

14 Neuroscience and Cognitive Ontology: A Case for Pluralism

Joseph McCaffrey and Jessey Wright

14.1 Introduction

Developing taxonomies—systems for classifying phenomena or objects of study—is an important scientific activity. By categorizing and classifying phenomena, scientists can more effectively organize research programs, and can even discover new phenomena and relations between them. Cognitive psychology seeks to explain and predict human behavior. The mental domain is thus organized into categories such as attention, episodic memory, reward prediction, and cognitive control. Collectively, these categories comprise a “cognitive ontology”—a set of categories taken to reflect the mind’s organization (Price & Friston, 2005; Poldrack, 2010; Klein, 2012; Anderson, 2015). Cognitive neuroscience aims to understand how different brain structures carry out these psychological capacities.

This chapter examines the role of cognitive neuroscience in shaping our cognitive ontology. The term “cognitive ontology” sometimes denotes the informatics concept of a database mapping the domain of mental entities and the relationships between them (Price & Friston, 2005; Poldrack, Kitur, et al., 2011). Elsewhere, cognitive ontology refers more loosely to the taxonomy of mental kinds adopted by contemporary psychology. Worries about our cognitive ontology manifest as concerns about the terminology used to refer to cognitive phenomena, the mental kinds those terms and concepts refer to, or as conflicts between views about the structure of cognition itself (Janssen, Klein, & Slors., 2017).¹ These concerns are interrelated. For instance, a clear and shared ontology is important for fruitfully engaging in debates about the structure of cognition or bringing empirical results to bear on such debates. After all, scientists cannot integrate findings from

different research projects if they do not, at minimum, use terms and concepts in similar ways.

How do we know if we have the right concepts for mapping the mind? On one hand, psychology exhibits conceptual progress over time. For example, most psychologists would agree that response inhibition, a psychological construct related to impulse control and measured using tasks such as the go/no-go task (Morein-Zamir & Robbins, 2015), is a better psychological category than Sigmund Freud's id or pleasure principle (Freud, 1989).² On the other hand, there are numerous debates about whether particular elements and concepts deserve to be in our cognitive ontology. Researchers debate how many types of memory there are (Tulving, 2007), whether emotions such as fear and anger are real mental kinds (Barrett, 2006; Lindquist et al., 2012), whether task switching is a unique cognitive control element (Lenartowicz et al., 2010), and so forth. In many domains of psychology (social, affective, cognitive, etc.), there are debates about the foundational terms and concepts used to describe our mental abilities, and about how these terms and concepts align with the experimental tasks used to study them in the laboratory (Sabb et al., 2008, Scarantino & Griffiths, 2011; Sullivan, 2016, 2017; Quesque & Rossetti, 2020).

Recently, philosophers and cognitive scientists have expressed a more global concern that perhaps our entire cognitive ontology (or large territories of it) needs substantial revision (Price & Friston, 2005; Lenartowicz et al., 2010; Poldrack, 2010; Anderson, 2015). In this chapter, we present and critique the main arguments motivating large-scale cognitive ontology revision in the philosophical and neuroscientific literature. The first argument for large-scale ontological revision is methodological and conceptual. It asserts that psychological categories are folksy, unscientific, and otherwise carelessly deployed. Not only should we expect radical ontological revision as neuroscience makes progress, but the looseness of the ontology itself may be holding cognitive neuroscience back (Poldrack, 2010; Bunzl, Hanson, & Poldrack, 2010). The second argument for cognitive ontology revision is empirical. This argument alleges that neuroimaging data—for example, functional magnetic resonance imaging (fMRI), positron emission tomography, and magnetic electroencephalography studies—suggest that the processes specified in our cognitive ontology do not map onto structural or functional features of the brain (e.g., Lenartowicz et al., 2010).

Advocates for this position claim these findings challenge the status of cognitive terms as demarcating scientifically legitimate kinds.

We have three goals for this chapter. The first goal is to familiarize readers with the main issues and debates about cognitive ontology revision (section 14.2). The second goal is to show that while the master arguments for large-scale ontology revision highlight important theoretical and methodological issues in the mind–brain sciences (we agree that neuroscience will contribute to taxonomic revision in psychology and that psychologists should systematically clarify their conceptual categories), our existing ontology is not necessarily the problem (section 14.3). The third goal is to challenge the widespread assumption of taxonomic monism in the cognitive ontology literature. Most researchers working on this issue seem to hold that (1) each item in a scientifically valid cognitive ontology will map to a specific item in the neural ontology (the collection of categories for organizing and describing parts or functions of the brain), and (2) there exists one correct (or best) cognitive ontology. We think these assumptions are mistaken and instead echo and reinforce calls for taxonomic pluralism (Danks, 2015; Hochstein, 2016; Sullivan, 2016, 2017). Taxonomic pluralism, applied to these issues, is the view that cognitive neuroscience will need different ontologies for achieving different research goals and for providing different perspectives on the phenomena of interest. After advocating for taxonomic pluralism, we devote the end of this chapter to examining what that might look like for cognitive neuroscience (section 14.4).

14.2 The Case for Cognitive Ontology Revision

The question of how to classify mental kinds has deep historical roots in philosophy and psychology. The ancient Greek philosopher Plato argued that the soul has three parts: the rational part, the spirited part, and the appetitive part (Plato, *Republic IV*, 431a1 ff). Scottish philosopher Thomas Reid (1710–1796) recognized dozens more “mental faculties” such as memory, abstraction, judgment, and imagination (Brooks, 1976). Disputes about how to classify mental kinds are similarly longstanding. Nineteenth-century phrenologist Franz Gall (1758–1828), who advocated twenty-seven mental faculties (e.g., memory for words, sense of places, poetic talent, love of offspring,

etc.), chastised philosophers for developing their mental taxonomies through internal reflection instead of scientific investigation (Gall, 1835).³

The recent debate about cognitive ontology brings these perennial questions about classifying mental kinds into the contemporary era of big data, open science initiatives, and brain imaging (Poldrack, 2010; Poldrack, Kittur, et al., 2011; Poldrack & Yarkoni, 2016). This emerging literature has three recurring themes. The first is a concern that our current psychological categories and the practices surrounding them may be holding back scientific research (Poldrack, Kittur, et al., 2011; Sullivan, 2016; Quesque & Rossetti, 2020). The second is that developing cognitive ontologies will improve science by making conceptual disagreements explicit and permitting researchers from disparate laboratories and research traditions (e.g., psychiatry and cognitive neuroscience) to integrate their findings (Price & Friston, 2005; Sabb et al., 2008; Poldrack, Kittur, et al., 2011; Sullivan, 2017). The third concerns how our cognitive ontology aligns with neuroimaging data (Anderson, 2010, 2014; Klein, 2012; Rathkopf, 2013; McCaffrey, 2015; Viola, 2017; Khalidi, 2017). Some propose using neuroimaging to determine what elements belong in our cognitive ontology (Lenartowicz et al., 2010; Poldrack 2010). This raises important questions about “the proper role of neuroscientific evidence in determining what mental entities there are” (Viola, 2017, p. 163).

There are varying views about how neuroscience might transform our cognitive ontology. Some researchers envision modest reform, while others envision a radical overhaul and a new beginning (Anderson, 2015). There are many research agendas with different theoretical orientations, experimental methods, and analytic tools. Some initiatives are primarily aimed at charting our existing cognitive ontology to achieve greater terminological and conceptual consistency across laboratories and fields (e.g., Poldrack, Kittur, et al., 2011; Quesque & Rossetti, 2020). Other work calls for testing our existing ontology using brain data (e.g., Lenartowicz et al., 2010; Lindquist et al., 2012), or using brain data to mine novel relationships between tasks and brain activation patterns (Poldrack, Halchenko, et al., 2009; Anderson, Kinnison, & Pessoa, 2013; Yeo et al., 2015). Despite these differences, we can extract an overarching narrative from advocates of cognitive ontology revision (Price & Friston, 2005; Poldrack, 2010; Anderson, 2014, 2015). According to this narrative, our current cognitive ontology is systematically ill suited for explaining human behavior or for predicting patterns of brain activation in experiments *because* it was developed

in absence of consideration of the brain. We should rectify this by testing, refining, and potentially rebuilding our cognitive ontology using the tools of neuroscience (Lenartowicz et al., 2010; Poldrack, 2010; Anderson, 2015). As Michael Anderson (2015) puts it, “the brain can (and should) act as one arbiter of the psychologically real” (p. 70).

There are two main arguments for this position. The first argument for revision is that we should distrust our current ontology because it contains concepts that are provisional, vague, or borrowed from prescientific or “folk” psychology (Bunzl et al., 2010; Poldrack, 2010). As cognitive neuroscientist Russel Poldrack laments, “the fundamental problem is our stone age psychological ontology” (Bunzl et al., 2010, p. 54). Moreover, cognitive psychologists often deploy these concepts heterogeneously and imprecisely (Poldrack, Kittur, et al., 2011; Quesque & Rossetti, 2020) or fail to analyze the relationships rigorously between tasks (activities performed by human subjects in a laboratory) and the mental processes those tasks recruit (Figdor, 2011; Sullivan, 2017). The second argument for revision is the empirical observation that neuroimaging studies show virtually no one-to-one mappings between brain structures or activation patterns and mental processes (Price & Friston, 2005; Anderson, 2010; Poldrack, 2010). This implies that our current taxonomy of mental processes fails to map onto the neural structures underlying cognition. We discuss these arguments in turn.

14.2.1 Why Our Cognitive Ontology Is Mistaken: Conceptual Woes

The first major argument for large-scale cognitive ontology revision is that we have good antecedent conceptual reasons for doubting our existing cognitive ontology. According to this proposal, our cognitive ontology contains numerous items that are vague, ill defined, or borrowed from folk psychology (Bunzl et al., 2010; Poldrack, 2010). Furthermore, different laboratory groups frequently use cognitive terms and concepts interchangeably (Poldrack, Kittur, et al., 2011; Quesque & Rossetti, 2020) and/or fail to provide adequate analyses of the tasks used to study cognition and the cognitive processes those tasks allegedly recruit (Figdor, 2011; Swick, Ashley, & Turken, 2011; Sullivan, 2016, 2017). In other words, psychology may have many flawed concepts and a set of embedded social practices that are uncondusive to weeding them out. We call these issues “conceptual woes.”

