

Folk psychological and neurocognitive ontologies

Joe Dewhurst (joseph.e.dewhurst)
Munich Center for Mathematical Philosophy

Abstract

It is becoming increasingly clear that our folk psychological ontology of the mental is unlikely to map neatly on to the functional organisation of the brain, leading to the development of novel ‘cognitive ontologies’ that aim to better describe this organisation. While the debate over which of these ontologies to adopt is still ongoing, we ought to think carefully about what the consequences for folk psychology might be. One option would be to endorse a new form of eliminative materialism, replacing the old folk psychological ontology with a novel neurocognitive ontology. This approach assumes a literalist attitude towards folk psychology, where the folk psychological and neurocognitive ontologies represent competing and incompatible ways of categorising the mental. According to an alternative approach, folk psychology aims to describe coarse-grained behaviour rather than fine-grained mechanisms, and the two kinds of ontology are better thought of as having different aims and purposes. In this chapter I will argue that the latter (coarse-grained) approach is a better way to make sense of everyday folk psychological practice, and also offers a more constructive way to understand the relationship between folk psychological and neurocognitive ontologies. The folk psychological ontology of the mental might not be appropriate for describing the functional organisation of the brain, but rather than eliminating or revising it, we should instead recognise that it has a very different aim and purpose than neurocognitive ontologies.

Introduction

This chapter will introduce the threat posed to folk psychology by novel neurocognitive ontologies and respond to this threat by arguing that we should adopt a coarse-grained understanding of folk psychology. Rather than conceiving of folk psychology as aiming to literally describe the fine-grained structure of (neuro)cognition, we ought to understand it as aiming to interpret, predict, and explain the behaviour of whole persons by attributing dispositional states to them. Adopting this conception of folk psychology will insulate it from any threat posed by cognitive ontology revision, and offer a more appealing picture of the relationship between folk psychology and cognitive neuroscience.

In section 1 I introduce what I mean by folk psychology, and in section 2 I describe the ongoing ‘cognitive ontology’ debate, which has arisen due to the failure of our current mental categories to map neatly onto the structure of the brain. A common kind of response to this problem is to argue that we should adopt a novel cognitive ontology, i.e. a taxonomy of cognitive states that is better suited to the apparent functional architecture of the brain. In section 3 I will describe how the adoption of a novel cognitive ontology might threaten folk psychology with a new form of eliminative materialism, based on the assumption that folk psychology aims to literally describe the same kinds of functions that cognitive neuroscience is interested in investigating. In section 4 I will argue that we can avoid this threat by adopting an alternative coarse-grained approach, which conceives of folk psychology as aiming to describe the behaviour of whole persons, rather than the underlying neural mechanisms that generate this behaviour. Finally, in section 5 I will consider the implications that adopting this approach would have for our understanding of the relationship between folk psychology and cognitive neuroscience. The real lesson we should take from cognitive ontology revision is not that folk psychological and neurocognitive ontologies are incompatible, but rather that they are best applied in different domains, and need not compete with one another.

1. Folk psychology and neuroscience

In the philosophical literature, ‘folk psychology’ is often conflated with ‘propositional attitude psychology’, and it is under this guise that the traditional debate about eliminative materialism has taken place. However, more recent work on folk psychology has drawn attention to the many other ways in which we might understand one another, including not only propositional attitudes but also other kinds of mental states (such as emotions or non-propositional attitudes), character traits, narratives, and normative constraints (see Spaulding 2018 for a general overview of these developments). In this section I will introduce both conceptions of folk psychology and the cognitive ontologies that they might support, and then consider how these could have influenced neuroscientific ontologies. This will set the scene for the next section, where I will discuss a recent debate about whether, and to what extent, our existing ‘cognitive ontology’ requires revision.

The term ‘folk psychology’ only became popular in the philosophical literature from around the 1980s onwards, following Paul and Patricia Churchland’s arguments that it constitutes a primitive and largely unsuccessful theory of how the mind works, one that ought to be replaced with a new theory drawn from “the conceptual framework of a completed neuroscience” (P.M. Churchland 1981: 67; see also his 1979 and P.S. Churchland 1986). This “eliminative materialism” stood in contrast with Fodor’s defense of folk psychology as a necessary framework for understanding the mind, for which he thought that we have no conceivable alternative (Fodor 1987: 132). Both Fodor and the Churchlands followed Lewis’ earlier characterisation of common-sense psychology (as Lewis called it) as a proto-scientific theory. Lewis argued that our everyday language for talking about the mind could be treated “as a term-introducing scientific theory” (1972: 256), such that we could simply read off an ontology of mental states from the way that we talk about the mind.¹ Here Lewis focused primarily on the ‘propositional attitudes’, i.e. attitudes such as belief and desire that one can hold towards a proposition, and Fodor (1975) followed this approach in constructing his ‘language of thought hypothesis’, according to which cognition consists in the manipulation of folk-psychologically characterised propositional attitudes. A similar emphasis is found in the ‘theory-theory’ in social cognition, which argues that our understanding of other minds is guided by an implicit ‘theory of mind’, although not necessarily one identical to the language of thought hypothesis (see Gopnik & Wellman 1994; cf. Premack & Woodruff 1978, who first introduced the term ‘theory of mind’ to scientific psychology). For example, according to this theory I might come to attribute to you a belief about the location of some object based on theoretical inferences informed by your behaviour (looking in a certain place and seeming surprised, etc.) The theory-theory is based on a literal understanding of the common-sense propositional attitude theory invoked by Lewis, i.e. a very particular (and somewhat peculiar) philosophical interpretation of a much broader cultural practice of self- and other understanding.

It was this approach that framed the original debate over eliminative materialism in the 1980s, which took it for granted that folk psychology was in the business of attributing propositional attitudes to people in a proto-theoretical manner, such that it could be understood as a literally true or false theory, amenable to scientific investigation. The eliminativists, such as Paul and Patricia Churchland,² argued that our best neuroscience would demonstrate that this theory was false, whereas the realists denied that this could be possible, in Fodor’s case going so far as to argue that folk psychology (and

¹ Lewis explicitly denied that folk psychology originated as a theory of this kind, but rather followed Sellars (1956) in treating it as a “good myth” that might help us better understand the mind (Lewis 1972: 257).

² The other most notable eliminativist is Stich (1983), who later repudiated his version of the view due to concerns raised by Lycan (1988) about the reference of folk psychological terms. Eliminativism also has historical antecedents in Feyerabend (1963) and Rorty (1965). Stich did acknowledge that folk psychology might be broader than just propositional attitude psychology, but glossed over this by suggesting all non-propositional mental terms could simply be restated in a propositional format (1983: 217).

psychology more generally) was autonomous from neuroscience in a way that shielded it from empirical refutation at this level of analysis (Fodor 1974). Nonetheless, both the eliminativists and the realists took it for granted that folk psychology was trying to literally explain how cognition functions in a relatively fine-grained manner. In section 3 I will argue that current debates over cognitive ontology revision pose a novel eliminativist threat, in a similar manner to the original eliminative materialism of the 1980s.

Subsequent work on folk psychology and social cognition has recognised that our common-sense understanding of other minds might not just consist in the attribution of propositional attitudes (see Lavelle 2019 for a general introduction to the topics discussed here). In social cognition, there has been an increased emphasis on non-theoretical means of understanding one another, such as simulation (Gordon 1986, Heal 1986), direct perception (Gallagher 2008a), and interaction (De Jaegher & Di Paolo 2007, Gallagher 2008b). Whether these are truly distinct from the theory-theory is a complicated question (see e.g. Lavelle 2012), and more recently there has been a shift towards endorsing some version of a hybrid theory that acknowledges the role of both theory and simulation in social cognition (see e.g. Mitchell 2005, Apperly 2008). There has also been a related shift away from focusing on just propositional attitude attribution and towards seeing folk psychology as a multifaceted phenomenon, consisting not just of propositional attitude attribution but also other means of understanding one another, such as character traits (Westra 2018), narrative structure (Bruner 1990, Hutto 2008), and normative constraints (Mameli 2001, McGeer 2007, Zawidzki 2013, Andrews 2015). This broader understanding of folk psychology, I will argue, might give us the resources to reconceive of the relationship between folk psychology and neuroscience in a way that avoids the threat of eliminative materialism posed by cognitive ontology revision.

