section 4 I will argue that the distinction between propositional attitude psychology and folk psychological discourse makes space for us to revise our scientific usage of folk psychological concepts without thereby eliminating folk psychology altogether.

## 1 Introduction

This paper will focus on Andy Clark and Jakob Hohwy's presentations of the Bayesian brain hypothesis, which are not entirely identical, but I will note when they differ. I will use the term predictive processing to refer to their versions of this hypothesis. Other versions of the Bayesian brain hypothesis exist (see [Spratling in press](#), for an overview), and these differ from Clark and Hohwy's in many important ways, but I will not be discussing them here. Below I present a very brief introduction to predictive processing (see [Hohwy 2013](#), or [Clark 2016](#), for a more detailed overview). I will also introduce what I mean by folk psychology, which includes both the classical understanding of folk psychology as propositional attitude psychology, and a more general phenomenon that I call 'folk psychological discourse'.

### 1.1 Predictive Processing

Predictive processing inverts conventional assumptions about the flow of information in the brain. Rather than starting with raw perceptual inputs that are gradually processed into refined models of the world, it begins with a rich, internally generated model that predicts incoming sensory data. These predictions are then compared with the actual inputs, and the model is updated accordingly. Overall

### Keywords

Belief | Cognitive ontology | Desire | Eliminativism | Folk psychology | Instrumentalism | Predictive processing | Propositional attitude psychology | Realism | Revision

### Acknowledgements

I would like to thank Dan Calder, Jonny Lee, Suilin Lavelle, Dave Ward, Chris Michel, Alessio Bucci, Adrian Downey, and two anonymous reviewers, all of whom provided detailed comments on earlier drafts of this paper.

the system aims to minimise prediction error, which can be accomplished in two distinct ways. The model can be revised so as to more accurately predict incoming stimuli (passive inference), or the system can act on its environment in order to make its own predictions more accurate (active inference). Which kind of inference is performed (active or passive) will depend on higher-level predictions of the best way to reduce error in the current situation. Thus Clark summarises predictive processing as positing “core perception-attention-action loops in which internal models of the world and their associated precision expectations play key action driving roles” (Clark 2016, p. 71). By uniting action and perception in this way, predictive processing aims to provide a general account of cognition.

There are a few further features of predictive processing that are especially relevant to my discussion of folk psychology. Predictions can be regarded as more or less precise by the system, with the level of precision being taken into account when updating the predictions. For example, a less precise prediction will be expected to generate some error, and so may not need to be modified too much when an error signal is received. Changes in precision weighting can also drive the system to attend more or less to different sources of stimuli (Clark 2016, chapter 2). Finally, the predictive processing systems described by Hohwy and Clark are hierarchical; they consist of a nested hierarchy of precision/error units, with each level of the hierarchy predicting the current state of the unit below, which is then compared to the actual state of that unit and updated (in the next iteration) in response to any error signals that it receives. This hierarchy bottoms out in units that predict inputs received via sensory transduction, and tops out with a very abstract model, perhaps just predicting general causal laws or regularities. I will describe some further features of predictive processing in more detail as I go on to compare it with folk psychology.

## 1.2 Folk Psychology

In philosophy of mind and cognitive science, ‘folk psychology’ is typically taken to be synonymous with ‘propositional attitude psychology’, i.e. a theory of the mind based on a folk understanding of propositional attitudes such as belief and desire. According to this interpretation of folk psychology, our everyday understanding of how other people (and ourselves) think is constituted by a theory of how propositional attitudes interact with one another to cause behaviour, along with a capacity to attribute propositional attitudes to other people. A propositional attitude is an attitude, such as belief or desire, towards a proposition, such as ‘the sky is blue’. When I see you pick up your umbrella, I may attribute to you the belief that it is going to rain and the desire not to get wet, which together produce the behaviour of picking up your umbrella. One can be committed to this interpretation of folk psychology with or without being committed to the additional claim that it is a true account of how the mind actually works. Whilst Fodor (Fodor 1975) and Churchland (Churchland 1979, Churchland 1981) both agree that folk psychology consists in the attribution of propositional attitudes, they disagree about whether or not there are actually propositional attitudes ‘in the head’. Historically the debate between realists and eliminativists about folk psychology has shared a foundational assumption that folk psychology is equivalent to propositional attitude psychology, and that propositional attitude psychology aims to literally describe the structure of human cognition. For this reason I refer collectively to both realists and eliminativists about propositional attitude psychology as ‘literalists’. There are a number of alternative positions that one can take towards propositional attitude psychology, including Dennett’s intentional stance (Dennett 1989), Davidson’s anomalous monism (Yalowitz 2014), and various kinds of dualism or non-physicalism. In this paper I will be focusing primarily on the literalist interpretation of propositional attitude psychology, as it has the most immediate consequences for cognitive scientific practice, contrasted with the other interpretations that typically have less to say about what sorts of things should populate our scientific ontology. As such, in section 2 I will be assessing whether predictive processing is compatible with the existence of propositional attitudes ‘in the head’.