Psychologists constantly debate the validity of their concepts. Good psychological concepts are supposed to explain current behavior, predict

future behavior or behavior in different contexts, and be distinct from one another (Feest, 2020). For example, the social psychology concept “grit” (i.e., perseverance toward long-term goals when faced with obstacles) is not intended to *describe* the fact that some people succeed in spite of difficulties but to *explain* in part why they do and to *predict* who is likely to (Duckworth et al., 2007). For proponents, grit should be in our cognitive ontology because (1) its measures (e.g., the Grit Scale) are independent of, and have independent predictive value compared to, measures for other personality traits (e.g., self-control), and (2) the construct predicts patterns of performance in multiple laboratory tasks and in real-world behavior. Common reasons for jettisoning psychological concepts include a lack of ecological validity (laboratory measures do not predict real-world behavior) and or redundancy with existing concepts. For example, Vazsonyi and colleagues (2019) found that scales for measuring grit and self-control have basically the same power for predicting success—thus, beyond terminology, “grit might be indistinguishable from self-control” (p. 224).

One would hope that the routine, iterative process of construct validation (Sullivan, 2016; Feest, 2020) would converge on a plausible cognitive ontology. But many have recently argued that current practices in cognitive psychology are un conducive to this goal of conceptual progress (Poldrack, Kittur, et al., 2011; Sullivan, 2017; Quesque & Rossetti, 2020). One worry is a lack of consistency in the terms and concepts used to describe cognitive processes. A second concern is a lack of consistency in the tasks used to study those putative processes and a lack of engagement with whether those tasks actually tap into the process in question. A final concern is that the terms and concepts psychologists employ are borrowed from prescientific folk psychology or other outdated psychological theories (Poldrack, 2010).

First, there is significant heterogeneity in the terms and concepts used to describe cognitive processes (Sullivan, 2017). Poldrack and Yarkoni (2016) claim that unlike other sciences, “cognitive neuroscience is awash in a sea of conflicting terms and concepts” (p. 588). Psychologists frequently use different terms to refer to the same cognitive process. For example, social cognition researchers use a motley of terms such as “mindreading,” “mentalizing,” “theory of mind,” “perspective taking,” “cognitive empathy,” and so on to denote the same hypothetical construct (Quesque & Rosetti, 2020). The same term can also have wildly diverging meanings—for example, the term “empathy” has more than forty different definitions and refers to at

least nine putative cognitive constructs (Quesque & Rossetti, 2020). Poldrack, Kittur, and colleagues (2011) highlight the diverging meanings of “working memory,” which is sometimes construed as manipulating information in an online fashion, sometimes merely holding that information, and other times as memory for “temporally varying” aspects of a task (p. 1).

Second, there is significant heterogeneity regarding the experimental tasks used to study cognitive concepts. Researchers frequently use different experimental tasks to measure the same cognitive process—for example, the go/no-go task (a person is given a standing rule to respond on certain trials and hold back on others), and the stop-signal task (a person is instructed to respond on trials until a stop signal) are both used to measure response inhibition (Swick et al., 2011; Morein-Zamir & Robbins, 2015). Likewise, researchers sometimes use the same task to measure different putative cognitive processes. In different laboratories, the Stroop task is used to measure response inhibition, response selection, and conflict detection.

Conceptual progress in cognitive neuroscience requires linking task performance in experimental settings (e.g., performance on a go/no-go task) to cognitive processes (e.g., response inhibition) and functional neuroanatomy (the brain areas involved in the task). These links must be made within individual experiments but also across various laboratories with different experimental protocols (Sullivan, 2016). Psychologists’ tendency to use terms, concepts, and experimental tasks interchangeably creates barriers to integrating findings across different laboratories and subdisciplines. If one group uses the Stroop task to study “conflict detection” and another group uses it to study “response inhibition,” then there could be surprising neuroanatomical and behavioral overlap between these (allegedly) distinct cognitive kinds. Avoiding this problem requires task analysis—that is, articulating and testing different theories of the cognitive processes involved in the task (Sullivan, 2016; Poldrack, 2010).

In some cases, researchers have likely not undertaken these task analyses because there is no widely used repository for keeping track of these proposed connections. Where these relationships are explicitly tested, the data are troubling. For example, Swick and colleagues (2011) report that two popular response inhibition tasks—the go/no-go task and the stop-signal task—elicit distinct patterns of brain activity. This challenges whether these tasks are truly equivalent measures of the same neurocognitive process. Poldrack, Kittur, and colleagues (2011) summarize this dire situation as

follows: “this lack of consistency in the way that tasks and concepts are treated in the literature makes it difficult to draw meaningful inferences from existing literature and limits the cumulative value of the knowledge represented in this literature” (p. 2).

Finally, some worry that our cognitive ontology is suspect because it derives from folk psychology and other prescientific cognitive theories (Bunzl et al., 2010; Poldrack, 2010). Poldrack (2010, p. 754) poses a vivid thought experiment: suppose phrenologist Franz Gall had fMRI at his disposal. Poldrack (2010, p. 754) claims Gall could have performed fMRI investigations of his phrenological faculties (e.g., sense of property, poetic talent) under contemporary labels for actual neuroimaging experiments (e.g., hoarding behavior, generation of creative vs. uncreative narrative). The narrow warning here is we should not reify a cognitive kind in our ontology merely because something in the brain “activates” during tasks meant to recruit it. The broader concern is that we should distrust our cognitive ontology because it comes from folk psychology rather than neuroscience. We revisit this claim in section 14.3.

14.2.2 Why Our Cognitive Ontology Is Mistaken:

No One-to-One Mappings

The second main argument for large-scale taxonomic revision is the empirical observation that our current cognitive kinds do not perform well at predicting or explaining why particular brain structures are, or are not, involved in various cognitive tasks.⁴ Some claim that if our cognitive ontology is correct, we should expect one-to-one mappings between those categories and brain structures—for example, regions, networks, or activation patterns (Price & Friston, 2005; Poldrack, 2010). As Poldrack (2010) puts it, “correctness of the ontology would be reflected in selective associations between structures and functions” (p. 754). The assumption here is that since the brain carries out cognition, we should expect to map items in a cognitive ontology onto brain structures. Likewise, we should be suspicious if our ontology contains an item that cannot be mapped onto some element of our neural ontology. But here we have a major problem: “a review of the neuroimaging literature suggests that selective association between mental processes and brain structures is currently impossible to find” (Poldrack, 2010, p. 274). We call this the problem of no one-to-one mappings.

A traditional view holds that each brain structure is associated with a single mental process (e.g., the amygdala is a “fear area”). Indeed, the assumption that “each mental entity should correspond to a single neural entity (and vice versa) is deeply embedded in cognitive neuroscience” (Viola & Zanin, 2017, p. 947). However, these simple mappings are not what neuroimaging studies show. Instead, there are one-to-many mappings in which each brain structure is associated with multiple cognitive processes (Poldrack, 2006; Anderson, 2010; Klein, 2012; McCaffrey, 2015; Burnston, 2016) and many-to-one mappings in which each cognitive process is associated with several different brain areas (Pessoa, 2014). The term “many-to-one” encompasses both the idea that cognitive processes map onto brain networks rather than regions (Pessoa, 2014) and the concept of neural degeneracy—that is, different brain structures can perform the same cognitive process (Price & Friston, 2002). Brain structures are typically associated with multiple cognitive processes, and cognitive processes are typically associated with multiple brain structures.

One way forward is to revise our cognitive ontology. Several researchers have proposed that if we altered our cognitive ontology to capture better what brain structures actually contribute to cognition—for example, in computational or mechanistic terms—we would achieve selective or one-to-one mappings (Price & Friston, 2005; Poldrack, 2010; Shine, Eisenberg, & Poldrack, 2016). The case of the so-called visual word form area exemplifies this proposal. Price and Friston (2005) claim that while the posterior lateral fusiform gyrus (pLFG) is reputed to process the visual form of written characters, it performs (or contributes to) many other functions such as reading by touch and processing non-word visual stimuli. They argue that the pLFG is not a word area or a visual area, but rather one that performs a kind of sensorimotor integration (p. 267).

Broca’s area, once thought to be specialized for speech production (i.e., control of muscles involved in articulating speech), is implicated in myriad functions, including tool use, comprehending arm gestures, syntax comprehension, and even musical syntax (Tettamanti & Weniger, 2006; Gentilucci & Volta, 2008). Tettamanti and Weniger (2006) claim that Broca’s involvement in these seemingly disparate cognitive processes is explained by the fact that all of them involve representing hierarchical structures—that is, Broca’s area computes the kinds of hierarchical structures that are common

to musical syntax, comprehending grammatical rules, and orchestrating speech.

There are proposals to revise the functions of many regions beyond the traditional psychological labels commonly ascribed to them (McCaffrey, 2015). Examples include claims that the pIFG is for sensorimotor integration (Price & Friston, 2005), Broca's area is a hierarchical processor (Tettamanti & Weniger, 2006), the anterior insula is a salience detector (Menon & Uddin, 2010), the parahippocampal cortex functions in contextual processing (Aminoff, Kveraga, & Bar, 2013), and so on. For these authors, while our current ontology fails to find one-to-one mappings between brain structures and cognitive processes, we can revise our ontology to achieve them. Proponents of large-scale cognitive ontology revision (e.g., Poldrack, Kittur, et al., 2011; Anderson, 2014) think we should revise the bulk of our cognitive ontology in this way. Klein (2012) objects that this proposal results in vague functional labels that are poor substitutes for cognitive terms, worrying that sensorimotor integration could describe what any part of the brain does, and hence fails to make specific predictions about the pIFG. While we agree in practice, in principle, we think terms such as "sensorimotor integration" or "salience detection" are placeholders for some as yet discovered either mechanistic or computational contribution to cognition.⁵

Some resist this conclusion, suggesting that these misaligned mappings will realign if we focus on the correct brain structures. Some authors have responded to the apparent multi-functionality of brain areas by arguing that we need to zoom in by dividing the region into functionally distinct areas (Scholz et al., 2009) or zoom out to map functions onto large-scale brain networks (Pessoa, 2014) to achieve one-to-one mappings. Increasing precision about our neural ontology is important, since there is no guarantee that traditional anatomical divisions (e.g., Brodmann areas) are the most cognitively interesting ones. But these simple fixes are unlikely to achieve neat one-to-one mappings, since multi-functionality is observed at many scales of brain organization.⁶ Large-scale brain networks also appear to be multifunctional just as individual regions are (Pessoa, 2014; Viola, 2017).