Whether we take a broad or narrow view of folk psychology, we can ask what kind of ontology of mental states it provides us with, and thus what kind of theory of mind and cognition it entails. The ontological commitments of pre-theoretic folk psychology are at the very least unclear, and perhaps even simply indeterminate, but philosophers (and cognitive scientists) have nonetheless tried to interpret and ‘clarify’ them (and what we take folk psychology to be ‘literally’ committed to will depend on how this interpretation is carried out). Lewis proposed reading off a set of theoretical commitments from the “everyday platitudes” of common-sense psychology (1972: 252), which in contemporary philosophy of mind is often assumed to be reducible to some version of belief-desire psychology. A similar approach is reflected in Fodor’s language of thought hypothesis, which expects the structure of cognition to match up (in some sense) with our everyday language for talking about the mind (understood by Fodor in narrow terms, i.e. propositional attitude attribution). It is important to note here that Fodor does *not* expect the reverse to be true, i.e. that folk psychology should conform to whatever our best scientific theory of mind and cognition is, but rather just that the folk theory is likely to be roughly correct in the first place.³

Classical eliminativism shares the realist assumption that we can simply read off an ontology of mental states from folk psychology, and then check whether this matches up with the empirical discoveries of our best cognitive neuroscience. A crucial difference is whether or not one expects this ontology to match up with the findings of neuroscience (in the case of the eliminativist) or with some more abstract psychological theory (in the case of the Fodorian realist), but the commitment to a ‘literal’ interpretation of folk psychology is apparent in both cases. Understood more broadly, folk psychology could be interpreted to support an ontology consisting of not just propositional attitudes, but also emotions, character traits, and perhaps even roles in a social narrative. There is then a further

³ This means that Fodor’s position, *qua* revisions to our folk ontology, is actually quite similar to that which I will present in sections 4 and 5 of this chapter. I thank J. Brendan Ritchie for pressing me on this point.

question of what kind of relationship this ontology has with the various cognitive scientific disciplines, including neuroscience.

Fodor saw folk psychology as being a precursor to our scientific *psychological* ontology, which he thought was wholly autonomous from neuroscience, whereas the eliminativists thought that our folk psychological ontology ought to be judged against our best neuroscience and revised or eliminated if it failed to match up.⁴ The reality is probably somewhat more complex. On the one hand, it seems increasingly implausible that psychology could be wholly autonomous from neuroscience (see e.g. Boone & Piccinini 2016; cf. Piccinini & Craver 2011; Knoll 2018 for a dissenting opinion), meaning that revisions to our neuroscientific ontology might also entail revisions to our (folk) psychological ontology (see section 3). On the other hand, the move away from a narrow understanding of folk psychology as *just* propositional attitude psychology means that we can now conceive of a more sophisticated relationship between folk psychology and neuroscience than mere one-to-one mapping.

In sections 4 and 5 I will argue that the dichotomy between folk psychological realism and eliminativism rests on the mistaken assumption that we should take folk psychology *literally*, i.e. understand it as being involved in the same kind of project as our scientific investigation of the mind and brain. We should adjust our perspective and reconceive of folk psychology as being in the business of interpreting the coarse-grained behaviour of whole persons, rather than the fine-grained mechanisms that generate that behaviour. From this alternative perspective it turns out that folk psychological and neuroscientific ontologies have such different aims, methods, and standards that it would be a mistake to directly compare them. This is not a new proposal, having antecedents in the idea that the application of folk psychological concepts within the context of neuroscience might constitute a kind of category mistake (see e.g. Bennett & Hacker 2003; cf. Ryle 1949). The novelty of my argument here is firstly in applying this idea to the specific case of cognitive ontology revision, and secondly in providing a distinctive kind of rationale for taking this approach, based not so much on linguistic or grammatical concerns, but rather on concerns to do with the nature of folk psychology itself, which seems more concerned with interpreting the behaviour of whole persons than with identifying the neural mechanisms responsible for that behaviour.

2. Ontology revision in cognitive neuroscience

Aspects of our folk psychological ontology have historically influenced cognitive neuroscientific ontologies, even if sometimes only in a subtle and indirect manner. Many basic psychological concepts such as memory, attention, and belief originate in folk psychology, and although most of these concepts have undergone technical revision over the years, they still bear some traces of their folk psychological origins. When designing neuroimaging studies it is necessary to define a task that is intended to operationalise the cognitive function that you are attempting to investigate, and most of these tasks are broadly folk psychological in flavour, even if the functions they are intended to track are often more precisely defined. For example, a study investigating the neural correlates of written language processing might deploy a reading task, where ‘reading’ is understood as a single kind of cognitive function that is expected to map neatly onto a region of the brain. This kind of influence

⁴ As noted above, Stich (1983) also endorsed a form of eliminativism, but he later realised that if one adopts a causal theory of reference then changes to the scientific ontology might instead give us reason to *revise* (rather than *eliminate*) the folk psychological ontology (see e.g. Stich 1996; cf. Lycan 1988). As I will argue in sections 4 and 5, I think this move misunderstands the relationship between scientific and folk ontologies in just the same way that (folk psychological) eliminativism does. This debate about “arguments from reference” (Mallon *et al* 2009) dominated much philosophical discussion of folk psychology in the 1990s and 2000s, and I hope to bypass it entirely here by focusing more on the practical differences between scientific and folk ontologies.

would be innocuous if it turned out that such categories were in fact appropriate for neuroscience, but as we will now see it is becoming increasingly clear that this might not be the case

The term ‘cognitive ontology’ was coined by Price & Friston (2005), who use it to refer to the set of cognitive functions that we appeal to when conducting neuroimaging studies.⁵ Ideally, they claim, this ontology should support a one-to-one mapping between functions and structures, such that “structures predict functions and functions predict structures” (Price & Friston 2005: 263). Each cognitive function (identified by getting a subject to perform a related task) should be correlated with activation in just one neural structure, and each structure should be implicated in just one kind of task or function. This would allow us to make clear statements about where in the brain each function is performed, vindicating the accuracy of our cognitive ontology (and, if the two were identical, also our folk psychological ontology). There is perhaps an (implicit) assumption of mind/brain identity underlying this approach, and more generally the approaches to cognitive ontology discussed in this section, although interpreted cautiously their aim is only to establish *correlations* between structures and functions, not identity relations.⁶ However, even under this more cautious interpretation, it is still assumed that there ought to be a correlation between cognitive functions and structures of the *brain*, rather than of the brain-and-body, or brain-body-and-world, or some other set of physical structures.

Unfortunately, it turns out that the cognitive ontologies applied in most neuroimaging studies do not support correlations of this kind. Typically we find cases of one-to-many mappings (where a single function appears to activate many structures), many-to-one mappings (where a single structure is implicated in many functions), and many-to-many mappings (where many different functions simultaneously cross-correlate with many different structures). Price & Friston see this as a problem, and one of the aims of their paper was to develop a way to revise our cognitive ontology in order to make it better match up with the structure of the brain. Their proposal is that we should develop a novel ontology by grouping together seemingly distinct functions that have similar activation profiles, coming up with more general labels for these new functions that captures their performance across different kinds of task. This would allow us to preserve one-to-one mapping at the expense of our original ontology, which would become subsumed under the new, more general functional categories.