The philosophical literature on propositional attitudes has tended to focus fairly exclusively on beliefs and desires, but these are not the only propositional attitudes that could exist. As well as believing or desiring that it is raining, we could also hope that it is raining, or expect that it is raining, or suspect that it is raining, and so on indefinitely (cf. Stich 1983, p. 217). Stich suggests that what many theorists do at this point is simply to redefine ‘belief’ and ‘desire’ such that they cover any potential propositional attitude, but notes that this may just distort the terms so far that they cease to bear any resemblance to their folk usage (Stich 1983, p. 218). Like Stich, I will not take sides in this debate — what I have to say about belief and desire will generalise to other propositional attitudes.

Whilst it is typically equated with propositional attitude psychology, folk psychology is in fact a far more complex phenomenon, consisting of (at least) behavioural predictions, mental state attributions, narrative competency, and normative constraints (cf. Ratcliffe 2006). I will refer to this complex phenomenon collectively as ‘folk psychological discourse’, in order to capture the sense in which our everyday descriptions of behaviour consist of more than just theoretically motivated ascriptions of propositional attitudes. We also describe and predict behaviour in non-mentalistic terms, situate our descriptions and predictions in an on-going narrative structure, and make explicit normative judgments about how one ought to behave — where “ought” is understood as carrying both ethical and rational weight. The traditional characterisation of folk psychology as the attribution of propositional attitudes is only one component of this discourse. My claim that folk psychology ought to be understood in this broader sense is somewhat controversial, but I do not intend to defend it here. Those who want to restrict the use of the term ‘folk psychology’ to refer only to propositional attitude psychology might want to refer to these broader phenomena in some other way, such as ‘mentalising’ or ‘common sense psychology’. Section 3 will focus on the relationship between predictive processing and folk psychology in this broader sense.

Folk psychology in both senses is closely related to social cognition, i.e. the cognitive mechanisms that facilitate social interaction and interpersonal understanding. Quadt (Quadt 2017) discusses social cognition in the context of predictive processing, and I will briefly touch on how predictive processing could influence our understanding of social cognitive mechanisms, insofar as those are relevant to folk psychological discourse, but this is not the main aim of my paper.

## 2 Predictive Processing and Propositional Attitude Psychology

Both Clark and Hohwy have suggested in informal discussion that predictive processing might be incompatible with the folk psychological conception of cognition. Clark has described the content of the predictions as “alien” and “opaque”<sup>1</sup>, and Hohwy has acknowledged the challenge posed by predictive processing to “folk psychological notions of perception, belief, desire, decision (and much more)”<sup>2</sup>. Clark has also written that predictive processing “may one day deliver a better understanding even of our own agent-level experience than that afforded by the basic framework of ‘folk psychology’” (Clark 2013, p. 17, repeated in Clark 2016, p. 82). I take it that what both of them have in mind when they refer to folk psychology is propositional attitude psychology. This is the interpretation of folk psychology that has traditionally been of most interest to philosophers, and as such it is a good place to begin my assessment of predictive processing and folk psychology. In this section I will focus on belief and desire, although as I have already noted the issues raised here will generalize to other propositional attitudes.

1 Comment made during BPPA (British Postgraduate Philosophy Association) Masterclass on Action-Oriented Predictive Coding, University of Edinburgh, 26<sup>th</sup>-27<sup>th</sup> October 2013.

2 From the comments section of a Brains Blog featured scholar post: <http://philosophyofbrains.com/2014/06/22/is-prediction-error-minimization-all-there-is-to-the-mind.aspx>. See also (Hohwy 2013, p. 2) for a statement of how predictive processing might lead us to “radically reconceptualize who we are”.

## 2.1 Belief

According to the conventional account of propositional attitudes, a belief is a state consisting of a proposition coupled with a positive epistemic attitude, i.e. one that regards it as true, and it will interact with other mental states so as to generate actions in accordance with the state of affairs captured by the proposition being true. It is sometimes said that beliefs have a ‘mind-to-world’ direction of fit — that is to say, a belief should be modified in response to how the world is, and not vice versa.

On the face of it this kind of mental state seems to fit nicely into the predictive processing story. It is natural to interpret predictions as beliefs about the world, albeit ones that are first generated and then tested, rather than being generated in response to sensory input. Indeed, some researchers have described predictions as beliefs, including Friston (see e.g. [Hobson and Friston 2014](#)), Hohwy ([Hohwy 2012](#)), and occasionally Clark himself ([Clark 2016](#), p. 129). Given that the term ‘belief’ is used in a technical sense in Bayesian theory, this might be passed off as a harmless usage; however, simply equating predictions with beliefs in the everyday sense would be to ignore a crucial difference between folk psychological beliefs and the predictions invoked by the predictive processing story. The former are usually understood as determinate (you either believe something or you do not), whereas the latter are inherently probabilistic. Rather than simply believing that it is raining, a predictive processing system will assign a level of probability to it raining, and act in accordance with this probability. As Clark puts it,

Instead of simply representing ‘CAT ON MAT’, the probabilistic Bayesian brain will encode a conditional probability density function, reflecting the relative probability of this state of affairs (and any somewhat-supported alternatives) given the available information. ([Clark 2016](#), p. 41)

As a consequence of this, adopting the predictive processing framework will require either a re-conception of the relationship between folk psychological beliefs and the brain, or an acceptance that beliefs are in fact probabilistic rather than determinate. I explore the first option later on in this paper, where I will argue that it is a mistake to think that folk psychology has ever been in the business of describing the structure of cognition at the same level of detail as a cognitive scientific theory like predictive processing does. Something like the second option has been explored by Pettigrew ([Pettigrew 2015](#)), who considers the epistemological implications of adopting a probabilistic notion of belief alongside the more conventional determinate notion. In section 4 I will also consider the possibility that we should modify our understanding of ‘belief’ rather than eliminating it from our scientific ontology.