Some authors attempt to resolve the problem of no one-to-one mappings by revising our assumptions about neural functioning rather than or in addition to our cognitive ontology (Klein, 2012; Anderson, 2014; McCaffrey, 2015; Burnston, 2016; Viola, 2017; Hutto, Peeters, & Segundo-Ortin, 2017). In particular, we should abandon the idea that each brain structure

performs a single function. For example, Burnston (2016) argues that brain regions perform different functions and computations in a context-sensitive fashion. The middle temporal visual area (MT) is usually thought to process motion. However, as Burnston notes, MT is also involved in processing color, fine depth, texture, and other visual properties. While a proponent of computational specificity (e.g., Shine et al., 2016) might argue that we need to revise our cognitive ontology to include some common computation MT is performing, Burnston argues that models of MT suggest different computations underlying these distinct visual properties.

We examine these calls for neurofunctional revision more closely in section 14.3. The main point, for revisionists, is that failures to map elements of our cognitive ontology neatly onto the brain supply evidence that our cognitive ontology is flawed. Thus, we need to revise our cognitive ontology if we hope to predict brain activation patterns on the basis of cognitive functioning or to describe correctly what brain regions contribute to cognition (Price & Friston, 2005; Poldrack, 2010; Rathkopf, 2013).

14.2.3 Proposed Remedies

Here, we briefly review some work from this burgeoning literature that makes concrete recommendations for resolving the problems with our cognitive ontology that were outlined above (Price & Friston, 2005; Sabb et al., 2008; Lenartowicz et al., 2010; Poldrack, Kittur, et al., 2011; Anderson et al., 2013; Yeo et al., 2015; Poldrack & Yarkoni, 2016).

Efforts to articulate our existing cognitive ontology range from local, informal pleas to use terms and concepts consistently—for example, Quesque and Rossetti (2020) in social cognition research—to formal, global efforts to chart our whole ontology in a searchable database. For example, the *Cognitive Atlas* (Poldrack, Kittur, et al., 2011) is a wiki-inspired infrastructure that aims to depict the current ontology of cognitive neuroscience. One goal of the *Cognitive Atlas*, which currently charts 868 cognitive concepts, 775 tasks, and proposed relationships between them (is a, is part of, is measured by, etc.), is to facilitate data aggregation and collaboration between different laboratory groups by providing “a systematic characterization of the broad range of cognitive processes” (Poldrack, Kittur, et al., 2011, p. 2).

Some researchers have tested cognitive ontologies directly against fMRI data in a top-down fashion. By “top-down,” we mean that they first articulated a cognitive ontology and then tested its elements using neuroimaging

methods. In a widely discussed study, Lenartowicz and colleagues (2010) tested an ontology of cognitive control against fMRI data. They performed a meta-analysis of hundreds of fMRI studies of cognitive control constructs such as working memory, response selection, response inhibition, and task switching. They trained a machine learning classifier on this data set to see if it could reliably discriminate patterns of brain activation associated with each construct. If the classifier can discriminate between the brain activation patterns for two cognitive kinds (e.g., working memory and response inhibition), this suggests they correspond to distinct neural processes. If the classifier cannot discriminate the patterns, this suggests the cognitive kinds may be different labels for the same neural process. Interestingly, Lenartowicz and colleagues (2010) report that the classifier can readily distinguish these cognitive control concepts, with the exception of task switching. The classifier was unable to reliably discriminate task switching from response inhibition and response selection. This, the authors argue, suggests that task switching is not a distinct neural process and should be removed from our cognitive ontology. Task switching may exist “only in the minds of cognitive scientists” (p. 690).

Others have used data-driven approaches to mine for novel conceptual categories that link patterns of brain activation to cognitive tasks (Poldrack, Halchenko, et al., 2009; Anderson et al., 2013; Yeo et al., 2015). Data-driven approaches often begin with data drawn from databases of task-elicited brain activity such as *Neurosynth* (Yarkoni et al., 2011) or *BrainMap* (Laird et al., 2011). Information about the cognitive processes is set aside, and the brain activation data are grouped and analyzed using dimensionality reduction techniques such as factor analysis, principal components analysis, or multidimensional scaling. Various machine learning methods, including classification (Yeo et al., 2015) and neural networks (Poldrack, Halchenko, et al., 2009), have also been used to explore alternative ways of categorizing neural data. This ideally results in a novel way of grouping together the brain data that either echoes existing categories or inspires new cognitive classifications and terms.

For example, Yeo and colleagues (2015) performed an analysis of roughly 10,000 fMRI experiments in the *BrainMap* database (www.brainmap.org). They built a mathematical model (an author-topic hierarchical Bayesian model) linking the fMRI activation patterns to eighty-three task categories (e.g., N-back task, visual pursuit task, etc.) via latent variables (which

correspond to cognitive components). The model estimated the probability that a task would recruit a cognitive component and the probability that a cognitive component would activate a particular region (set of volumetric pixels or “voxels”). The model was built to formalize mathematically the plausible notions that: (1) tasks can recruit a number of different cognitive components, (2) each cognitive component can participate in multiple tasks, and (3) cognitive components can map onto distributed brain regions. The model uncovered twelve hypothetical cognitive components that predicted their data set. These components were “discovered” by the model and do not correspond to traditional cognitive kinds. As Yeo and colleagues (2015) write, they “have refrained from explicitly labeling the cognitive components in order to not bias the readers’ interpretation” (p. 3661).

Data-driven approaches such as that those briefly discussed above are attractive vehicles for a conceptual revolution because they allow for the adoption of an agnostic stance toward the cognitive ontology. That is, the focus of these approaches is on identifying categories that have the potential to solve the problems motivating ontology revision—both conceptual woes and no one-to-one mappings—without running headlong into them. A machine learning classifier or an exploratory factor analysis procedure can, in theory, categorize data without referencing or considering the cognitive capacities and tasks used to produce it in the first place. In the next section, we will critically evaluate the prospects of large-scale ontology revision.

14.3 A Knot with Three Ends: Assessing the Case for Large-Scale Taxonomic Revision

So far, we have presented arguments for revising our cognitive ontology using the tools of neuroscience (Price & Friston, 2005; Poldrack, 2010; Anderson, 2015). Next, we evaluate the case for large-scale cognitive ontology revision. Rather than pick apart individual studies, we will address the central philosophical arguments motivating revision. We claim while the arguments presented in section 14.2 provide an important theoretical and methodological critique of cognitive neuroscience and a vital source of inspiration for moving forward, they are not universally damaging to our existing ontology. We propose that the major problems raised by revisionists—(1) our cognitive ontology is in doubt, (2) our neurofunctional ontology is in doubt, and (3) psychology needs more terminological, conceptual,

and experimental rigor—are so intertwined that blaming a lack of progress specifically on our cognitive ontology is premature.

First, we address whether our cognitive ontology *must* be flawed because it descends from folk psychology rather than neuroscience (e.g., Bunzl et al., 2010). We pose whether radical cognitive ontology revision is somehow inevitable as neuroscience progresses. Then, we discuss the tangled knot of empirical and methodological arguments motivating cognitive ontology revision. We identify a problem of mutual interdependence between refining our cognitive ontology, testing our theories of neural functioning, and facilitating terminological and conceptual consistency among different experimental groups and literatures. Many in the cognitive ontology debate hope we can hold our assumptions about the brain's functional organization and our research practices fairly constant, and adjust our cognitive ontology accordingly. Contrary to this hope, we think these problems are so interdependent that fixing one at a time (e.g., revising our cognitive ontology while holding our neurofunctional ontology fixed) or even identifying the main source of error (e.g., our cognitive ontology is what blocks us from achieving one-to-one mappings) is likely impossible.

14.3.1 Cognitive Ontology and Folk Psychology: Is Radical Revision Inevitable?

Philosophers have long pondered the fate of our folk psychological concepts as neuroscience progresses (P. S. Churchland, 1986; Hochstein, 2016; Francken & Slors, 2018; Dewhurst, 2020). Folk psychology is often regarded as a nonscientific ontology that is vague, familiar, and uninformed by cognitive science. The ongoing concern that folk psychology exerts a corrupting influence on cognitive science is reflected in Bunzl and colleagues' (2010) comment that “describing the [inferior parietal lobe] in some familiar and yet vague folk psychological terms creates a hopeless muddle of claims and agendas that get fossilized in the journals and training of graduate students” (p. 54).

The proposal that scientific progress will require replacing our current set of mental concepts with neuroscientific concepts is timely, provocative, and not entirely novel. In the 1980s, philosophers Patricia Churchland and Paul Churchland famously argued for “eliminative materialism,” which held that the folk psychological concepts we use to explain human behavior—for example, beliefs, desires, and intentions—will be discarded

for neuroscientific concepts (P. M. Churchland, 1981; P. S. Churchland, 1986). For example, the folk concept “memory” lumps together phenomena (e.g., remembering how to ride a bike, remembering that Paris is the capital of France, remembering your third birthday) that cognitive neuroscience splits via multiple memory systems (e.g., procedural memory for motor skills and semantic memory for facts; Michaelian, 2011). Thus, the folk concept of “memory” is ripe for replacement—perhaps all folk psychological concepts are too.

The Churchlands (P. M. Churchland, 1981; P. S. Churchland, 1986) argued that the nature of scientific progress virtually guarantees the replacement of folk psychological concepts by neuroscientific ones. First, they argued that folk psychology is part of a thoroughly prescientific world view, likening “beliefs” and “desires” to “miasmas” or “witchcraft.” Second, they argued that the theoretical model of scientific reduction that would preserve folk psychology is wrong.

In the traditional philosophical model of reduction, a higher-level scientific theory “reduces” to a lower-level theory when all of the entities and empirical regularities of the former can be identified with those of the latter. For instance, classical genetics would reduce to molecular genetics if hereditary patterns equate to facts about the cellular transmission of genetic material (Schaffner, 1969). This form of reduction may be folk psychology friendly. Just as identifying molecular genes need not eliminate hereditary concepts, identifying brain systems need not eliminate folk psychological kinds.

P. S. Churchland (1986) thinks this “smooth” sort of inter-theoretic reduction where higher-level entities can be identified with lower-level ones is rare. Instead, the history of science is riddled with cases where the concepts initially used to understand some phenomena were eliminated as science progresses. For instance, the term “caloric” (a hypothetical substance that passes from hotter to colder bodies) was eliminated from chemistry with the kinetic theory of heat (P. S. Churchland, 1986). The same is true of the terms “ether,” “phlogiston,” and many other once cherished explanatory posits. The Churchlands argued that neuroscientific models of behavior will supplant prescientific folk concepts such as intentions, beliefs, or desires, which will go the way of caloric, ether, and phlogiston (P. M. Churchland, 1981).