To illustrate this approach, they focus on one example: the different kinds of function currently attributed to the left posterior lateral fusiform (LPLF). These include processing visual information about written words in reading tasks (Cohen et al 2000); processing the visual attributes of animals in semantic categorisation tasks (Martin & Chao 2001); and processing visual/tactile information more generally (Amedi et al 2002). The result is a case of many-to-one mapping, where a single structure (the LPLF) supports at least three different kinds of functional attribution. In order to avoid this, Price & Friston suggest reclassifying the function of the LPLF as ‘sensorimotor integration’, which they claim is able to accommodate each of the subsidiary functions attributed to it in different kinds of task. Their approach has since been criticised somewhat in the philosophical literature, with a common response being that ‘sensorimotor’ integration is just too broad a functional category to explain anything, and that we should instead attribute functions in a task or context sensitive manner (see e.g. Klein 2012, McCaffrey 2015, Burnston 2016). While Price & Friston acknowledge that there is a practical benefit to attributing more specific functions in the context of particular tasks, they still think it is beneficial to have a general functional category that preserves one-to-one mapping, such as ‘sensorimotor’ integration, because “it is more useful to label a region

⁵ They also refer to it as a ‘functional ontology’, but ‘cognitive ontology’ seems to be the terminology that is now used most commonly in the literature.

⁶ See Towl (2011) and Nathan (this volume) for further discussion, and Vernazzani (this volume) for a historical perspective.

with a function that explains all patterns of activation” (Price & Friston 2005: 268).⁷ Insofar as their motivation here is primarily pragmatic, it could be seen as an example of McCauley & Bechtel’s (2001) ‘heuristic identity theory’, which conceives of proposed “psycho-neural identities” as tools for generating new hypotheses and guiding experimentation in a manner that is fully compatible with a pluralistic cognitive ontology. However, for my purposes it is the general strategy and framing of the problem that is important, not the specific details of this case, and even the context sensitive mapping strategies will end up having counterintuitive consequences for our folk psychological ontology (which I will discuss in more detail in the next section).

Since Price & Friston first identified this problem, there have been many different proposals for how to resolve it, which can be broadly classified as ‘top-down’ (holding fixed our cognitive ontology and revising our understanding of neural structure) and ‘bottom-up’ (holding fixed our understanding of neural structure and revising our cognitive ontology). My focus here will be on the latter kind of approach, which if adopted would have the most significant impact on our folk psychological ontology (see Machery & McCaffrey 2016 for some general criticism of this kind of approach). In the rest of this section I will introduce two further bottom-up strategies for cognitive ontology revisions, each of which would threaten our folk psychological ontology in quite different ways. The first of these, advocated for by Russ Poldrack and colleagues, follows Price & Friston in aiming to preserve one-to-one mapping, while the second, developed by Michael Anderson, takes a more flexible approach based on the phenomenon of neural reuse, but nonetheless ends up with something very different to our current ontology.

Poldrack has proposed (and initiated) the development of a ‘Cognitive Atlas’, which aims to develop “a comprehensive, formally specified ontology of mental processes” (2010: 756), better suited for mapping to the structural organisation of the brain. This takes the form of an online database where different labs can upload their experimental protocols and results (www.cognitiveatlas.org), which can then be compared and analysed using data mining techniques. Poldrack & Yarkoni (2016) describe this approach in more detail, arguing that large-scale analyses of neuroimaging data can be used to overcome several challenges facing cognitive neuroscience, and emphasizing the role that “formal cognitive ontologies” can play in this process. They note that “all else being equal, we believe that a model of psychological processes that also maps systematically onto known biological structures is strongly preferable over one that does not” (*ibid*: 599); i.e., they give priority to biological or structural factors over functional or task-specific factors when determining their ontology.

For example, Lenartowicz *et al* (2010) analysed neuroimaging results associated with the construct ‘cognitive control’, finding five relevant key terms: working memory, response selection, response inhibition, task switching, and cognitive control itself (*ibid*: 682). In order to check which, if any, of these terms correspond uniquely to a neural structure, they performed a meta-analysis of neuroimaging studies in which they occur, and discovered that while there was a clear distinction between the patterns of activation corresponding to response selection on the one hand, and between working memory, response inhibition and cognitive control on the other, there was no clear distinction between the tasks associated with the latter group (the data corresponding to task switching was unclear). Based on this analysis they conclude that response selection is a distinct function associated with the precentral gyrus and middle frontal gyrus, and that cognitive control, response inhibition and working memory may together constitute a second distinct function associated with “a

⁷ An alternative kind of response that I do not have space to consider here is to develop an ontology based on the evolutionary origins of these neural structures. Barrett (2012) proposes that the LPLF should be understood as performing “category specific object recognition”, a functional attribution that he argues can accommodate the different kinds of task that this region is correlated with (see Rathkopf, this volume, for further discussion of this kind of evolutionary approach).

right-lateralized network involving frontal and subcortical regions” (Lenartowicz *et al* 2010: 688). They acknowledge that their data was somewhat noisy, and so do not present these results as conclusive, but nonetheless take them to be indicative of the kinds of revision that we should make to our ontology of cognitive control. Based on an initial ontology of five functions, they end up with a reduced ontology of only two or three, by grouping three items together into a new functional construct. This is comparable to the way in which Price & Friston proposed grouping the various functions associated with the LPLF under a single unified function, and if applied more widely could have similarly major repercussions for our folk psychological ontology.

An alternative possibility, suggested by Figdor (2011), is that the kind of novel ontology proposed by Lenartowicz *et al* shouldn't be understood as a replacement for either our neurocognitive or folk psychological ontology, but rather as a kind of novel 'task ontology' that can be used to mediate between psychology and neuroscience. Once equipped with this improved task ontology, we could then begin the process of refining our cognitive ontology proper, by eliminating (or revising) “cognitive labels that cannot be stably operationalized” (*ibid*: 225), a process which is likely to involve modifications to some parts of the folk ontology, but could leave other parts intact. Figdor (2018) also argues for a literalist attitude towards the resulting ontology, which might qualify as a partial vindication of folk psychology, depending on how heavily it ends up being modified. Nonetheless, as I will argue in section 4, if we reconceive of folk psychology in more coarse-grained terms then there is no need to worry about the eliminativist implications of cognitive ontology revision in the first place. Figdor's literalist project could happily coexist alongside this re-conception, as an alternative but non-competing way of making sense of the human cognitive system.

The third approach to cognitive ontology revision I want to introduce here is somewhat different, as it doesn't aim to map a single functional category onto each neural structure (thus preserving one-to-one mapping), but rather acknowledges that each neural structure might be implicated in multiple different tasks, and tries to construct an ontology that reflects this. Focusing on the phenomenon of neural reuse, Michael Anderson has proposed that we should characterise neural structures in terms of their 'personalities' rather than their functions, where personalities are understood as “the functional dispositions of individual regions, their underlying causal powers, and their propensities to cooperate with sets of other regions” (Anderson 2014: 114). A region that was previously identified as performing a single, discrete function might instead be characterised in terms of the general kind of contribution it makes to a wide range of tasks, where this contribution does not neatly correspond to anything that we might recognize as a cognitive function. More technically this proposal involves the generation of multidimensional “fingerprint plots” that represent the full range of functional properties associated with the brain (*ibid*: 118). These fingerprint plots closely resemble the diagrams used to represent human personality traits, and are intended to predict activation in a region across a wide range of tasks. For example, the plot for the left inferior parietal sulcus shows the most activation on inhibition tasks, somewhat less activation on vision, motor learning, observation, and preparation tasks, and so on. Rather than coming up with a novel functional description that predicts this behaviour, Anderson wants to give a multidimensional characterisation that accounts for the contributions of this region to a diverse range of tasks. Like Poldrack, he also suggests using statistical techniques to uncover the underlying dimensions that are principally responsible for a region's functional contributions, but these are also going to be unpredictable and opaque from a folk psychological perspective – i.e., Anderson does not envision dimension reduction as a route to the recovery of the folk psychological ontology, but rather as a tool for constructing an alternative. The envisioned outcome is an ontology of 'personalities' rather than functions, preserving one-to-one mapping at the expense of our pre-existing functional categories. Instead of saying that a structure performs a single function like 'word identification', each region of the brain will be given a complex, dispositional analysis that tells us the extent to which it is likely to be implicated in various kinds of

task (for some examples see Anderson 2015: 118). The resulting ontology will look very different to that which we find in folk psychology, consisting of complex, multidimensional descriptions of dispositional properties, rather than simple functional attributions.