The predictions involved in predictive processing may also be individuated at a much finer level of detail than folk psychological belief attributions usually allow for. Whilst a paradigmatic belief might be about whether or not it is raining, the content of the predictions at some levels of the hierarchy are more likely to be cashed out in terms of fine-grained details of the external world, predicting features such as edges and light gradients rather than the ‘middle sized dry goods’ that populate the folk ontology. Even at higher levels of the hierarchy, the content of the predictions are still somewhat unusual, as they incorporate multi-modal, emotional, bodily, and other contextual associations. Combined with the probabilistic nature that I described above, the predictions posited by predictive processing begin to look less like the everyday notion of a belief. Clark expresses something like this view himself when he writes that the “looping complexities” involved in predictive processing “will make it hard (perhaps impossible) adequately to capture the contents or the cognitive roles of many key inner states and processes using the terms and vocabularies of ordinary daily speech” ([Clark 2016](#), p. 292).

However, Hohwy ([Hohwy 2013](#), p. 60) describes how the relationship between a predictive processing system and a dynamically evolving world could give rise to higher-level regularities that might come to resemble something more like folk psychological contents. He gives the example of perceiv-

ing a partially occluded cat, but forming a prediction of a whole cat based on feedback from seeing different parts of this cat at different points in time as it moves behind the occluder. The content of the whole-cat prediction is relatively coarse-grained, but it would in turn predict lower-level perceptions of parts of cats that change over time. The system is thus able to account both for the diachronic appearance of rapidly changing parts of cats, and the more abstract notion of a whole cat who stands behind the occluder and is temporally extended. So, if Hohwy is correct, we might expect to see something resembling folk psychological states at the higher levels of a predictive processing hierarchy, even if these states are different to how we usually conceive of them (i.e. their content is non-linguistic, abstract, and probabilistic, rather than consisting of linguistic propositions with determinate content).

Finally, it is important to recognise the dual nature of predictions. Predictions function both as representations of the world and of ways that the system can act in the world. Via the mechanism of active inference, predictions can be used to motivate and generate actions, a feature that is usually associated more with desires than beliefs. Clark likens this feature of high-level hypotheses to Millikan's (Millikan 1996) "pushmi-pullyu" representations, which have "both descriptive and imperative content" (Clark 2016, p. 187). At this point there is a sense in which beliefs, if they were to be identified with predictions, would begin to blur into what we might more naturally characterise as desires. Hohwy himself suggests that perception and belief might both be reconceived as a single notion of expectation (Hohwy 2013, p. 72), which could go some way towards reconciling predicting processing with propositional attitude psychology, although it would require that we adopt a revisionary approach towards folk psychology. Taken a step further this revisionary approach could also involve collapsing desire in with perception and belief, leaving us with a single kind of mental state that encompasses all aspects of cognitive processing.

Belief, understood as a positive epistemic attitude towards a proposition, does not straightforwardly fit in to the ontological framework of predictive processing. Whilst proponents of predictive processing have occasionally described predictions as beliefs, they have in mind something quite different to the traditional propositional attitude interpretation of folk psychology. Folk psychological beliefs are typically determinate and take linguistic propositions as their argument, whilst predictions are probabilistic and refer to a wide range of distinct contents, most of which are likely to be non-propositional. Nonetheless, it is plausible that we might find something closer to the folk psychological notion of belief at higher levels of the predictive processing hierarchy, where coarse-grained predictions about stable features of the environment are to be found. In the next section I will turn to desire, which in the context of action-oriented predictive processing is also constituted by predictions about the world.

## 2.2 Desire

Much like a belief, a desire consists of a proposition coupled with an attitude; only this time the attitude has a world-to-mind direction of fit, and will function accordingly. If I desire that it is raining, I will not pick up my umbrella (as I would if I believed it was raining), but I might sigh deeply and complain about the heat, or invest in experimental cloud seeding technologies.

As I mentioned above, the predictions posited by Clark and Hohwy's versions of predictive processing are action-oriented. This means that as well as providing a model of the world, they also serve to motivate the system to act via the mechanism of active inference. In this latter capacity they seem to fulfil a role very much like that played by desires in the traditional account of folk psychology. They represent how the system would like the world to be, and coupled with beliefs about the current state of the world, they generate the appropriate actions to bring about this desired state of affairs. Understood in this way we might conclude that it is viable to adopt a mild revisionism, where it turns out that beliefs and desires are both instantiated by a single kind of state, an 'action-oriented prediction'.