However, the Churchlands’ arguments are not decisive in retrospect. Not every prescientific concept is discarded as science advances. While the

contemporary explanation for Earth's day/night cycle differs from ancient ones (i.e., it invokes the Earth's rotation rather than gods), science has not eliminated the concept of "a day." Second, reductionism is no longer the dominant view of how higher- and lower-level sciences relate. Instead, anti-reductionist views predominate in the philosophy of mind, biology, and even physics (e.g., Fodor, 1974; Batterman, 2000).

Some calls for cognitive ontology revision are Churchlandian in spirit (e.g., Bunzl et al., 2010; Poldrack, 2010). They share with eliminative materialism: (1) worries that the concepts currently governing the science are folksy and ripe for replacement (Poldrack, 2010), and (2) an optimism that neurobiological models will provide grounds for radically revising the existing ontology (Lenartowicz et al., 2010) or starting anew (Yeo et al., 2015). Many in this literature (e.g., Bunzl et al., 2010) argue that folk psychological concepts pervade our cognitive ontology, and that their influence is detrimental to advancing our understanding of cognition.

The idea that folk psychology can and does play a productive role in cognitive science is, like the idea that folk concepts should be eliminated, not novel. Several philosophers have recently argued that folk psychological concepts can be productive participants in cognitive science research (Sullivan, 2014; Hochstein, 2017; Francken & Slors, 2018; Dewhurst, 2020). There are at least three problems with the suggestion that folk psychology corrupts our cognitive ontology. First, human cognitive agents possess internal subjective states, which are what folk concepts aim to capture. Theories that fail to account for these internal states ultimately fail to describe human capacities categorically, and so fail to describe the naturalistic human traits and capacities that cognitive science aims to explain (e.g., Francken & Slors, 2018; Dewhurst, 2020). Second, contrary to the intuition that folk psychology can only corrupt cognitive neuroscience, folk psychology—or a version of it that is open to scientific revision and reconceptualization—remains a productive part of theorizing about human and animal cognition (Sullivan, 2014; Hochstein, 2017). Third, theories that focus on the emergent level of cognition, which is often cashed out in folk psychological terms, can achieve practical outcomes in settings, such as a clinic, where more detailed mechanistic accounts that spell out the chemicals and biological pathways correlated with behavior do not (e.g., Tabb & Schaffner, 2017).

Tension between folk psychology and cognitive science is easy to find. Brain training programs that make bold promises to "change" participants'

brains through the power of “neural plasticity,” ongoing debates about whether brain scans can serve as lie detectors, or any other click-bait headline describing a finding in cognitive science provide examples of folk terms and scientific terms failing to individuate the same phenomena. It is not just in scientific reporting that these ontological frameworks are at odds. Consider the character Dory from *Finding Nemo*. Dory is a fish described in the film as having a problem with short-term memory. However, as is observed in scene after scene, Dory has no problems with working memory, as she is able to solve problems she faces and to make considered decisions. Instead, speaking scientifically, she has a problem with consolidation in long-term memory. This is a case in which a scientific term, “short-term memory,” is colloquially used in a manner that does not correspond at all with scientific research.⁷

Francken and Slors (2018) call this a “translation problem,” and locate some of the responsibility for addressing it on the shoulders of cognitive scientists. In particular, they note that a many-to-many relationship holds between common-sense cognitive concepts (i.e., folk concepts) and scientific concepts (i.e., our cognitive ontology). This is not, on their view, grounds to eliminate the folk from cognitive science, but instead grounds for cognitive scientists to reevaluate how their research relates to the everyday concepts of which it purports to advance understanding.

A central goal of cognitive science is to clarify how concepts studied in the laboratory, such as response inhibition and task switching—relate to real-world mental phenomena such as multitasking while driving or impulsive behavior in addiction. Folksy concepts often delineate the phenomena in the real world that cognitive scientists aim to understand better through controlled laboratory experiments. Indeed, research in cognitive science is often motivated by and framed around real-world examples that lean on common sense, or folk psychological, intuitions about cognitive capacities. A talk about response inhibition may open with a colloquial example of impulsive shopping or stopping a car when a child runs out into the road.

These examples are not just rhetorical tools used to prime other scholars to think about a tricky scientific question or to convince funding agencies of the relevance of research. They can, and often do, become entrenched cases around which intuitions, hypotheses, and theories are refined and evaluated. For example, recognition memory, which is the capacity to identify previously encountered objects, people, or situations as such, is often

treated as having two sub-component processes: recollection (or remembering) and familiarity (or knowing). Recollection is described as the capacity to retrieve details associated with the previous experience of the object, person, or event in question. Familiarity, on the other hand, merely delineates the feeling that the target has been previously encountered, and does not include the additional contextual details that may allow one to “place” the memory. This distinction, and empirical methods for investigating its biological realizers, is frequently motivated by appeal to the “butcher on the bus” scenario:

Consider seeing a man on a bus whom you are sure that you have seen before; you “know” him in that sense. Such a recognition is usually followed by a search process of asking, in effect, Where could I know him from? . . . Eventually the search may end with the insight, That’s [sic] the butcher from the supermarket! (Mandler, 1980, pp. 252–253).

This observation leads us toward the second reason folk psychology should not be hastily jettisoned from cognitive science: it is a framework with considerable empirical utility.

Sullivan (2014), through a detailed analysis of the history of the Morris water maze in rodent memory research, argues that establishing the reliability of an experiment, and developing and refining tasks more generally, “for individuating a discrete cognitive function requires a consideration of ‘what’ an organism trained in the paradigm is learning” (p. 58). She proposes that this is done by attributing beliefs, desires, and intentions to that organism. Furthermore, she notes that ascribing internal cognitive states to research subjects (by adopting the “intentional stance”) also assists researchers with inferring from observations of behavior to claims about cognitive functions.⁸

In a similar vein, Hochstein (2017) defends the use of folk psychology (specifically, the ascriptions of propositional attitudes to research subjects) by arguing that folk psychological concepts are indispensably used to frame comparative psychology research—for example, disputes about whether chimpanzees or scrub jays have theory of mind concern whether one animal can *know* what another animal *knows* or what another animal *intends* (italics denote folk concepts). As Hochstein (2017) notes, the fact that folk concepts are interwoven into the linguistic practices of researchers is not evidence that they are indispensable. Nonetheless, “[t]o be dismissive of such theories simply because of the terminology they choose to invoke would be to cut ourselves off from decades of psychological research that

we know empirically provides exactly the sort of information that neuroscientists need in order to refine and improve their models” (p. 1142).

If the above arguments are correct, then folk psychological concepts are motivation for, targets of, and supporting scaffolds for research in cognitive science. The scientific and folk ontologies are not in direct conflict, and may even be symbiotically related. When ontologies of the mind are viewed pluralistically, this situation is less surprising. After all, different ontologies serve different aims. Where the ontology of cognitive science has been developed to individuate and explain human cognitive capacities, folk psychology is used to predict and explain the behavior of whole organisms (Dewhurst, 2020). Tabb and Schaffner (2017) note that, at least in the case of schizophrenia research and treatment, different ontologies may be more important in different contexts. They notice that while folk psychiatry categorizes patterns most readily visible in a clinical context (after all, it is through the lens of folk psychiatry that patients view themselves), the aims and circumstances of laboratory research may be better facilitated by different ontological commitments (p. 355).

Even if our cognitive ontology contains folk concepts, nothing guarantees their elimination as neuroscience progresses. Furthermore, the targets of cognitive ontology revision are the empirical posits of cognitive psychology (e.g., response inhibition) rather than true folk concepts (e.g., impulsivity). Though some psychologists study familiar mental concepts—for example, lust, attachment, and romantic love (Fisher et al., 2002)—other terms and concepts—for example, “priming,” “semantic memory,” or “bottom-up attention”—may be alien to most people. And folksy terms such as “working memory” often refer to elaborate cognitive models (Baddeley & Hitch, 2019) rather than self-reported experiences. Our general point here is that the mere fact that a cognitive concept has folk psychological roots does not corrupt its scientific value. Furthermore, the persistence and utility of folk psychological concepts and categories strongly suggests that instead of calling for the elimination of folk psychology, those concerned with the ontology of cognitive science ought to pay more careful attention to the nuanced relationship between folk cognitive concepts and scientific ones.

14.3.2 Mutual Interdependence and the Case for Revision

Shorn from philosophical arguments about the nature of scientific progress and the explanatory deficiencies of folk psychology, claims that our

psychological taxonomy will be replaced with neuroscientific concepts reflect a stance rather than a guarantee. We concede that brain research may spur extensive taxonomic revision. Psychology has made and will continue to make taxonomic progress. Perhaps our current ontology will someday seem as provincial and unscientific as Gall's phrenological faculties strike contemporary readers. Neuroscience will likely inform this revisionary process. Although philosophers usually resist the idea of psychoneural reduction, the mere fact that "evidence from neuroscience is *relevant* to models of cognition is sufficient to license the thought that it could be used to revise psychological constructs" (Anderson, 2015, p. 70, emphasis added). The work reviewed above demonstrates that neuroscience can productively challenge our cognitive theories and models.

So, we agree that cognitive ontology revision is a worthy project and that the existing literature raises foundational issues for improving cognitive neuroscience (Poldrack, 2010; Sullivan, 2016, 2017; Anderson, 2015). Nevertheless, we think the major arguments for cognitive ontology revision—no one-to-one mappings and conceptual woes—fail to establish that our cognitive ontology needs massive revision. For one, there are limitations with existing proposals.

Consider attempts to identify novel cognitive constructs by mining task-based patterns of brain activation (Poldrack, Halchenko, et al., 2009; Yeo et al., 2015). While this bottom-up methodology promises to reveal the brain's functional preferences in a manner unbiased by our existing ontology, it has significant drawbacks. First, the dimensionality reduction techniques these approaches employ do not have unique solutions in terms of the number of constructs identified. Thus, it is premature to reify these mathematical posits as real mental entities (McCaffrey & Machery, 2016). Second, the constructs are (deliberately) uninterpretable from a cognitive standpoint, meaning the constructs that maximally *predict* brain activation patterns may do less work *explaining* what is going on cognitively. But our goal is not to criticize individual studies, as the work is promising and novel, and future iterations may assuage these concerns.