Regardless of what kind of solution one endorses to the problem of cognitive ontology revision, it seems likely that we will have to abandon, or at least revise, our existing cognitive ontology in response to it. In the next section I will consider what impact this might have on folk psychology itself, which is the source of the existing ontology, and thus might seem to be threatened by any potential revisions to it.

3. The threat to folk psychology

Having presented three different approaches to cognitive ontology revision, I will now consider the *prima facie* threat that such revision poses to folk psychology. As I suggested in the previous section, this threat arises because our existing cognitive ontology is at least somewhat inspired by folk psychology. If this ontology were successful, enabling one-to-one mappings between functions and structures, it could be seen to vindicate or naturalise our folk psychological categorisation of mental states and processes (at least under the literalist interpretation of folk psychology). However, if it is unsuccessful in the ways described in the previous section, requiring replacement or revision, then folk psychology might require a similar treatment. This would essentially constitute a novel form of eliminative materialism, with developments in cognitive neuroscience threatening to replace or revise our folk psychological ontology.⁸

The one-to-one mapping aspired to by Price & Friston is the ideal target aimed at by much contemporary cognitive neuroscience, at least implicitly. A typical approach to investigating the neural correlates of some cognitive function involves operationalizing that function with a particular task, and then measuring a subject's neural activity while they perform that task. The functions chosen for these studies typically still bear at least a passing resemblance to folk psychological categories, and the tasks that are intended to probe them are clearly inspired by a common-sense interpretation of the function. For example, in the LPLF studies described by Price & Friston, we see functions such as 'processing visual information about written words' and 'processing the visual attributes of animals' which, while expressed in a somewhat more technical manner than we might be used to, at least make some sense from a folk perspective. The hope of the folk psychological realist is that if these kinds of functions could be localised to discrete neural structures, then the folk psychological taxonomy of mental states and processes would be to some extent vindicated or naturalised.

In contrast, each of the proposals for cognitive ontology revision that I considered in the previous section would replace these (relatively) common-sense categories with something that it is much harder to make sense of from a folk psychological perspective. Price & Friston's proposal was to unite the various functions of the LPLF under the umbrella category 'sensorimotor integration', but even if this were more explanatory, it is not at all obvious that such a category has any place in our folk psychological ontology. Similarly, Poldrack's proposed Cognitive Atlas project could find patterns in large quantities of neuroimaging data that would not necessarily bear any direct resemblance to the kinds of behavioural patterns that folk psychology is sensitive to.⁹ Based on such an analysis, Lenartowicz *et al* (2010) proposed grouping together cognitive control, response inhibition, and working memory, each of which might individually make sense to folk psychology,

⁸ I have previously considered similar concerns arising from the predictive processing framework (Dewhurst 2017), and Clark (2019) considers whether this framework would entail the elimination of the folk psychological construct 'desire', responding in part to concerns raised by Klein (2018). Adopting the coarse-grained approach that I advocate here would dissolve concerns of this kind.

⁹ Poldrack discusses some of these issues himself in a blogpost: <http://www.russpoldrack.org/2016/04/how-folksy-is-psychology-linguistic.html>

but when combined do not appear to form a folk-psychologically meaningful cluster. Finally, Anderson's 'neural personalities' describe in quantitative terms the contributions made by each region to a diverse range of tasks, and certainly don't bear any close resemblance to folk psychological categories. So if we are going to have to replace our cognitive ontology with something resembling one of these proposals, then it seems like we will have to abandon the initial hope that we could vindicate or naturalise folk psychology by mapping the states and processes that it identifies on to the activity of neural structures.

Taken one step further, the failure of cognitive neuroscience to vindicate our folk psychological ontology could be used as the basis for a novel argument for eliminative materialism. If our most successful neuroscience requires a revised cognitive ontology composed of neural personalities, or novel categories such as a sensorimotor integration, then it suggests that our original, folk psychologically inspired ontology was also inaccurate, and perhaps deserving of elimination. At the very least this folk psychological ontology will require fairly radical revision if it is going to match up to the novel ontologies introduced in the previous section. Such a revision might subsequently influence how we conceive of ourselves and others, and how we go about predicting and explaining our everyday behaviour. Churchland suggests that, having rejected the folk theory of mind as inadequate, "one might learn to comprehend and report one's internal states and activities within a different and more adequate framework" (1979: 99), i.e. the framework provided by our best cognitive neuroscience. A proponent of a revised cognitive ontology might similarly suggest that we ought to start talking about one another's mental lives in terms of these new categories rather than those of folk psychology.

For the folk psychological realist this possibility will simply constitute a *reductio* of some aspect of the cognitive ontology project, and indeed there have been several alternative responses, such as suggesting that we might instead want to revise our understanding of the mapping relation towards something more context sensitive (Klein 2012, McCaffrey 2015, Burnston 2016; cf. Dewhurst 2019), or adopt a more flexible understanding of the functional structure of the brain (see e.g. Glymour & Hanson 2016). Nonetheless, for the realist who wants to try and identify one-to-one mappings between folk psychologically inspired cognitive functions and discrete neural structures, the kind of evidence appealed to by Price & Friston (2005) does seem to present a serious problem. Either they must accept that our folk psychological ontology is somewhat inaccurate compared with the functional structure of the brain, or they must give up on this particular kind of naturalisation project.

Both the realist and eliminativist interpretations of the relationship between folk psychology and cognitive neuroscience reflect a literalist approach towards our folk psychological ontology. That is, the hope that our folk psychological ontology might be naturalised or otherwise vindicated by neuroscience assumes that folk psychology was always aiming to *literally* describe what is going on in the head, and this assumption is similarly reflected in our feeling of disappointment when it fails to do so. Realist and eliminativist attitudes towards folk psychology can be seen as two sides of the same coin, sharing the basic assumption that the success or failure of folk psychology will depend on its eventual scientific vindication (or lack thereof). We must bear this assumption in mind when we are considering the impact of cognitive ontology revision on folk psychology. In the next section I will consider what an alternative might look like, and how it could make a difference to the implications of the cognitive ontology debate for folk psychology.

4. Towards a coarse-grained folk psychology

I am not the first to draw a connection between proposals for cognitive ontology revision and the threat of a novel eliminative materialism. In this section I will consider two previous engagements with this issue and argue that both point towards a similar solution: rather than embracing eliminativism as a consequence of cognitive ontology revision, we ought to adopt a more coarse-

grained approach, where folk psychology is understood as aiming at predicting and explaining the behaviour of whole persons rather than saying anything about the functional organisation of their brains. I will now present this alternative picture of folk psychology and explain how it avoids the threat from cognitive ontology revision, before exploring its broader implications for the relationship between folk psychology and neuroscience.

The idea that our folk theories might not be in direct conflict with our empirical ones is of course not entirely novel. Similar proposals have been made previously with regard to e.g. emotion categories (Griffiths 1997), biological taxonomies (Dupre 1981), and concepts understood as psychological kinds (Machery 2009). More generally, anti-essentialist theories of natural kinds such as Boyd's (1999) homeostatic property cluster theory and Slater's (2015) stable property cluster theory would seem to support the idea that different 'kinds' of kinds might be appropriate in different social or epistemic contexts (cf. Ludwig 2017 on indigenous and scientific kinds). The attitude towards folk psychology and neurocognitive ontologies that I present here and in the next section is fully compatible with this general trend in the literature on natural kinds towards partial or local eliminativisms/revisionisms, where we can accept changes to our scientific ontology in some domain without thereby threatening the associated folk ontology.