However, as Clark draws attention to, there is another sense in which predictive processing seems to do away with desire entirely. Friston et al. 2011 write, "crucially, active inference does not invoke any



‘desired consequences’” (Friston et al. 2011, p. 157), which Clark interprets as “a world in which value functions, costs, reward signals, and *perhaps even desires* have been replaced by complex interacting expectations that inform perception and entrain action” (Clark 2016, p. 129, emphasis added). The key issue here is that predictive processing inverts the conventional ordering of action causation assumed by folk psychology. Rather than a desire generating behaviour that leads to an expected outcome, a prediction of an expected outcome is generated first, which then goes on to cause behaviour that brings about that outcome. Desire seems to be relegated to a phenomenal sensation associated with this sequence of events, and does not seem to play any causal or functional role in generating either the behaviour or the outcome.

Clark argues that there need not be any contradiction here. Instead of eliminating desire from our ontology, we can reconceive of it as a consequence of the interaction between predictions and the environment (Clark 2016, p. 129). Insofar as it allows us to recognise the differences between predictive processing and folk psychology without simply eliminating the latter, I will go on to endorse something like this position, but first I want to mention a further issue that it raises. Reconceiving desire as a consequence rather than a cause of action has the potential for a deeply counterintuitive picture of personal level agency. Rather than being a distinct source of actions, agency (in the guise of active inference) turns out to be nothing more than a tool used by the system to minimise prediction errors.<sup>3</sup> We do not do things because we want to do them; we *feel* like we want to do things *because* doing them will minimise prediction error. As Hohwy puts it, “[w]hat drives action is prediction error minimisation [...] rather than what the agent wants to do” (Hohwy 2013, p. 89). Hohwy presents this as a positive result, unifying perception and action under one single mechanism (Hohwy 2013, p. 76), but for many this will seem like a sleight of hand, akin to Dennett’s attempts to reconcile free will with a deterministic cognitive architecture (see Dennett 1984, Dennett 2004). Perhaps this is just a symptom of a mistaken folk conception of agency, or perhaps it points towards confusion between two distinct modes of explanation — either way, the folk concept of desire would turn out not to be doing any significant work in our cognitive scientific explanations of action generation.

The folk psychological concept of desire as an action-motivating attitude is encompassed by the predictive processing notion of an action-oriented prediction, which via the mechanism of active inference is able to act on the world in order to make itself come true. Thus, predictive processing differs from the folk notion of desire in two crucial respects: firstly, beliefs and desires are implemented by a single kind of state, an action-oriented prediction; secondly, desire is relegated to a secondary status, as it is prediction-error minimization, rather than any personal goals of the system, that drive action. Hohwy presents this as a positive result, offering the possibility of unifying perception and action under one single mechanism. I think instead that what it indicates is that the project of trying to naturalise folk psychology by identifying propositional attitudes with the theoretical posits of our best cognitive science is a mistaken one, as it misconstrues the aim and purpose of folk psychology. In the next section I will consider how a broader interpretation of folk psychology fits with the predictive processing framework.

### 3 Predictive Processing and Folk Psychological Discourse

Propositional attitude psychology is only one small part of folk psychological discourse. In our everyday lives, we are often able to make predictions about how someone is likely to behave, both in the short and long term. Sometimes we may also attribute mental states when predicting behaviour, although more typically such ascriptions are used in an attempt to explain or justify either past or future behaviour. Another type of explanation or justification can be given by folk psychological narratives, which are distinct from either behavioural or mentalistic language. All three of the above kinds of folk psychological discourse can also be used to impose normative constraints on behaviour, i.e. by

<sup>3</sup> Colombo (Colombo 2017) makes a similar point, in the context of challenging the empirical foundations of the Humean theory of motivation.

using them in an imperative rather than descriptive mode. I will now describe each component of folk psychological discourse in more detail, before considering what impact the success of predictive processing would have on this discourse.

### 3.1 Folk Psychological Discourse

At a very basic level, we are able to avoid bumping into strangers on the street and can make use of physical cues to understand what someone is about to do, or what they expect us to do. We can improve on this basic capacity for behaviour prediction by engaging in explicit theorising. Rather than just predicting future behaviour on the basis of current behaviour, I can supplement my prediction with a model of the kind of situation that I am observing, and how it normally plays out. This is distinct from mental state attribution or propositional attitude psychology in that it may not require anything other than a simple grasp of behavioural regularities, something that even non-human animals without a full-blown theory of mind seem to be able to do (see e.g. [Rosati and Hare 2010](#)).

When I see my colleague stand up from her desk and head empty-handed towards the fridge that stands in the corner of our office I can safely predict that she will open it and take something out. Of course, predictions of this kind only go so far — without any additional information I probably couldn't predict what she was going to get out of the fridge, although once I knew what she had got out I could probably predict what she was going to do with it. The additional information required to predict what she might get out is precisely what is provided by the other components of folk psychological discourse. For example, if I had seen her put some chocolate in there earlier in the day, and if she had just told me that she fancied a snack, I might be able to successfully predict not only that she would open the door, but also that she might take out the chocolate, break off a piece, and eat it. If I was further aware of her kindly nature, and that she knew I liked chocolate, I might predict that she would offer me some as well. So behavioural predictions that go beyond very simple and immediate circumstances seem to typically require further, non-behavioural information.