Our major criticism is that the cognitive ontology debate points to such a tangle of issues that we cannot revise our cognitive ontology in a vacuum, or even pinpoint whether our existing ontology is in jeopardy. This literature raises foundational doubts about: (1) our practices for mapping cognitive kinds onto neural structures, (2) our cognitive ontology, and (3) the

terminological and conceptual rigor of experimental psychology. The overall situation is knotted indeed. But since these problems are mutually interdependent (resolving one requires simultaneously grappling with the others), we cannot definitively say that our cognitive ontology is the problem. For all we know, our cognitive ontology will emerge fairly unscathed while we massively revise our views of how the brain carries out psychological functions (Anderson, 2014; Burnston, 2016) and/or our practices for linking terms to constructs and constructs to tasks (Figdor, 2011; Burnston, 2016).

Neurofunctional revision or taxonomic revision? The fact that brain structures are implicated in multiple cognitive processes (and vice versa) is widely considered evidence of deficiencies in our ontology, since existing concepts fail to capture what brain areas are doing (Price & Friston, 2005; Anderson, 2010). But as many authors (e.g., Klein, 2012; McCaffrey, 2015; Viola, 2017) note, neurofunctional revision (revising our views of the brain's functional organization) may undercut the case for ontological revision. For example, Burnston (2016) argues that brain regions perform multiple functions in a context-sensitive manner (i.e., according to neuromodulatory effects and/or their neural context). But if areas can perform multiple functions, then a region's involvement in multiple functions is not itself an indictment of our existing ontology. The multi-functionality of brain areas only weighs against our ontology if we assume the area performs some unified computation or function (Anderson, 2010; Shine et al., 2016), but this assumption is debatable (Anderson, 2014; Viola, 2017). These problems only worsen if we consider more radical models in which cognitive processes are carried out by different coalitions of brain regions depending on the circumstances (Anderson, 2014; Hutto et al., 2017).

Neurofunctional revision or revising psychological practice? Thus, neurofunctional revision may provide means for resisting taxonomic revision. It is unclear whether we have failed to find one-to-one mappings between brain regions and cognitive functions because we have the wrong ontology or because such mappings do not exist (Anderson, 2014; Viola, 2017). But the evidence motivating neurofunctional revision is bound in concerns about psychology's conceptual practices. For example, formal meta-analyses play an important role in claiming that cortical regions are typically multifunctional (Anderson, 2010). But these meta-analyses (e.g., showing that brain areas are implicated in many different cognitive domains) compare

brain activation patterns to the cognitive terms used in particular studies. And as we have seen, psychologists deploy these terms in a heterogeneous matter. Perhaps if psychologists deployed their concepts, terms, and tasks consistently, the relationship between mental categories and neural entities would not appear so messy (Poldrack, Kittur, et al., 2011; Sullivan, 2016).

Taxonomic revision or revising psychological practice? Efforts to revise our cognitive ontology using brain data likewise depend on psychologists' task analyses and conceptual practices (Sullivan, 2016). For example, Lenartowicz and colleagues (2010) claim that since a classifier applied to brain data cannot distinguish task-switching patterns from response selection or response inhibition ones, we should question the underlying reality of task switching. But as Figdor (2011) notes, this meta-analysis pools studies using different tasks (e.g., go/no-go task vs. stop-signal task) to study their target construct (e.g., response inhibition). These tasks may not be equivalent measures of the same neural process—for example, the go/no-go task and the stop-signal task elicit distinguishable patterns of brain activity (Swick et al., 2011). Therefore, one must address the relationship between tasks and constructs while trying to address the relationship between constructs and patterns of brain activity (Figdor, 2011).

Proponents of large-scale cognitive ontology revision often claim that neuroscience suggests our ontology is mistaken. We agree that the current situation is tangled. However, assuming that fixing our ontology will fix problems such as our inability to find one-to-one mappings between brain structures and cognitive processes is like finding a particularly tangled knot with three ends and believing you know which one to pull. We think the problems with structure-function mappings, our cognitive ontology, and conceptual practices in psychology are mutually interdependent—they cannot be solved independently. The situation is like a knot with three ends that need to be simultaneously worked on. It is likely that progress in the mind–brain sciences will require simultaneously untangling these three issues rather than addressing any individually.

14.4 Atlases, Not Maps: A Case for Taxonomic Pluralism

Aligning our cognitive ontology with brain data will require simultaneously reckoning with three issues: (1) achieving consistency and conceptual clarity about mental kinds and the tasks that engage them, (2) revising our views of

the brain's functional neuroanatomy, and (3) adjusting our cognitive ontology through lumping, splitting, discovery, or elimination of members. Will progress on these three fronts achieve a unified cognitive ontology in which mental kinds and brain structures map smoothly onto one another?

The problem actually runs much deeper. Many authors implicitly or explicitly assume that there is a correct cognitive ontology in which every cognitive kind maps onto a particular brain mechanism (Poldrack, 2010; Anderson, 2015; Viola, 2017). In other words, an assumption of taxonomic monism (there is one true taxonomy of human cognition) pervades the cognitive ontology literature. For example, Anderson (2015) describes the issue as a “debate in the cognitive sciences over the *right taxonomy* for understanding cognition—the *right theory of* and *vocabulary for* describing the structure of the mind” (p. 68, emphasis added). Additionally, many researchers think the right cognitive ontology is one where each element maps onto a brain structure. Price and Friston (2005) write, “a systematic ontology for cognition would *facilitate the integration of cognitive and anatomical models* and organise the cognitive components of diverse tasks into a *single framework*” (p. 262, emphasis added).

Here, we challenge this widespread assumption of taxonomic monism, echoing and reinforcing authors advocating taxonomic pluralism in science generally (e.g., Dupré, 1995; Danks, 2015) and in cognitive neuroscience specifically (Hochstein, 2016; Sullivan, 2017). Taxonomic pluralism, applied to these issues, holds that we will need multiple cognitive ontologies to capture the diverse aims of researchers in the mind–brain sciences. Therefore, we disagree that cognitive neuroscience ultimately will, can, or even should aim to, converge on a unified cognitive ontology charting the entire mental domain.⁹

An orthodox view in philosophy holds that there is a single, correct way to classify empirical domains—for example, chemical substances, diseases, and species (Wilkerson, 1988; Boyd, 1991). According to this view, a correct taxonomy identifies a set of “natural kinds”—that is, categories that “carve nature at its joints” independently of researchers' goals and interests.¹⁰ Recent work in philosophy of science challenges this traditional picture, pointing to various ways in which taxonomic categories (1) partly depend on the goals and interests of particular researchers and (2) are not fixed by understanding the domain's metaphysical (e.g., causal or mechanistic) structure (Dupré, 1995; Craver, 2009; Danks, 2015; Hochstein, 2016; Sullivan, 2016, 2017; Plutynski, 2018).¹¹ Next, we review these developments

in the philosophy of science (section 14.4.1) and illustrate with examples how they motivate goal-dependent pluralism about cognitive ontologies (section 14.4.2).

14.4.1 Goal Dependence and Pluralism about Cognitive Ontology

A perennial debate in philosophy concerns the status of natural kinds (Wilkerson, 1988; Boyd, 1991; Khalidi, 2013). Natural kinds are allegedly categories reflecting the world's real structure in a mind-independent way (i.e., the categories do not depend on human thoughts or goals). Paradigmatic cases of natural kinds include biological categories (e.g., mollusks), chemical substances (e.g., gold), and so on as opposed to conventional categories that depend on our goals, judgments, and preferences—for example, weeds are plants we don't like, and jocks are people who privilege athletic prowess. Essentialist views of natural kinds—dating to John Locke (1632–1704)—propose that while attending to superficial qualities leads to merely conventional categories (e.g., gold and silver are “sparkly metals”), discovering hidden essences (e.g., gold and silver each consist of atoms with a specific atomic number) leads to natural kinds sharing innumerable properties and supporting innumerable inferences (e.g., all members of the category “gold” have the same atomic number, melting point, conductivity, etc.; Wilkerson, 1988).

A famous example is jade (Putnam, 1975). Jade was originally classified by observable properties such as color, malleability, and so on. But jade actually consists of two separate minerals: jadeite ($\text{NaAlSi}_2\text{O}_6$) and nephrite ($\text{Ca}_2(\text{MgFe})_5\text{Si}_8\text{O}_{22}(\text{OH})_2$). Jadeite and nephrite differ in chemical composition and physical properties, including density, hardness, and refractive index (Harlow & Sorensen, 2005). A standard conclusion is that “jade” is a merely conventional category (useful for ceremonial purposes or commerce), while “jadeite” and “nephrite” are natural kinds that license numerous inductive generalizations (generalizations about what will happen if the rock is heated, scratched, subjected to pressure, etc.). The jade example shows how uncovering microstructural differences (Putnam, 1975) can motivate splitting folk taxonomic categories; another case would be dividing “fish” into “bony fishes” versus “cartilaginous ones.” One can also lump together once-separate categories after discovering deeper similarities (e.g., lumping whales, dogs, and bats into the category “mammal” based on homologies).

There are many philosophical views about what natural kinds are and how we identify them—for example, whether members of a natural kind

share microstructural essences (e.g., Putnam, 1975) or homeostatic property clusters (e.g., Boyd, 1991). What these views share is the commitment that questions of scientific taxonomy—for example, how many species of insects there are—are settled by properly understanding the metaphysical structure of the world. Thus, there is a single, correct taxonomy that carves nature at its joints.

Philosophers and cognitive scientists have employed the natural kinds framework to theorize about whether our mental categories correctly reflect the mind's structure (Hochstein, 2016). Michaelian (2011) questions whether memory can be a natural kind, since there are multiple memory systems (e.g., episodic vs. procedural memory). Barrett (2006) argues that basic emotion categories (e.g., anger and fear) are not natural kinds, given evidence against specific neural and physiological signatures for particular emotions. Machery (2005) claims that the mental kind concept is not a natural kind, since different mental representations (e.g., prototypes, exemplars, theories) housed in different brain systems actually do the work psychologists attribute to concepts.

Some authors recommend lumping or splitting cognitive kinds based on neural structures or representational features. As an example of kind splitting, Michaelian (2011) argues that memory is not a natural kind, since declarative memory and non-declarative memory are computationally distinct (the former involves representing the past, while the latter arguably does not) and have distinct neural implementations (explicit memory involves the hippocampus and associated structures, while implicit memory involves the basal ganglia). Thus, perhaps we can make interesting scientific generalizations about declarative memory but not memory as a whole (Michaelian, 2015). This is a psychological analogue to the jade versus jadeite/nephrite case. As a potential example of lumping, De Brigard (2014b) argues that recalling the past (i.e., episodic memory), envisioning the future, and pondering counterfactual scenarios rely on a common set of brain areas. For De Brigard (2014b, p. 179) this raises (though does not settle) the possibility that these disparate cognitive abilities are manifestations of one general capacity for projecting the self into different hypothetical situations.