Francken & Slors (2014, see also their 2018) describe how what they call "commonsense cognitive concepts" (i.e. folk psychological concepts) get incorporated into neuroscientific explanations, and how this might give rise to various kinds of problem. They identify an "implicit realism" about commonsense cognitive concepts as being the basis for this incorporation (*ibid*: 253-4), giving rise to the apparent dichotomy between folk psychological realism and eliminativism that I identified in the previous section. Their proposed solution is to instead adopt an 'interpretivist' approach, inspired by Davidson (1980) and Dennett (1987), whereby folk psychology is understood as tracking behavioural patterns rather than aiming to identify discrete states and processes in the brain. This would allow us to acknowledge the failure of folk psychological (or 'commonsense cognitive') concepts at accomplishing the latter task, while also preserving a positive role for folk psychology in interpreting the behaviour of whole persons, and thereby avoiding the eliminativism/realism dichotomy.

Murphy (2017a, see also his 2017b) paints a similar picture, distinguishing between three options that are available to us with regard to folk psychology and the cognitive ontology debate: integration, elimination, or autonomy. Integration is essentially what I have been calling literal realism, where folk psychology is assumed to make empirical claims about the structure of cognition and is therefore vulnerable to the mapping concerns raised by the likes of Price & Friston (2005). Elimination would be the consequence if integration fails, or requires such extensive revisions that our cognitive concepts no longer resemble their folk psychological origins in any meaningful way. Finally, autonomy offers a way out of the integration/elimination dichotomy, by conceiving of the role of folk psychology in a way that does not make it hostage to empirical success. This third option could be accomplished by adopting the interpretivist approach favoured by Francken & Slors (2014), which can help make sense of how folk psychology could be 'autonomous' from neuroscientific details but nonetheless predictive and explanatory of human behaviour. In the rest of this section I will develop this approach in more detail, connecting it with contemporary dispositional approaches and arguing that it is compatible with a certain kind of (non-literal) realism about folk psychology.

As Francken & Slors (2014) note, their interpretivist proposal is probably best developed in Dennett's (1987) intentional stance approach, which conceives of folk psychology as being a particular kind of interpretive 'stance' that one can take towards a complex system, alongside the 'design' and 'physical' stances. The predictive and explanatory success of these stances, according to Dennett, depends on the existence of 'real patterns' in the behaviour of these complex systems, which can only be identified and acted on by interpreting them at a certain level of abstraction. So, the

intentional stance (and thus folk psychology) succeeds by considering the coarse-grained behaviour of a whole person understood as a rational agent, rather than focusing on fine-grained neurophysiological details (which, indeed, we did not even have access to for most of our evolutionary and cultural history). Cognitive neuroscience, in contrast, might have more success by focusing on more fine-grained details, but this does not invalidate the intentional stance, or require that it should be revised in light of its failure to map onto the functional structure of the brain. Indeed, Dennett's approach can explain why our folk psychological ontology might be so different to the revised neurocognitive ontology, as there is no *prima facie* reason to think that the same kinds of concepts are going to be suited for picking up on real patterns at different levels of grain.

Interpretivism, including Dennett's intentional stance approach, also has a lot in common with dispositional approaches, which conceive of mental states (as attributed by folk psychology) as dispositions (behavioural or otherwise) rather than discrete entities. Schwitzgebel (2002) presents a modern defense of dispositionalism about belief, inspired by Ryle (1949), which allows for not only behavioural dispositions but also phenomenal and cognitive dispositions. This kind of account could be extended to other propositional attitudes and folk psychology more generally, and can explain the explanatory and predictive success of folk psychology without committing it to making empirical claims that might be at odds with our neurocognitive ontology. Both interpretivism and dispositionalism also enjoy some empirical support of their own, insofar as it seems like the folk might not actually be committed to making any claims about internal mental states (Curry 2018), and can make better sense of the language we use to express propositional attitudes (Matthews 2017, 2011, see also his 2007/2010).¹⁰ Given that it can also avoid the threat posed by cognitive ontology revision, I think we have good reason to adopt this kind of approach towards folk psychology. The implication of the literal approach is that if folk psychology fails to match up with our best cognitive neuroscience, then it ought to be eliminated or revised. Yet even if this were the case, it seems obvious that 'the folk' (including cognitive scientists themselves) could carry on interpreting one another's behaviour in just the same way that they have been doing for millennia, putting pressure on the realist/eliminativist dichotomy posed by the literatist.

According to the coarse-grained account of folk psychology that I think we ought to endorse, when we attribute mental states to someone we are doing no more than saying that they are disposed to behave in certain ways. Crucially, we are making no commitment as to the structure of the mechanisms responsible for that behaviour, and as such folk psychology should not be taken to aim at literally describing these mechanisms. When I say that someone believes something, I just mean that they are likely to act as though it were true, and when I say that someone is brave or intelligent I am just making a general statement about the kinds of behaviour and competencies they are likely to exhibit. In neither case am I committing to any details about what is going on 'in the head', nor am I saying anything that could possibly be falsified by cognitive scientific discoveries. Even if it turned out that there was nothing resembling the structure of belief-desire psychology going on in the brain, or that character traits like bravery and intelligence were not stable scientific constructs, we could nonetheless go on attributing these concepts to one another in a meaningful manner. It might turn out that their meaningfulness has more to do with sociocultural constraints on behaviour than with anything mechanistic (see e.g. Zawidzki 2013), but it would nonetheless be meaningful. Understood in this way, folk psychology is in principle not vulnerable to scientific refutation, for so long as we continue to exhibit the correct behavioural dispositions, it will remain a coherent kind of social practice.

It is also important to note that both the interpretive and dispositional account of folk psychology can be construed as 'realist', albeit of a distinct form from the literal realism that I

¹⁰ Although see Quilty-Dunn & Mandelbaum (2018) for some recent criticism of dispositionalism.

identified previously (i.e. the kind of realism that forces a dichotomy with eliminativism). Interpretivism about folk psychology is realist insofar as it claims that when we attribute mental states to one another, we are doing so on the basis of Dennettian real patterns, which is to say we are identifying and interpreting real patterns in the behaviour of those we attribute mental states to (see Dennett 1981, 1991; cf. Ross 2000).¹¹ Similarly, dispositionalism about folk psychology is realist insofar as the dispositions we attribute to one another are just as real as any other dispositions, such as that of a soluble object (like a sugar cube) to dissolve when placed in water. Even if folk psychology does not correctly identify the fine-grained functional structure of the brain, it can nonetheless correctly identify behavioural patterns and dispositions which are just as real as those described by neuroscience.

5. The relationship between folk psychology and neuroscience

Adopting a more coarse-grained approach to folk psychology can help us to re-evaluate the relationship between folk psychology and cognitive neuroscience. Rather than conceiving of neuroscience as aiming to identify the neural correlates of more-or-less folk psychological categories, we ought to instead conceive of it as aiming to uncover the complex mechanisms that give rise to the kind of *behaviour* that folk psychology describes, predicts, and explains.¹² There is no reason to think that these mechanisms will conform to the categories of folk psychology, but equally no reason for concern when they fail to do so. The two kinds of ontology (folk psychological and neuroscientific) are simply so different that we cannot (and should not) even try to directly compare them. McDowell (1994) makes a similar point about the relationship between subpersonal mechanisms and personal level perceptual experience: the former somehow enable the latter, but there is no conflict between the two, and no reason to think that one ought to be reducible to the other. Furthermore, there is also no reason to think that we should have any kind of privileged access to subpersonal mechanisms (either our own or those of other people), such that their structure might be reflected in our folk psychological categories.¹³

The literal interpretation of folk psychology, on the other hand, makes a commitment to certain kinds of empirical discoveries (i.e., neural structures with a functional architecture that can be mapped onto folk psychology). If our best cognitive ontology turns out to be radically different to the folk psychological ontology, then this might mean that the latter must be eliminated or revised. There are of course other options available to the literal realist. They could accept a limited amount of revision to the folk psychological ontology, stopping short of full-blown eliminativism. They could also insist that it must be the neuroscience itself that is wrong, adopting a ‘top-down’ strategy and revising our interpretation of the neuroimaging data in order to match up with the folk ontology. There is a lot of interpretive work that must be done when conducting neuroimaging studies, all of which gives us some room for manoeuvre. For example, by switching to a network analysis of the functional relevance of neural activity (see e.g. Glymour & Hanson 2016; see also Wright, this volume), we could avoid the need to map cognitive functions directly onto neural structures, and thus perhaps preserve the neuroscientific relevance of the folk psychological ontology.