There is a thin line between behavioural predictions of this kind and more complex attributions of mental states as hidden causes of behaviour. The latter are what have typically been emphasised in previous philosophical discussions of folk psychology, normally under the more specific guise of propositional attitude attributions. It is important to distinguish between mental state attributions in general and the particular case of propositional attitude attributions, if only to make room for the in-principle possibility that there could be non-propositional mental states attributed by the folk. It is extremely natural to supplement behavioural descriptions with mentalistic language, to the point where not doing so can in fact feel somewhat artificial. Consider again my prediction of what my colleague will do when she stands up from her desk and walks towards the fridge. Based only on behavioural assumptions, I can predict that she will open it and take something out, and perhaps even predict what she will take out if I saw her put something there earlier, or if she only ever uses the fridge to store one item, but my predictions immediately become much more powerful if I have access to mentalistic data. Now I can predict that she will take out some chocolate and eat, because I know that she is hungry, and that she believes there to be some chocolate in the fridge. I can also predict that she will offer me some, because I know that she is kind, and I know that she knows that I am hungry. This is only a very simple case, but the complexity begins to increase rapidly as we add in details, especially once we get to recursive attributions (“I know that she knows that...”).

Note that there are what appear to be two kinds of attributions in the above vignette, at least prior to further analysis. We have propositional attitude attributions: “she believes  $x$ ” and “she knows  $x$ ”. We also have something like emotional or dispositional attributions: “she is hungry” and “she is kind”. Whilst these can easily be reinterpreted as propositional attitude attributions — “she desires food” and “she wants to please others” — I think that doing so mischaracterises the nature of folk psychological discourse, as we typically interpret dispositional attributions as having a wider remit than proposi-

tional attitude attributions. “She is hungry” implies not only that she desires food, but also that she might be somewhat irritable, and that she might use food-related examples when making philosophical arguments, for instance.

Whilst I can predict my colleague’s behaviour by engaging in crude behaviourism or by attributing mental states, it is often easier to simply situate her actions in an on-going narrative structure, one that I have built up over the weeks, months, and years that I have known her (see Bruner 1990, and Hutto 2008). This narrative allows for predictions, as if I am familiar with the narrative then I know what comes next, but it also provides a contextual justification for her behaviour. Narrative competency is distinct from mental state attribution because it does not rely on any particular theory of how the mind works, or even an explicit awareness of other people as mental agents at all.

Folk psychology can also serve as a normative or regulative discourse. Morton (Morton 1980) first articulated the suggestion that folk psychology might be partially normative, and McGeer (McGeer 2007), Zawidzki (Zawidzki 2013), and Andrews (Andrews 2015) have all explored it more recently. Folk psychological discourse can be considered to be playing a regulative role whenever it causes us to adjust our behaviour in some way. Zawidzki lists several different forms that this can take, “including imitation, pedagogy, norm cognition and enforcement, and language based regulative frameworks, like self- and group-constituting narratives” (Zawidzki 2013, p. 29). Note that ‘mindshaping’, as Zawidzki calls it, spans the whole range of folk psychological discourse, from “self- and group-constituting narratives” right down to the basic, perhaps pre-folk psychological, imitation of the behaviour of conspecifics. One particularly interesting case that Zawidzki explores in some detail is the way in which our explicit attributions of mental states to ourselves and to others ends up becoming self-fulfilling prophecies, as we consequently feel social pressure to conform to those attributions and thus to maintain at least a kind of surface level consistency (Matthews 2013, p. 111, makes a similar suggestion). This reversal of the usual way in which folk psychology is thought to operate, from mindreading to mindshaping, resembles the relationship between passive and active inference under predictive processing accounts. An explicit mental state attribution that leads to conformative behaviour could be seen as a kind of active inference, whilst folk psychology understood as mindreading is more akin to a form of passive inference.

Folk psychology in the more complex sense that I have described above is less obviously threatened by predictive processing. Regardless of whether or not folk psychological discourse paints a literally correct picture of how the brain works, it is self-evidently capable of predicting behaviour across a wide range of everyday situations, and it would continue to be able to do this even if the way in which it reached these predictions appealed to an implicit theory of cognition that turned out to be false. It is not even clear that folk psychological discourse is primarily in the business of explaining cognition, as the likes of Fodor and Churchland seem to take it to be. Rather it can be characterised as being primarily in the business of predicting behaviour under normal conditions, much as folk physics is able to make successful predictions under a range of everyday circumstances despite making several false assumptions and failing to accurately explain the dynamics of physical systems (cf. Churchland 1979).