The cognitive ontology debate essentially takes these local concerns about whether memory or emotion categories are natural kinds and applies them to the entire taxonomy of psychology. The hope is that we can articulate our cognitive ontology and then use brain data to lump, split, eliminate,

or otherwise revise its elements until we carve the mind at its natural joints (Poldrack, 2010; Anderson, 2015). Success would entail completely mapping the cognitive domain, with each concept aligning with an element of our neural ontology (Price & Friston, 2005). This taxonomy could be represented in a database akin to the *Diagnostic and Statistical Manual of Mental Disorders* or the *International Classification of Diseases*. We support these efforts, and agree that brain data are relevant to our cognitive ontology (Michaelian, 2011; Anderson, 2015; Khalidi, 2017). But we caution against the idea that revising our cognitive ontology will lead to a Periodic Table of the Mind—that is, a monolithic database representing the true structure of human cognition.

Recent philosophy of science suggests that while developing taxonomies in biology, chemistry, psychiatry, and so on requires correctly tracking the world's metaphysical structure, understanding that structure does not determine what taxonomy we should have. Instead, scientific ontologies unavoidably depend on researchers' goals and interests (Dupré, 1995; Craver, 2009; Danks, 2015; Hochstein, 2016; Plutynski, 2018). Since these goals and interests are plural, so too are our taxonomies. As Danks (2015) puts it, "the world does not come to us 'carved up' into its constituent objects, and we can 'segment' the world (through our scientific theories) in different, incompatible ways depending on our goals and purposes" (p. 3603). Therefore, "the pursuit of a unified (scientific) ontology is fundamentally misguided" (p. 3603).¹²

Consider questions about biological individuals. Some researchers envision a natural way to type or count biological individuals—for example, there is some definitive answer to whether ant colonies (e.g., with one reproductive queen, some reproductive male drones, and many sterile workers) are one big organism or many little ones (e.g., using genetic, immunological, or evolutionary criteria). Kovaka (2015) argues that instead biologists classify individuals variously based on their idiosyncratic goals and aims. For example, models of ant foraging may construe ant colonies as collections of separate individuals that make and interpret communications, while some evolutionary models consider the whole colony as effectively one individual. Taxonomic pluralism holds that neither way of thinking is strictly correct, and both are warranted in different circumstances.

Philosophers have recently imported these insights to cognitive neuroscience, arguing that taxonomic pluralism is required for meeting the science's diverse aims (Hochstein, 2016; Sullivan, 2017). Sullivan (2017)

proposes that cognitive neuroscience needs coordinated pluralism to progress. Sullivan (2017, p. 141) argues that cognitive neuroscience currently exhibits an uncoordinated or self-defeating pluralism in which researchers unwittingly have different cognitive ontologies and preferred tasks for studying their proposed cognitive elements (see section 14.2.1). While this situation hinders research, Sullivan claims that deliberate, coordinated pluralism can facilitate it.

For Sullivan, current practices can give the false appearance that researchers are studying the same cognitive concept—for example, response inhibition. In fact, different laboratories often use different tasks and protocols to study them—for example, the go/no-go task and the stop-signal task. The stop-signal task and the go/no-go task may not recruit the same cognitive process (Swick et al., 2011). Not realizing this fosters miscommunication and stymies research. But we cannot solve this problem by simply curtailing some lines of research and insisting that laboratories use the same task. Researchers want to know, for example, what role response inhibition plays in addiction and other risky behaviors (Morein-Zamir & Robbins, 2015). There are many live possibilities for how the inhibitory processes targeted by these tasks relate to real-world behavior (perhaps they each target a different process relevant to addiction, perhaps only one is relevant to addiction, but the other is relevant to other risky behaviors, etc.). Instead of cutting down lines of research, we need a pluralism in which researchers develop different taxonomies and employ different methods while building the physical structures (e.g., databases) and social structures (e.g., research initiatives) needed to facilitate communication and coordination.

One might reply that taxonomic pluralism may be useful now, but this results from our relative ignorance. Perhaps as our understanding of the brain matures, neuroscience will achieve a unified neurocognitive ontology. Hochstein (2016; see also Craver, 2009) argues that even when we understand the underlying neurobiology, pragmatic considerations necessarily factor into classifying mental kinds. Hochstein argues that classifying mental kinds (e.g., lumping or splitting cognitive concepts) based on neural mechanisms requires answering whether the mechanisms in question are the same or different. The problem is that brain mechanisms—within individuals, between individuals, and between different species—are similar and different in innumerable respects. According to Hochstein (2016), whether we highlight the similarities or the differences “depends largely on

how abstract or detailed our descriptions are, and what the interests and goals of individual scientists are when describing them" (p. 754).

Consider the neural circuits governing digestion or swimming/crawling behavior in invertebrates (e.g., crabs, lobsters, and leech). In these circuits, "similar rhythms can arise from different mechanisms" (Marder, Goeritz, & Otopalik, 2015, p. 156). Put differently, from animal to animal, the same circuit (a neural network with a largely similar oscillation pattern) can be built out of neurons with different properties connected in different ways. For most purposes, these subtle differences do not matter. Most of these circuits respond similarly to changes in temperature, neuromodulatory input, lesions, and so on. But for other cases, small differences in circuit composition critically matter. For example, whether these circuits will break down when subjected to unusually high or low temperatures depends on individual circuit properties (Marder et al., 2015, p. 158). Do different crabs have the same neural circuits? The answer is that it depends. If you are interested in whether the circuits produce a generally similar rhythm or whether the rhythm will persist under minor temperature fluctuations, the answer is yes. If you are interested in whether they produce the same exact rhythm or will persist under the same large temperature fluctuations, the answer is no. So, whether these animals have the same circuit seems to depend on what you mean by "the same" and for what purpose you are asking.

If this analysis is correct (even for animals with relatively simple nervous systems), then we should not expect studying brain mechanisms to settle questions about our cognitive ontology. Instead, pragmatic considerations will invariably weigh into our decisions about whether we are dealing with the same cognitive kinds. Next, we present two examples from cognitive neuroscience motivating taxonomic pluralism. We illustrate how neuroscientists can justifiably categorize cognitive processes differently according to their particular aims and interests.

14.4.2 Two Examples of Goal Dependence: Translational Psychiatry and Cognitive Modeling

Here, we give two kinds of examples motivating taxonomic pluralism in discussions of cognitive ontology. The first comes from studies of children with disruptive behavioral disorders (DBD) in translational psychiatry (White et al., 2016; Blair et al., 2018). This case suggests that the same behavioral disposition can arise from distinct neural underpinnings. While

this example deals specifically with the neural basis of mental disorders, the general considerations apply to attempts to link personality traits (e.g., impulsiveness or introversion) to their neural bases. The second comes from cognitive models of reading (Seghier et al., 2012; Price, 2018) and working memory (Baddeley & Hitch, 2019).

Divergent neural underpinnings of reactive aggression White and colleagues (2016) performed behavioral and neuroimaging studies designed to test the neural underpinnings of reactive aggression (aggression in response to perceived threats and provocations) in children with DBD such as conduct disorder and oppositional defiant disorder. Children with DBD tend to exhibit higher levels of reactive aggression than controls. In laboratory settings, they have a higher propensity to retaliate in a social ultimatum game, even when it is economically counterproductive (e.g., if retaliation leads to losses). Among children with DBD, some are also high in callous-unemotional traits (e.g., they have reduced prosocial capacities such as guilt and empathy [DBD/+CU]), while others score low on callous-unemotional traits [DBD/-CU]. Previous studies demonstrated that DBD/+CU children may have heightened emotional responses to perceived threats—for example, they seem to have heightened amygdala response when shown pictures of fearful faces (Viding et al., 2012). This raises the possibility that while DBD/-CU and DBD/+CU children both have heightened rates of reactive aggression, the neural underpinnings of this tendency for reactive aggression (partly) differ between the groups.

This is precisely what White and colleagues (2016) found. During ultimatum game retaliation, DBD/-CU children exhibited heightened responses in brain regions associated with reactive aggression and threat detection in animals (e.g., the amygdala and periaqueductal gray), whereas DBD/+CU children did not. Furthermore, activity levels in these regions were predictive of the propensity to retaliate for DBD/-CU but not DBD/+CU, children. This suggests that reactive aggression arises in DBD/-CU children partly due to a heightened threat response, where DBD/+CU may undertake reactive aggression in a more cold and calculated manner.

White and colleagues (2016) conclude that these results show “differences in the underlying neurobiology of maladaptive reactive aggression” (p. 282). Furthermore, these differences may matter for clinical practice—for example, they argue that interventions designed to reduce emotional responses would only help mitigate reactive aggression in DBD/-CU children. At this juncture,

it might be tempting to suggest that DBD/–CU and DBD/+CU children are entirely distinct clinical populations with distinctive neural underpinnings. But there is more to the story. While some of the mechanisms underlying retaliatory behavior differ, others may be the same. White and colleagues (2016) suggest that while the temptation to carry out retaliation may arise by different means in CBD/–CU and CBD/+CU children, the propensity to retaliate is also partly rooted in a shared inability to appreciate the consequences of retaliation fully. They hypothesized that alterations in the ventromedial prefrontal cortex (vmPFC) may render CBD children more likely to escalate a situation without appreciating the ensuing negative consequences. Indeed, they found that connectivity patterns involving the vmPFC corresponded to a propensity to punish in CBD youths. Thus, they claim that while there are clinically important neural differences in reactive aggression in CBD/–CU and CBD/+CU youths, there is also a shared mechanism that is also a worthwhile clinical target.

Plural taxonomies in cognitive models Early cognitive models of reading involved a single pathway for articulating sounds on the basis of orthography (written characters). However, patients with impaired reading exhibited a puzzling pattern of deficits in which some patients had difficulty reading real words with exceptional phonological features (e.g., *pint*), while others had trouble reading non-words with typical phonological features (e.g., *pord*). This led to modeling two pathways from orthography to phonology: a sub-lexical pathway going from orthography to phonological rules, and a lexical pathway going from orthography to a whole-word phonological store. Patients with deficits for atypical words (e.g., *pint*) were hypothesized to have damage to the lexical pathway, while patients with deficits for non-words were hypothesized to have damage to the sub-lexical pathway (Price, 2018).