However, regardless of whether a strategy like this is successful, by moving to a more coarse-grained understanding of folk psychology we can avoid the threat of eliminativism entirely. One way to think of this approach is simply as a restatement of the idea that applying folk psychological

¹¹ Whether or not Dennett himself should be interpreted as a realist is a complicated question which I do not intend to get into here. It is sufficient for my purposes that there is a sense in which his approach to folk psychology can be understood as realist.

¹² See Raja & Anderson (this volume) for further discussion of the relationship between neuroscience and behaviour.

¹³ For more on the personal/subpersonal distinction, see Dennett (1969) and Drayson (2012, 2014).

concepts to neuroscience constitutes a category mistake (cf. Bennett & Hacker 2003), or that it somehow mixes up the kinds of language used to describe our manifest and scientific images of the world (cf. Sellars 1963). Our folk psychological ontology reflects the manifest image, our cognitive ontology reflects the scientific image, and there is no in-principle reason to think that they should be reconcilable. Of course, this approach would also rule out any straightforward reduction of the mental to the physical, although that is not to say that the mental states picked out by folk psychology are entirely independent of the physical states studied by cognitive neuroscience. There is more work to be done on how to make sense of this relationship in a naturalistic manner, but my own preferred approach is to see folk psychology as picking out (real) patterns in person-level behaviour that are generated by neuroscientific mechanisms (cf. Dennett 1991). Looked at in this way there is no need to eliminate, or even revise, folk psychology in response to developments in cognitive neuroscience, as it will remain just as good as it ever has been at picking out person-level patterns.¹⁴ In some cases the folk *are* interested in something more fine-grained, such as when they pursue a clinical intervention from a neurosurgeon, but in these cases I think we should understand them as deferring to the expertise (and ontology) of the scientific community, rather than as adopting a more fine-grained ontology.

The coarse-grained approach does still allow for a kind of *partial* eliminativism, which acknowledges the failure of folk psychological concepts at tracking fine-grained neuroscientific states and processes (i.e., the mapping problem), and allows that they might need to be revised, replaced, or eliminated from this explanatory context. Hence adopting this approach is compatible with calling for the revision of our cognitive ontology for neuroscientific purposes, and this might mean we will end up with a neuroscientific ontology that is very different to our folk psychological ontology. An ontology such as that envisioned by Price & Friston, Poldrack, or Anderson would likely only be translatable into folk psychological terms with considerable effort, and might not even be translatable in any meaningful sense at all. Rather than seeing this a problem to be avoided, I think we should instead try to come to terms with it, by being honest about the limitations of the scientific image for making sense of everyday experience, without thereby taking this to mean that everyday experience is somehow inexplicable. This is something that we have already had to deal with in other domains, such as physics, where our best ontology has no clear correspondence to everyday experience. The issue is perhaps more pressing when it comes to folk psychology, which is both more immediate and more personal than folk physics, but this doesn't mean that such an approach cannot be made to work.¹⁵

It is also possible that empirical and theoretical developments in cognitive neuroscience *will* eventually have an impact on the folk psychological ontology, in the same way that psychoanalytic concepts like 'the unconscious' entered the folk ontology during the 20th century (cf. Richards 2000). Something similar may have happened in recent decades with the adoption (by the folk) of neurochemical terminology when describing and attributing certain kinds of mental states. e.g. statements like 'I'm not feeling great today, my serotonin levels are a bit low' (cf. Rodriguez 2006, Rose & Abi-Rached 2013, Francken & Slors 2018). It seems plausible (as Murphy 2017b suggests)

¹⁴ Which is not to say that it is *very* good at this. It is plausible that the success of folk psychology is at least somewhat overrated, especially when it comes to edge cases like mental illness and socially disruptive behaviour (see e.g. Matthews 2013 for some discussion of these issues, and the benefits of taking a dispositional approach to them). However, it is clearly successful at least some of the time, and the approach taken here can help make sense of how this could be true even if it fails to track the fine-grained structure of neural processing.

¹⁵ One strategy, which I will not pursue here, would be to use our neurocognitive ontology to *explain* why our folk psychological ontology is the way that it is, without treating such an explanation as a route to elimination or reduction. This would be a non-eliminativist version of the so-called 'illusionist' approach to conscious experience (see e.g. Frankish 2017), although as noted by Graziano (2016: 112-3), the label 'illusionist' might be somewhat misleading in this context.

that the most extensive changes to the folk ontology might come in response to psychiatric research, which can sometimes offer satisfying explanations for otherwise disturbing or inexplicable behaviour. In the non-pathological cases, where folk psychology is relatively successful at predicting and explaining behaviour, there is no need for it to conform to novel scientific categories, but in the pathological cases, where this behaviour is perhaps harder for it to explain, it might be more susceptible to influences from the cognitive scientific ontology (cf. Matthews 2013).

Here the difference from classical eliminativism is that these adjustments are not mandated or required by philosophers, but rather occur naturally as a process of linguistic or conceptual development, and may often not accurately reflect the neuroscience that they are inspired by. Murphy (2017a: 141) suggests that the real impact of cognitive neuroscience on folk discourse might have more to do with the ethical and political implications of our changing self-conception, but I think that such concerns are best addressed independently of empirical questions about what kind of ontology is best suited for scientific practice.¹⁶ Adopting a more coarse-grained approach would allow us to keep questions about our neuroscientific ontology separate from questions about our folk ontology, even if the latter is sometimes informally influenced by the former (and *vice versa*, for better or worse). To be clear, my view is not that the folk psychological ontology is necessarily static and unchanging, but rather that changes to it are likely to be at best indirectly related to changes to our neurocognitive ontology, via the unpredictable medium of social and cultural interpretation of psychology, psychiatry, and neuroscience. Any deliberate attempts to change the folk psychological ontology are likely to be ineffective at best and/or to have unintentional (and potentially harmful) outcomes at worst.

How does all this relate to the broader conception of folk psychology introduced in section 1? Character traits are plausibly just another kind of interpretation of whole persons, and already fitted well into the dispositional picture, as being ‘brave’, for example, can be understood in terms of being disposed to behave bravely when circumstances require it. Folk psychological narratives typically concern the actions of persons, not parts of their brain, and can again be understood as ways of interpreting and explaining those actions. Taking on a certain role in a narrative will also mean being disposed to behave in certain ways, and the social understanding that we can gain from these shared narratives does not depend on the structure of the (neural) mechanisms that generates that behaviour. Finally, the normative constraints imposed by folk psychology can help to ensure that our behaviour conforms to folk psychological expectations, regardless of the structure of subpersonal mechanisms. In this sense folk psychology can be thought of as a kind of self-fulfilling prophecy, sometimes acting to generate the very same behaviour that it predicted.¹⁷ This also means that when we attribute folk psychological states to one another, we may have more than merely epistemic aims in mind (we might also be aiming to influence each other’s behaviour, for example). Folk psychology has a broader social function that goes beyond mere prediction and explanation, and this function is not necessarily threatened by revisions to our neurocognitive ontology. Even if it turned out that the language of beliefs and desires, hopes and fears, norms and narratives, and so on, was completely unsuited to our

¹⁶ Knobe (2007) explores some ways in which moral judgements might both influence and be influenced by folk psychology, and suggests that neuroscientific concepts could not play the same kind of role. For some recent considerations of the broader moral and social implications of contemporary neuroscience, see Caruso & Flanagan (2018).

¹⁷ See Andrews 2015 for further discussion of what she calls “the folk psychological spiral”, where our explanation of some unusual behaviour might commit us to acting more predictably in the future. Zawidzki (2013) presents a more general account of how what he calls “mindshaping” might help to regulate our behaviour in a way that makes predicting and explaining it computationally tractable. Understood in this way, folk psychological concepts would constitute socially constructed “human kinds”, in Hacking’s (1995) sense.

analysis of neuroimaging studies, it would not stop being useful for our understanding of whole persons, and there would be no reason to think that we ought to revise or eliminate it.