There is a potential concern here, that folk psychology’s predictive success might be sufficient reason to say that it does in fact give correct explanations of behaviour, and thus constitutes an accurate theory of cognition. According to some theorists in philosophy of science, the predictive success of a theory gives us good reason to be realists about that theory (see e.g. Putnam 1975; cf. Chakravartty 2015, sec. 2.1). If this were true then the predictive success of folk psychological discourse would be a reason to adopt a realist attitude towards it. I am sympathetic to this concern, but I think that it risks conflating two ways in which folk psychology could be successful. Rather than denying that folk psychological discourse describes real entities and processes, I want to deny that these entities and processes are of the sort that should serve an explanatory role in our scientific theories of cognition.



### 3.2 The Impact of Predictive Processing

With all that in mind, we can consider how well the predictive processing story aligns with folk psychological discourse in this more general sense. [Friston and Frith 2015](#) have explored behaviour reading and prediction, understood in the context of Bayesian inference. They argue that predicting the behaviour of another requires synchrony between the two brains in question (that of the predictor and the predicted), allowing predictions to be made based entirely on the current state of one's own brain. This sounds somewhat like the simulation theory in social cognition (see e.g. [Goldman 2006](#)), which claims that we understand other minds by analogy with our own mind, although [Friston and Frith 2015](#) do not make this connection. Given that predictive processing can be interpreted as generating a simulation of the target domain, this similarity is perhaps unsurprising, although there's also a sense in which the heavy emphasis on inference puts predictive processing closer to the theory-theory (see [Ravenscroft 2010](#), sec. 2.1), which posits a literal theory of how minds work as the main mechanism for social cognition. One possibility here is that adopting the predictive processing framework would contribute to the development of a hybrid theory that includes elements of both theory-theory and simulation theory. (See [Quadt 2017](#), for further discussion of predictive processing and social cognition.)

Narratives and social norms both seem to fit very comfortably into the predictive processing framework. Clark writes that individuals may “actively constrain their own behaviours so as to make themselves more easily predictable by other agents” ([Clark 2016](#), p. 286), a suggestion that fits very neatly into the mindshaping account of social cognition presented by Zawidzki ([Zawidzki 2013](#)). Clark also suggests that personal narratives might “function as high-level elements in the models that structure our own self-predictions, and thus inform our own future actions and choices” ([Clark 2016](#), p. 286). This is very close to the role envisioned for narratives in personal and social cognition by Hutto ([Hutto 2008](#)), and thus entirely consistent with my broader characterisation of folk psychological discourse. Hohwy ([Hohwy 2013](#), p. 163) describes how difficult even simple behavioural predictions can be, and suggests that a failure to take into account broader contextual features might help explain the social cognitive deficit found in people with autism. So understanding folk psychological discourse within the predictive processing framework might involve telling a story about how high-level predictions of behaviour will involve complex models spanning not only individual agents but also the social and cultural environments that contribute to their behaviour.

The predictive processing account of cognition might even be reliant upon folk psychological narratives and social interaction more generally. Clark has written elsewhere about the importance of niche construction and cognitive scaffolding for human cognition ([Clark 2008](#)), and he devotes a chapter of his book on predictive processing to discussing these issues ([Clark 2016](#), chapter 8). By conceiving of folk psychological discourse as a form of cognitive scaffolding we can retain an important space for it in our explanations of cognition, even if folk psychological explanations themselves are sometimes hard to reconcile with predictive processing. For example, by helping to regulate human behaviour via the enforcement of social norms, folk psychological discourse might serve as a form of active inference, changing the social environment so as to make it easier to predict. It also provides shared narratives that can help make sense of the behaviour of others, as well as exerting a regulative influence in their own right (see [Andrews 2015](#)). So even if propositional attitude psychology turns out to be a bad model of the cognitive architecture required for predictive processing, it might continue to be pragmatically useful to conceive of people as the kinds of systems that have beliefs and desires. We are then left with the further question of whether the failure of folk psychological discourse to match up precisely to our current best theories of cognition gives us reason to eliminate it, or whether being pragmatically useful is enough to escape this fate.

## 4 Revision Without Elimination

Predictive processing seems to be at least partially incompatible with the model of cognition given by propositional attitude psychology. The action-oriented nature of predictions means that it is hard to maintain a principled distinction between beliefs and desires, and the fine-grained content of some levels of the predictive processing hierarchy means that the content of those predictions will be hard to describe in everyday terms. One response to this partial incompatibility, originally offered by Lycan (Lycan 1988) in response to Stich's (Stich 1983) eliminativism, is that we should simply reinterpret folk psychological terms such as belief and desire in line with empirical discoveries. The motivation for this response is given by the causal theory of reference advocated by Putnam and Kripke, under which the reference of a term is determined by its initial use, rather by a definite description. So it would turn out that propositional attitude psychology was referring to fine-grained, action-oriented predictions all along, and that the folk were simply confused about what it was that they were trying to refer to when they attributed beliefs and desires to each other.

Without wanting to enter into the details of this debate, I am perfectly happy to concede that this is a viable option when presented with evidence of incompatibility between a folk theory and a scientific theory. However, as I have tried to argue, applying this argument to folk psychology misconstrues the nature of what the folk are trying to do in the first place. Propositional attitudes were never meant to refer to fine-grained mental states, but are instead intended to pick out and predict coarse-grained behavioural patterns and dispositions. I should emphasise that I am not advocating an anti-realism with regard to the posits of propositional attitude psychology — I think that whilst they are perfectly real, they just aren't the kind of thing that we should look for inside people's heads. This position is not so different from that advocated by Dennett 1989, and seems to more accurately capture the nature of our everyday understanding of other people.