This implies that the same (at some level of description) component in a cognitive model can be carried out by different brain structures at different times—that is, retrieving phonology from orthography in healthy patients is done by at least two brain mechanisms (Price, 2018). This represents a broader pattern (called neural degeneracy; see sections 14.2 and 14.3) in which studies of neuropsychological deficits reveal that a cognitive process can be carried out by different brain structures (Price, 2018; Seghier & Price, 2018). Similarly, Seghier and colleagues (2012) found that while the parallel

letter processing required for skilled reading (measured via word identification under rapid presentation) is usually associated with the left ventral occipitotemporal cortex, some patients with damage to this area read using an alternate pathway involving the left superior temporal sulcus (STS). Interestingly, Seghier and colleagues (2012) then found that neurotypical individuals also utilize this STS pathway to varying degrees when reading.

The challenge to taxonomic monism is whether we consider retrieving phonology from orthography or parallel letter processing single items in a cognitive ontology or multiple items. On one hand, they achieve the same end in some sense. On the other hand, they are implemented in disparate neural systems (between and even presumably within individuals) that plausibly vary in computational properties. We think that for some purposes (e.g., studying memory with lists of familiar words), it does not matter whether reading involves a lexical and a sub-lexical route. But for other purposes (e.g., explaining the pattern of deficits above), it clearly matters a great deal. Once more, whether we lump or split the cognitive process in question seems to depend on what we are trying to investigate or explain.

A similar picture emerges when looking at models of working memory. Early models of working memory had components such as the central executive, the phonological loop, and the visuospatial sketchpad (e.g., Baddeley & Hitch, 1974). A neurotypical individual can attempt to rehearse short lists of numbers, words, and so on by different means, including conscious or unconscious subvocalization or visualization (Baddeley & Hitch, 2019). Does this mean that working memory is a single cognitive component or multiple ones (e.g., visual working memory, auditory working memory, etc.)? Once again, for some purposes (e.g., studying whether working memory impairments generally affect task switching), it might not matter whether people tend to use the visuospatial sketchpad or phonological loop. For other purposes, it matters very much. For example, there are tasks (e.g., verbal word repetition) that selectively interfere with the phonological loop; the same goes for the visuospatial sketchpad. Thus, questions about what tasks interfere with working memory require dividing the mechanism further (Baddeley & Hitch, 2019). To make matters even worse, there is evidence that speech and music may be processed differently in the phonological loop. Thus, some studies may want to divide the phonological loop further.

We have endeavored to show that there is no definitive answer to questions such as whether working memory is one cognitive kind or many

sub-kinds. Instead, whether we lump or split the cognitive kind in question depends on what we want to explain or investigate. If we care about some downstream interactions of working memory, we may want to consider its various implementations as a single entity. If, on the other hand, we care about what processes interfere with working memory online or about individual or clinical differences in working memory, we may need to split the kind in a more fine-grained way. One knee-jerk reply to this argument is that since every difference makes a difference, we should always split kinds when there is *any* difference in the neural mechanism. But as Hochstein (2016) and Craver (2009) point out, this has the unfavorable consequence of multiplying our ontology endlessly as we garner trivial mechanistic details. If every human hippocampus has a slightly different way of implementing episodic memory, then we need a new entry in our ontology for every brain hemisphere.

To navigate the Scylla of coarse-grained categories that obscure important differences and the Charybdis of multiplying categories in ways that obscure interesting similarities, cognitive ontologies will need to articulate the various investigational purposes for which they are constructed (Hochstein, 2016). The end result will be less like a giant map charting all of human cognition and more like an atlas that tells you what map you need at the moment. Sometimes, you need to know a region's topography and local customs in great detail; sometimes you just need a gas station.

14.5 Conclusion

In this chapter, we examined the growing literature on cognitive ontology revision, highlighting how neuroscientific theories and methods might be used to develop, shape, and refine psychology's taxonomy. Then, we challenged two of the debate's central assumptions. First, we challenged the assumption that failures to achieve one-to-one mappings between brain structures and cognitive functions necessarily imply that our ontology is mistaken. Instead, it is uncertain where the central problem lies. The numerous issues surrounding our cognitive ontology, the brain's functional organization, and conceptual practices in psychology are deeply intertwined and must be addressed jointly. Second, we challenged the assumption that achieving a single, correct cognitive ontology is possible or desirable. Instead,

we think researchers should embrace taxonomic pluralism—the need for different ontologies for different purposes (Hochstein, 2016; Sullivan, 2017).

In advocating taxonomic pluralism, we are not claiming that all taxonomies are equally worthy or that our psychological taxonomy should float free of brain science. But proposals to use neuroscience to revise our cognitive ontology are too often bound in the claim that counting brain mechanisms will reveal what categories psychology should have. Psychologists and neuroscientists have a diverse range of goals and interests. Some have clinical aims, while others want to understand typical behavior. Some are interested in individual or cultural differences, while others are interested in commonalities. Some strive to understand unique features of human cognition, while others aim to understand features shared with nonhuman animals. We doubt that a single cognitive ontology can serve these diverse purposes.

It is an exciting time to be thinking about these questions. By making the structure of our cognitive ontology (and its relationship to tasks and brain regions) explicit, initiatives such as the Cognitive Atlas can spur new studies and conversations about the nature of the human mind. We hope that initiatives such as these will increasingly incorporate concepts from diverse theoretical origins—for example, from embodied cognition, ecological psychology, non-Western psychology, and so on—and attend to the diverse predictive, explanatory, and clinical goals of the mind–brain sciences. Doing so may facilitate cross talk between siloed corners of cognitive science, and make our theories of social and moral behavior more neurally plausible.

Notes

1. By debates about the structure of cognition, we mean disputes about broad theoretical frameworks for understanding cognition, such as embodied approaches, predictive processing approaches, etc. vs. standard representationalist views (Janssen et al., 2017).
2. Sigmund Freud (1855–1939), the founder of psychoanalysis, divided the mind into three parts: the id, the ego, and the super-ego. In his cognitive ontology, the id (an instinctual drive for sexual and bodily pleasure) and its repression by the super-ego (an internalization of social rules) explains numerous psychological phenomena from dream symbols to mental disorders. Besides psychoanalysts, psychologists no longer use this concept. Response inhibition—the ability to resist an impulse (e.g., to suppress a desire to look toward a blinking light)—is an element of cognitive control (control over one's thoughts and behaviors) in contemporary cognitive ontologies.

3. Franz Gall is most known for founding phrenology—the belief that studying skull “bumps” reveals someone’s intellectual and moral character. Despite his reputation for promulgating pseudoscience, he presciently advocated the ideas that mental functions localize to parts of the cerebral cortex and that studying brain lesions informs theories of mental faculties.
4. For example, Anderson (2010) reports meta-analyses in which every cortical area is associated with numerous cognitive functions (e.g., reading, mathematics, etc.) spanning numerous domains of cognition (vision, perception, memory, etc.).
5. Researchers express this notion differently. Anderson (2010) claims that multi-functional brain regions contribute a common working to different traditionally defined psychological processes. Rathkopf (2013) describes a functional label for a brain area that does not reference psychological task as its intrinsic function. Shine and colleagues (2016) claim that regions contribute the same computation to different cognitive tasks. See McCaffrey (2015) for a detailed discussion.
6. Anderson (2014, chapter 1.4) argues that multi-functionality is observed in small human cortical regions as well as small neural circuits in invertebrates.
7. Thanks to Felipe De Brigard for this delightful example.
8. Dennett (1987) coined the notion of the intentional stance to capture how folk psychology aids in the evaluation and prediction of behavior. To adopt the intentional stance is to ascribe beliefs and desires to a target in order to make predictions about how the target will act or behave (see Dennett, 1987).
9. Our view echoes De Brigard’s (2014a) point that “the organizational principles of the brain might not mirror the categories we use to describe the mind’s many functions. The brain is not an atlas of the mind” (p. 43).
10. The idea that our best theories “carve nature at its joints” comes from Plato (*Phaedrus*, 265e1).
11. Interested readers should consult Dupré (1995) and Plutynski (2018) for examples of goal-dependent taxonomies in evolutionary biology and medicine.
12. As Danks (2015) notes, taxonomic pluralism does not entail that objects come into and out of existence as we classify them, or that all classification schemes are equal. Instead, it recognizes that dividing the world in different ways can support different but equally useful inductive inferences, depending on the context.

References

- Aminoff, E. M., Kveraga, K., & Bar, M. (2013). The role of the parahippocampal cortex in cognition. *Trends in Cognitive Sciences*, 17(8), 379–390.

- Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, 33(4), 245–266.
- Anderson, M. L. (2014). *After phrenology: Neural reuse and the interactive brain*. Cambridge, MA: MIT Press.
- Anderson, M. L. (2015). Mining the brain for a new taxonomy of the mind. *Philosophy Compass*, 10(1), 68–77.
- Anderson, M. L., Kinnison, J., & Pessoa, L. (2013). Describing functional diversity of brain regions and brain networks. *NeuroImage*, 73, 50–58.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic press.
- Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex*, 112, 91–106.
- Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science*, 1(1), 28–58.
- Batterman, R. W. (2000). Multiple realizability and universality. *The British Journal for the Philosophy of Science*, 51(1), 115–145.
- Blair, R. J. R., Meffert, H., Hwang, S., & White, S. F. (2018). Psychopathy and brain function: Insights from neuroimaging research. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 401–421). New York: Guilford Press.
- Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61(1), 127–148.
- Brooks, G. P. (1976). The faculty psychology of Thomas Reid. *Journal of the History of the Behavioral Sciences*, 12(1), 65–77.
- Bunzl, M., Hanson, S. J., & Poldrack, R. A. (2010). An exchange about localism. In S. J. Hanson & M. Bunzl (Eds.), *Foundational issues in human brain mapping* (pp. 49–54). Cambridge, MA: MIT Press.
- Burnston, D. C. (2016). A contextualist approach to functional localization in the brain. *Biology and Philosophy*, 31(4), 527–550.
- Churchland, P. M. (1981). Eliminative materialism and propositional attitudes. *The Journal of Philosophy*, 78(2), 67–90.
- Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind–brain*. Cambridge, MA: MIT Press.
- Craver, C. F. (2009). Mechanisms and natural kinds. *Philosophical Psychology*, 22(5), 575–594.