By adopting a coarse-grained approach to the folk psychological ontology, we can effectively inoculate it against any eliminativist threat, including not only the current threat from cognitive ontology revision, but also potential future threats from novel neuroscientific discoveries. At the same time, we ought to be sensitive to the misuse of folk psychological concepts within cognitive neuroscience, especially when such concepts do not pick out cognitive functions that map adequately onto the functional architecture of the brain. In such cases we should develop novel cognitive ontologies that better reflect this architecture, but doing so need not entail making any changes to analogous components of the folk psychological ontology. We can simply accept that the two ontologies have different targets (whole persons versus neural structures), and correspondingly different explanatory standards and predictive goals.

Conclusion

In section 1 I introduced some different ways of understanding folk psychology and argued that the dichotomy between folk psychological realism and eliminativism depends on a fine-grained interpretation, where folk psychological concepts are understood as literally aiming to describe the mechanistic structure of cognition. In section 2 I introduced the recent debate over cognitive ontology revision in neuroscience, and in section 3 I demonstrated how some existing responses to this debate could threaten our existing folk psychological ontology. In sections 4 and 5 I presented an alternative approach to folk psychology and considered how this might change our understanding of the relationship between folk psychological and neuroscientific ontologies. I argued that, in order to avoid the threat of eliminativism posed by cognitive ontology revision, we ought to reject the fine-grained, literal understanding of folk psychology and instead adopt a coarse-grained approach, where folk psychology aims to predict and explain the behaviour of whole persons rather than tracking the mechanistic structure of cognition. Doing so would insulate folk psychology from the threat posed by cognitive ontology revision, and it can also help us to better understand the relationship between folk psychology and cognitive neuroscience, which should be seen as different levels of description rather than competing ontologies.

Acknowledgements

Many thanks to Jonny Lee, Adrian Downey, E. Brown Dewhurst, and Carrie Figdor for providing helpful comments on earlier drafts, to J. Brendan Ritchie for his very helpful reviewer comments, and to Marco Viola and Fabrizio Calzavarini for hosting the *Neural Mechanisms* lecture series and editing this volume. Earlier versions of the material in this chapter have been presented at many workshops and conferences, including the *BSPS 2016 Annual Conference* in Cardiff, the *Early Career Mind Network Research Forum* in Durham in 2016, the “Symposium on Structure-Function Mappings in Cognitive Neuroscience” at the *14th Annual Conference of the Italian Society for Cognitive Science* in Bologna in 2017, and the *Colloquium on Consciousness and Cognition* at the Ruhr-Universität Bochum in June 2018.

References

- Amedi, A., Jacobson, G., Hendler, T., Malach, R., & Zohary, E. 2002. “Convergence of visual and tactile shape processing in the human lateral occipital complex.” *Cerebral Cortex*, 12: 1202-212.
- Anderson, M. 2010. “Neural reuse: A fundamental organizational principle of the brain.” *Behavioural and Brain Sciences*, 33/4: 254-61.
- Anderson, M. 2014. *After Phrenology*. Cambridge, MA: MIT Press.

- Andrews, K. 2015. “The Folk Psychological Spiral: Explanation, Regulation, and Language.” *The Southern Journal of Philosophy*, 53: 50-67.
- Apperly, I. A. 2008. “Beyond Simulation-Theory and Theory-Theory.” *Cognition* 107/1: 266-83.
- Barrett, H. C. 2012. “A hierarchical model of the evolution of human brain specializations.” *Proceedings of the National Academy of Sciences*, 109: 10733-40.
- Bennett, M. R. & Hacker, P. M. S. 2003. *Philosophical Foundations of Neuroscience*. Malden, MA: Blackwell Publishing.
- Boone, W., & Piccinini, G. 2016. “The cognitive neuroscience revolution.” *Synthese*, 193/5: 1509-34.
- Boyd, R. 1999. “Homeostasis, species, and higher taxa.” In Wilson (ed.), *Species: New Interdisciplinary Essays*. Cambridge, MA: MIT Press.
- Bruner, B. 1990. *Acts of Meaning*. Cambridge, MA: HUP.
- Burnston, D. 2016. “A contextualist approach to functional localization in the brain.” *Biology & Philosophy*, 31/4: 527-50.
- Caruso, G. & Flanagan, O. (eds.) 2018. *Neuroexistentialism*. Oxford: OUP.
- Churchland, P. M. 1979. *Scientific Realism and the Plasticity of Mind*. Cambridge, UK: CUP.
- Churchland, P. M. 1981. “Eliminative materialism and the propositional attitudes.” *Journal of Philosophy*, 78: 67-90.
- Churchland, P. S. 1976. *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press.
- Clark, A. 2019. “Beyond Desire? Agency, Choice, and the Predictive Mind.” *Australasian Journal of Philosophy*, DOI: [10.1080/00048402.2019.1602661](https://doi.org/10.1080/00048402.2019.1602661)
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M., & Michel, F. 2000. “The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients.” *Brain*, 123: 291–307.
- Curry, D. S. 2018. “Beliefs as inner causes: The (lack of) evidence.” *Philosophical Psychology*, 31:6, 850-77.
- Davidson, D. 1980. *Essays on actions and events*. Oxford: OUP.
- De Jaegher, H. & Di Paolo, E. 2007. “Participatory Sense-Making: An enactive approach to social cognition.” *Phenomenology and the Cognitive Sciences*, 6/4: 485-507.
- Dennett, D. 1969. *Content and Consciousness*. Routledge and Kegan Paul.
- Dennett, D. 1981. “True Believers.” In Haugeland (ed.), *Mind Design*. Cambridge, MA: MIT Press.
- Dennett, D. 1987. *The Intentional Stance*. Cambridge, MA: MIT Press.
- Dennett, D. 1991. “Real Patterns.” *The Journal of Philosophy*, 88/1: 27-51.
- Dewhurst, J. 2017. “Folk Psychology and the Bayesian Brain.” In Metzinger & Wiese (eds.), *Philosophy and Predictive Processing*. Frankfurt am Main: MIND Group.
- Dewhurst, J. 2019. “Context sensitive ontologies for a non-reductionist cognitive neuroscience.” *Australasian Philosophical Review*, 2/2: 224-8.
- Drayson, Z. 2012. “The uses and abuses of the personal/subpersonal distinction.” *Philosophical Perspectives*, 26/1:1-18.
- Drayson, Z. 2014. “The Personal/Subpersonal Distinction.” *Philosophy Compass*, 9/5: 338-46.
- Dupre, J. 1981. “Natural Kinds and Biological Taxa.” *The Philosophical Review*, 90/1: 66-90.
- Feyerabend, P. 1963. “Mental events and the brain.” *Journal of Philosophy*, 60: 295-6.