Accepting that there is an incompatibility between folk psychology and cognitive science does not necessitate any radical change to our everyday folk psychological practices. Much of folk psychological discourse, such as behavioural predictions, normative constraints, and narrative competency, need not give an entirely accurate picture of the mechanisms of cognition in order to contribute to our understanding of other people. Whilst the literalist interpretation of folk psychology as propositional attitude psychology is at least partially incompatible with predictive processing, this is only a problem for those (such as Fodor) who insist that our theories of cognition should be structured along the same lines as propositional attitude psychology. Similarly, the eliminative materialism advocated by Churchland and others relies upon a commitment to folk psychology *either* accurately describing the structure of cognition *or* being eliminated. Without this dichotomy we can accept that our scientific usage of folk psychological terms should be revised without feeling at all compelled to eliminate it from everyday usage.

Nonetheless, it is plausible that a popularised version of predictive processing could partially modify our folk psychological self-conception, in ways that could be hard to predict ahead of time. Consider, for instance, the impact that the psychoanalytic notion of the unconscious had on the popular conception of the mind (Richards 2000). Prior to the 20<sup>th</sup> century the idea that our behaviours might be guided by unconscious motivations was not widely held at all, but now it has become common to refer to 'Freudian slips' and so on in everyday descriptions of behaviour. Similarly, if predictive processing catches on we might one day find ourselves referring to high-level predictions in order to explain why someone mistook a disguised mule for a zebra. This would be an entirely natural process that is (mostly) outside of the control of philosophers and cognitive scientists, so unless we thought that there was some ethical imperative to change the way that people think of themselves,<sup>4</sup> we should not let this concern us.

<sup>4</sup> Michael L. Anderson suggested that someone like George Lakoff might think that we do have an ethical imperative of this sort – I am content to leave this as a question for the ethicists.

Quite aside from threatening to eliminate folk psychology, predictive processing presents us with an opportunity to revise our scientific usage of folk psychological concepts. We should aim to develop a novel conceptual taxonomy that more accurately reflects the structure of cognition and allows us to move beyond the limitations of folk psychological discourse. Understood in this way folk psychology could be used to identify interesting target phenomena and inspire scientific research (cf. [Turner 2012](#)), but should not be used as a source of technical cognitive scientific concepts. Some suggestions as to how to accomplish this transition can be found in recent literature on cognitive ontology revision (see e.g. [Price and Friston 2005](#), [Klein 2012](#), [Poldrack 2006](#), [Poldrack 2010](#), and [Anderson 2014](#), [Anderson 2015](#)). Here the basic idea is that by comparing data from a large number of studies we can work out what the functionally relevant states and processes are, and design new terminology that better reflects these states and processes. In the case of predictive processing this might mean acknowledging the bidirectional nature of predictions, which in a sense serve both as beliefs and desires, and identifying the relevant differences between the way that these states interact with one another and the way that propositional attitudes are thought to interact with one another. I will not discuss the details of this proposal any further in this paper, although I do think that a concerted effort will be required in this regard if we are to move beyond the conceptual taxonomy offered by folk psychology.

## 5 Conclusion

Hohwy and Clark are both correct to suggest that folk psychology, interpreted as propositional attitude psychology, is likely to turn out to be incompatible with the account of cognition presented by predictive processing. As a consequence of this incompatibility the adoption of predictive processing will require the development of a novel conceptual taxonomy that more accurately describes the states and processes involved in prediction error minimisation. Whilst it might be okay to continue referring to predictions as ‘beliefs’ when one is clear that this is a technical usage, it would be better to use a new term so as to avoid confusion with the folk psychological sense of ‘belief’ (much as the term ‘surprisal’ has been adopted to avoid confusion with the common-sense understanding of ‘surprise’). One option would be to use only the term ‘predictions’, although even here there is an implication of personal level agency that could perhaps be misleading. Better yet would be to coin an entirely new term that does not carry any unwanted associations.

Regardless of how we go about revising our conceptual taxonomy, folk psychology will not be threatened with elimination. This is because the traditional interpretation of folk psychology as propositional attitude psychology mischaracterises the nature of folk psychological discourse, which is actually a far more complex phenomenon consisting of behavioural predictions, narrative competency, and normative constraints, in addition to mental state attributions. Whilst a literalist reading of propositional attitude attributions is incompatible with predictive processing, there is no incompatibility between predictive processing and folk psychological discourse as a whole. If predictive processing turns out to be a successful theory of cognition (which is largely an empirical question), it may consequently have some impact on how we conceive of ourselves, and thus indirectly influence folk psychological discourse. The extent of this influence, however, is hard to predict, and in any case is not the sort of thing that philosophers or cognitive scientists should be in the business of attempting to regulate. We should content ourselves with ensuring that the language used to describe predictive processing is clear and accurate, and that the framework itself is coherent and supported by empirical evidence.

## References

- Anderson, M. L. (2014). *After phrenology*. Cambridge, MA: MIT Press.
- (2015). Mining the brain for a new taxonomy of the mind. *Philosophy Compass*, 10(1), 68-77.
- Andrews, K. (2015). The folk psychological spiral. *Southern Journal of Philosophy*, 53, 50-67.
- Bruner, J. S. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Chakravartty, A. (2015). Scientific realism. In E. N. Zalta (Ed.) *The Stanford encyclopedia of philosophy*.
- Churchland, P. M. (1979). *Scientific realism and the plasticity of mind*. Cambridge: Cambridge University Press.
- (1981). Eliminative materialism and the propositional attitudes. *The Journal of Philosophy*, 78(2), 67-90.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.
- (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(03), 181-204.
- (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. New York: Oxford University Press.
- Colombo, M. (2017). Social motivation in computational neuroscience: Or if brains are prediction machines then the Humean theory of motivation is false. In J. Kiverstein (Ed.) *Routledge handbook of philosophy of the social mind*. Abingdon, OX / New York, NY: Routledge.
- Dennett, D. C. (1984). *Elbow room: The varieties of free will worth wanting*. Cambridge, MA: MIT Press.
- (1989). *The intentional stance*. Cambridge, MA: MIT press.
- (2004). *Freedom evolves*. London: Penguin UK.
- Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Friston, K. & Frith, C. (2015). A duet for one. *Consciousness and Cognition*, 36, 390-405.
- Friston, K. Mattout, J. & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, 104(1-2), 137-160.
- Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. New York: Oxford University Press.
- Hobson, J. A. & Friston, K. J. (2014). Consciousness, dreams, and inference: The Cartesian theatre revisited. *Journal of Consciousness Studies*, 21(1-2), 6-32.
- Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in Psychology*, 3. <https://dx.doi.org/10.3389/fpsyg.2012.00096>.
- (2013). *The predictive mind*. Oxford: Oxford University Press.
- Hutto, D. D. (2008). *Folk psychological narratives: The socio-cultural basis of understanding reasons*. Cambridge, MA: MIT Press.
- Klein, C. (2012). Cognitive ontology and region-versus network-oriented analyses. *Philosophy of Science*, 79(5), 952-960.
- Lycan, W. G. (1988). *Judgement and justification*. Cambridge / New York / Melbourne: Cambridge University Press.
- Matthews, R. J. (2013). Belief and belief's penumbra. *New essays on belief* (pp. 100-123). Springer.
- McGeer, V. (2007). The regulative dimension of folk psychology. In D. D. Hutto & M. M. Ratcliffe (Eds.) *Folk psychology re-assessed* (pp. 137-156). Dordrecht, NL: Springer.
- Millikan, R. (1996). Pushmi-pullyu representations. *Philosophical Perspectives*, 9, 185-200.
- Morton, A. (1980). *Frames of mind: Constraints on the common-sense conception of the mental*. Oxford: Oxford University Press.
- Pettigrew, R. (2015). Pluralism About Belief States. *Aristotelian Society Supplementary Volume*, 89(1), 187-204.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10(2), 59-63.
- (2010). Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on Psychological Science*, 5(6), 753-761.
- Price, C. J. & Friston, K. J. (2005). Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology*, 22(3-4), 262-275.
- Putnam, H. (1975). *Mathematics, matter, and method*. London / New York: Cambridge University Press.
- Quadt, L. (2017). Action-oriented predictive processing and social cognition. In T. Metzinger & W. Wiese (Eds.) *Philosophy and predictive processing*. Frankfurt am Main: MIND Group.
- Ratcliffe, M. (2006). 'Folk psychology' is not folk psychology. *Phenomenology and the Cognitive Sciences*, 5, 31-52.
- Ravenscroft, I. (2010). Folk psychology as a theory. In E. N. Zalta (Ed.) *The Stanford encyclopedia of philosophy*.



- Richards, G. (2000). Britain on the couch: The popularization of psychoanalysis in Britain 1918 – 1940. *Science in Context*, 13(02), 183-230.
- Rosati, A. G. & Hare, B. (2010). Social cognition: From behavior-reading to mind-reading. In G. F. Koob, M. Le Moal & R. F. Thompson (Eds.) *Encyclopedia of behavioral neuroscience*. London, UK / Burlington, MA / San Diego, CA: Academic Press.
- Spratling, M. W. (in press). A review of predictive coding algorithms. *Brain and Cognition*. <https://dx.doi.org/10.1016/j.bandc.2015.11.003>.
- Stich, S. P. (1983). *From folk psychology to cognitive science: The case against belief*. Cambridge, MA: MIT Press.
- Turner, R. (2012). The need for systematic ethnopsychology: The ontological status of mentalistic terminology. *Anthropological Theory*, 12(1), 29-42.
- Yalowitz, S. (2014). Anomalous monism. In E. N. Zalta (Ed.) *The Stanford encyclopedia of philosophy*.
- Zawidzki, T. W. (2013). *Mindshaping: A new framework for understanding human social cognition*. Cambridge, MA: MIT Press.