- Figdor, C. 2011. “Semantics and Metaphysics in Informatics: Toward an Ontology of Tasks.” *Topics in Cognitive Science* 3: 222-6.
- Figdor, C. 2018. *Pieces of Mind*. Oxford: OUP.
- Fodor, J. 1974. “Special sciences (or: The disunity of science as a working hypothesis).” *Synthese*, 28/2: 97-115.
- Fodor, J. 1975. *The Language of Thought*. Cambridge, MA: MIT Press.
- Fodor, J. 1987. *Psychosemantics*. Cambridge, MA: MIT Press.
- Francken, J. & Slors, M. 2014. “From commonsense to science, and back.” *Consciousness and Cognition*, 29: 248-58.
- Francken, J. & Slors, M. 2018. “Neuroscience and everyday life.” *Brain and Cognition*, 120: 67-74.
- Frankish, K. (ed.) 2017. *Illusionism as a Theory of Consciousness*. Exeter: Imprint Academic.
- Gallagher, S. 2008a. “Direct perception in the intersubjective context.” *Consciousness and Cognition*, 17: 535-43.
- Gallagher, S. 2008b. “Inference or interaction.” *Philosophical Explorations*, 11/3: 163-74.
- Glymour, C. & Hanson, C. 2016. “Reverse Inference in Neuropsychology.” *The British Journal for the Philosophy of Science*, 67/4: 1139-153.
- Gopnik, A. & Wellman, H. 1992. “Why the Child’s Theory of Mind Really is a Theory.” *Mind and Language*, 7: 145-72.
- Gordon, R. 1986. “Folk Psychology as Simulation.” *Mind and Language*, 1/2: 158-71.
- Graziano, M. 2016. “Consciousness engineered.” *Journal of Consciousness Studies*, 23/11-12: 98-115.
- Griffiths, P. 1997. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago, IL: UCP.
- Hacking, I. 1995. “The looping effects of human kinds.” In Sperber, Premack, & Premack (eds.), *Causal Cognition*, Oxford: Clarendon Press.
- Heal, J. 1986. “Replication and Functionalism”. In *Language, Mind, and Logic* (ed. Butterfield). Cambridge: CUP.
- Hutto, D. 2008. *Folk Psychological Narratives*. Cambridge, MA: MIT Press.
- Klein, C. 2012. “Cognitive Ontology and Region- versus Network-Oriented Analyses.” *Philosophy of Science*, 79/5: 952-60.
- Klein, C. 2018. “What Do Predictive Coders Want?” *Synthese* 195/6: 2541–57.
- Knobe, J. 2007. “Folk Psychology: Science and Morals.” In Hutto & Ratcliffe (eds.), *Folk Psychology Re-Assessed*. Springer.
- Knoll, A. 2018. “Still Autonomous After All.” *Minds and Machines*, 28/1: 7-27.
- Lavelle, J. S. 2012. “Theory-theory and the direct perception of mental states.” *Review of Philosophy and Psychology*, 3: 213-30.
- Lavelle, J. S. 2019. *The Social Mind*. Routledge.
- Lenartowicz, A., Kalar, D. J., Congdon, E., & Poldrack, R. A. 2010. “Towards an ontology of cognitive control.” *Topics in Cognitive Science*, 2(4): 678-92.
- Lewis, D. 1972. “Psychophysical and theoretical identifications.” *Australasian Journal of Philosophy*, 50/3: 249-58.
- Ludwig, D. 2017. “Indigenous and scientific kinds.” *The British Journal for the Philosophy of Science*, 68/1: 187-212.
- Lycan, W. 1988. *Judgement and Justification*. Cambridge: CUP.
- Machery, E. 2009. *Doing Without Concepts*. Oxford: OUP.
- Mallon, R., Machery, E., Nichols, S., & Stich, S. 2009. “Against Arguments from Reference.” *Philosophy and Phenomenological Research*, 79/2: 332-56.

- Mameli, M. 2001. “Mindreading, mindshaping, and evolution.” *Biology and Philosophy*, 16/5: 595-626.
- Martin, A., & Chao, L. L. 2001. “Semantic memory and the brain: Structure and processes.” *Current Opinions in Neurobiology*, 11: 194-201.
- Matthews, R. 2007/2010. *The Measure of Mind: Propositional Attitudes and their Attribution*. Oxford: OUP.
- Matthews, R. 2011. “Measurement-Theoretic Accounts of Propositional Attitudes.” *Philosophy Compass*, 6/11: 828-41.
- Matthews, R. 2013. “Belief and Belief’s Penumbra.” In Nottlemann (ed.), *New Essays on Belief*. Palgrave Macmillan.
- Matthews, R. 2017. “The Elusive Case for Relationalism about the Attitudes.” *Philosophy and Phenomenological Research*, online first: doi: 10.1111/phpr.12380.
- McCaffrey, J. 2015. “The Brain’s Heterogeneous Functional Landscape.” *Philosophy of Science*, 82/5: 1010-22.
- McCaffrey, J., and Machery, E. 2016. “The reification objection to bottom-up cognitive ontology revision.” *Behavioral and Brain Sciences*, 39: e125.
- McCauley, & Bechtel, W. 2001. “Explanatory Pluralism and Heuristic Identity Theory.” *Theory & Psychology*, 11(6): 736-60.
- McGeer, V. 2007. “The Regulative Dimension of Folk Psychology.” In Hutto & Ratcliffe (eds.), *Folk Psychology Re-assessed*. Springer
- McDowell, J. 1994. “The Content of Perceptual Experience.” *The Philosophical Quarterly*, 44/175: 190-205.
- Mitchell, J. P. 2005. “The false dichotomy between simulation and theory-theory.” *Trends in Cognitive Science*, 9/8: P363-4.
- Murphy, D. 2017a. “Brains and Beliefs.” In Kaplan (ed.), *Explanation and Integration in Mind and Brain Science*. Oxford: OUP.
- Murphy, D. 2017b. “Can psychiatry refurbish the mind?” *Philosophical Explorations*, 20/2: 160-74.
- Piccinini, G., & Craver, C. 2011. “Integrating psychology and neuroscience: functional analyses as mechanism sketches.” *Synthese*, 183/3: 283-311.
- Poldrack, R. 2010. “Mapping Mental Function to Brain Structure.” *Perspectives on Psychological Science*, 5/6: 753-61.
- Poldrack, R. & Yarkoni, T. 2016. “From Brain Maps to Cognitive Ontologies.” *Annual Review of Psychology*, 67: 587-612.
- Premack, D. & Woodruff, G. 1978. “Does the chimpanzee have a theory of mind?” *Behavioral and Brain Sciences*, 4: 515-26.
- Price, C. J., & Friston, K. J. 2005. “Functional ontologies for cognition: The systematic definition of structure and function.” *Cognitive Neuropsychology*, 22/3: 262–275.
- Quilty-Dunn, J. & Mandelbaum, E. 2018. “Against dispositionalism: belief in cognitive science.” *Philosophical Studies*, 175: 2353-72.
- Richards, G. 2000. “Britain on the Couch: The Popularisation of Psychoanalysis in Britain 1918-1940.” *Science in Context*, 13/2: 183-230.
- Rodriguez, P. 2006. “Talking brains: a cognitive semantic analysis of an emerging folk neuropsychology.” *Public Understanding of Science*, 15: 301-30.
- Rorty, R. 1965. “Mind-Body Identity, Privacy, and Categories.” *The Review of Metaphysics*, 19/1: 24-54.
- Rose, N. & Abi-Rached, J. M. 2013. *Neuro: The new brain sciences and the management of the mind*. Princeton University Press.
- Ross, D. 2000. “Rainforest realism: A Dennettian theory of existence.” In D. Ross, A. Brook,

and D. Thompson (eds.), *Dennett's Philosophy: A Comprehensive Assessment*. Cambridge, MA: MIT Press.

- Ryle, G. 1949. *The Concept of Mind*. Hutchinson.
- Schwitzgebel, E. 2002. "A phenomenal, dispositional account of belief." *Nous*, 36: 249-75.
- Sellars, W. 1956. "Empiricism and the Philosophy of Mind." In Feigl & Scriven (eds.), *Minnesota Studies in the Philosophy of Science*. University of Minnesota Press.
- Sellars, W. 1963. *Science, Perception, and Reality*. New York: Humanities Press.
- Slater, M. 2015. "Natural Kindness." *The British Journal for the Philosophy of Science*, 66/2: 375-411.
- Spaulding, S. 2018. "Mindreading beyond belief: A more comprehensive conception of how we understand others." *Philosophy Compass*, 13/11.
- Stich, S. 1983. *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press.
- Stich, S. 1996. *Deconstructing the Mind*. Oxford: OUP.
- Towl, B. N. 2011. "Mind-brain correlations, identity, and neuroscience." *Philosophical Psychology*, 25/2: 187-202
- Westra, E. 2018. "Character and theory of mind: an integrative approach." *Philosophical Studies*, 175/5: 1217-41.
- Zawidzki, T. 2013. *Mindshaping*. Cambridge, MA: MIT Press.