- Danks, D. (2015). Goal-dependence in (scientific) ontology. *Synthese*, 192(11), 3601–3616.
- De Brigard, F. (2014a). The anatomy of amnesia. *Scientific American Mind*, 25(3), 39–43.
- De Brigard, F. (2014b). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese*, 191(2), 155–185.
- Dennett, D. (1987). *The intentional stance*. Cambridge, MA: Bradford.
- Dewhurst, J. (2020). Folk psychological and neurocognitive ontologies. In F. Calzavaran & M. Viola (Eds.), *Neural mechanisms: New challenges in the philosophy of neuroscience* (pp. 311–334). Heidelberg: Springer.
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087.
- Dupré, J. (1995). *The disorder of things: Metaphysical foundations of the disunity of science*. Cambridge, MA: Harvard University Press.
- Feest, U. (2020). Construct validity in psychological tests—The case of implicit social cognition. *European Journal for Philosophy of Science*, 10(1), 4.
- Figdor, C. (2011). Semantics and metaphysics in informatics: Toward an ontology of tasks. *Topics in Cognitive Science*, 3(2), 222–226.
- Fisher, H. E., Aron, A., Mashek, D., Li, H., & Brown, L. L. (2002). Defining the brain systems of lust, romantic attraction, and attachment. *Archives of Sexual Behavior*, 31(5), 413–419.
- Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, 28(2), 97–115.
- Francken, J. C., & Slors, M. (2018). Neuroscience and everyday life: Facing the translation problem. *Brain and Cognition*, 120, 67–74.
- Freud, S. (1989). The ego and the id. *TACD Journal*, 17(1), 5–22. (Original work published 1923)
- Gall, F. J. (1835). *On the functions of the brain and of each of its parts: With observations on the possibility of determining the instincts, propensities, and talents, or the moral and intellectual dispositions of men and animals, by the configuration of the brain and head* (Vol. 1). Boston, MA: Marsh, Capen & Lyon.
- Gentilucci, M., & Volta, R. D. (2008). Spoken language and arm gestures are controlled by the same motor control system. *Quarterly Journal of Experimental Psychology*, 61(6), 944–957.

- Harlow, G. E., & Sorensen, S. S. (2005). Jade (nephrite and jadeitite) and serpentinite: Metasomatic connections. *International Geology Review*, 47(2), 113–146.
- Hochstein, E. (2016). Categorizing the mental. *The Philosophical Quarterly*, 66(265), 745–759.
- Hochstein, E. (2017). When does “folk psychology” count as folk psychological? *The British Journal for the Philosophy of Science*, 68(4), 1125–1147.
- Hutto, D. D., Peeters, A., & Segundo-Ortin, M. (2017). Cognitive ontology in flux: The possibility of protean brains. *Philosophical Explorations*, 20(2), 209–223.
- Janssen, A., Klein, C., & Slors, M. (2017). What is a cognitive ontology, anyway? *Philosophical Explorations*, 20(2), 123–128.
- Khalidi, M. A. (2013). *Natural categories and human kinds: Classification in the natural and social sciences*. Cambridge: Cambridge University Press.
- Khalidi, M. A. (2017). Crosscutting psycho-neural taxonomies: The case of episodic memory. *Philosophical Explorations*, 20(2), 191–208.
- Klein, C. (2012). Cognitive ontology and region-versus network-oriented analyses. *Philosophy of Science*, 79(5), 952–960.
- Kovaka, K. (2015). Biological individuality and scientific practice. *Philosophy of Science*, 82(5), 1092–1103.
- Laird, A. R., Eickhoff, S. B., Fox, P. M., Uecker, A. M., Ray, K. L., Saenz, J. J., . . . & Turner, J. A. (2011). The BrainMap strategy for standardization, sharing, and meta-analysis of neuroimaging data. *BMC Research Notes*, 4(1), 1–9.
- Lenartowicz, A., Kalar, D. J., Congdon, E., & Poldrack, R. A. (2010). Towards an ontology of cognitive control. *Topics in Cognitive Science*, 2(4), 678–692.
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *The Behavioral and Brain Sciences*, 35(3), 121.
- Machery, E. (2005). Concepts are not a natural kind. *Philosophy of Science*, 72(3), 444–467.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87(3), 252–271.
- Marder, E., Goeritz, M. L., & Otopalik, A. G. (2015). Robust circuit rhythms in small circuits arise from variable circuit components and mechanisms. *Current Opinion in Neurobiology*, 31, 156–163.
- McCaffrey, J. B. (2015). The brain’s heterogeneous functional landscape. *Philosophy of Science*, 82(5), 1010–1022.

- McCaffrey, J. B., & Machery, E. (2016). The reification objection to bottom-up cognitive ontology revision. *Behavioral and Brain Sciences*, *39*, E125.
- Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function*, *214*(5–6), 655–667.
- Michaelian, K. (2011). Is memory a natural kind? *Memory Studies*, *4*(2), 170–189.
- Michaelian, K. (2015). Opening the doors of memory: Is declarative memory a natural kind? *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(6), 475–482.
- Morein-Zamir, S., & Robbins, T. W. (2015). Fronto-striatal circuits in response-inhibition: Relevance to addiction. *Brain Research*, *1628*, 117–129.
- Pessoa, L. (2014). Understanding brain networks and brain organization. *Physics of Life Reviews*, *11*(3), 400–435.
- Plutynski, A. (2018). *Explaining cancer: Finding order in disorder*. Oxford: Oxford University Press.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63.
- Poldrack, R. A. (2010). Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on Psychological Science*, *5*(6), 753–761.
- Poldrack, R. A., Halchenko, Y. O., & Hanson, S. J. (2009). Decoding the large-scale structure of brain function by classifying mental states across individuals. *Psychological Science*, *20*(11), 1364–1372.
- Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., . . . & Bilder, R. M. (2011). The cognitive atlas: Toward a knowledge foundation for cognitive neuroscience. *Frontiers in Neuroinformatics*, *5*, 17.
- Poldrack, R. A., & Yarkoni, T. (2016). From brain maps to cognitive ontologies: Informatics and the search for mental structure. *Annual Review of Psychology*, *67*, 587–612.
- Price, C. J. (2018). The evolution of cognitive models: From neuropsychology to neuroimaging and back. *Cortex*, *107*, 37–49.
- Price, C. J., & Friston, K. J. (2002). Degeneracy and cognitive anatomy. *Trends in Cognitive Sciences*, *6*(10), 416–421.
- 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). The meaning of “meaning.” *Minnesota Studies in the Philosophy of Science Papers*, *7* (131–193).
- Quesque, F., & Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. *Perspectives on Psychological Science*, *15*(2), 384–396.

- Rathkopf, C. A. (2013). Localization and intrinsic function. *Philosophy of Science*, *80*(1), 1–21.
- Sabb, F. W., Bearden, C. E., Glahn, D. C., Parker, D. S., Freimer, N., & Bilder, R. M. (2008). A collaborative knowledge base for cognitive phenomics. *Molecular Psychiatry*, *13*(4), 350–360.
- Scarantino, A., & Griffiths, P. (2011). Don't give up on basic emotions. *Emotion Review*, *3*(4), 444–454.
- Schaffner, K. F. (1969). The Watson–Crick model and reductionism. *The British Journal for the Philosophy of Science*, *20*(4), 325–348.
- Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E. N., & Saxe, R. (2009). Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. *PLoS One*, *4*(3), e4869.
- Seghier, M. L., Neufeld, N. H., Zeidman, P., Leff, A. P., Mechelli, A., Nagendran, A., . . . & Price, C. J. (2012). Reading without the left ventral occipito-temporal cortex. *Neuropsychologia*, *50*(14), 3621–3635.
- Seghier, M. L., & Price, C. J. (2018). Interpreting and utilising intersubject variability in brain function. *Trends in Cognitive Sciences*, *22*(6), 517–530.
- Shine, J. M., Eisenberg, I., & Poldrack, R. A. (2016). Computational specificity in the human brain. *Behavioral and Brain Sciences*, *39*, e131.
- Sullivan, J. A. (2014). Is the next frontier in neuroscience a “decade of the mind”? In C. Wolfe (Ed.), *Brain theory: Essays in critical neurophilosophy* (pp. 45–67). London: Palgrave Macmillan.
- Sullivan, J. A. (2016). Construct stabilization and the unity of the mind–brain sciences. *Philosophy of Science*, *83*(5), 662–673.
- Sullivan, J. A. (2017). Coordinated pluralism as a means to facilitate integrative taxonomies of cognition. *Philosophical Explorations*, *20*(2), 129–145.
- Swick, D., Ashley, V., & Turken, U. (2011). Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. *NeuroImage*, *56*(3), 1655–1665.
- Tabb, K., & Schaffner, K. (2017). Causal pathways, random walks, and tortuous paths: Moving from the descriptive to the etiological in psychiatry. In K.S. Kendler, & J. Parnas (Eds.), *Philosophical Issues in Psychiatry IV* (pp. 342–360). Oxford: Oxford University Press.
- Tettamanti, M., & Weniger, D. (2006). Broca's area: A supramodal hierarchical processor? *Cortex*, *42*(4), 491–494.

Tulving, E. (2007). Are there 256 different kinds of memory? In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III* (pp. 39–52). Hove, UK: Psychology Press.

Vazsonyi, A. T., Ksinan, A. J., Jiskrova, G. K., Mikuška, J., Javakhishvili, M., & Cui, G. (2019). To grit or not to grit, that is the question! *Journal of Research in Personality, 78*, 215–226.

Viding, E., Sebastian, C. L., Dadds, R. M., Lockwood, P. L., Cecil, D. A., De Brito, S. A., & McCrory, E. J. (2012). Amygdala response to preattentive masked fear in children with conduct problems: The role of callous-unemotional traits. *American Journal of Psychiatry, 169*(10), 1109–1116.

Viola, M. (2017). Carving mind at brain's joints. The debate on cognitive ontology. *Phenomenology and Mind, (12)*, 162–172.

Viola, M., & Zanin, E. (2017). The standard ontological framework of cognitive neuroscience: Some lessons from Broca's area. *Philosophical Psychology, 30*(7), 945–969.

White, S. F., VanTieghem, M., Brislin, S. J., Sypher, I., Sinclair, S., Pine, D. S., . . . & Blair, R. J. R. (2016). Neural correlates of the propensity for retaliatory behavior in youths with disruptive behavior disorders. *American Journal of Psychiatry, 173*(3), 282–290.

Wilkerson, T. E. (1988). Natural kinds. *Philosophy, 63*(243), 29–42.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods, 8*(8), 665–670.

Yeo, B. T., Krienen, F. M., Eickhoff, S. B., Yaakub, S. N., Fox, P. T., Buckner, R. L., . . . & Chee, M. W. (2015). Functional specialization and flexibility in human association cortex. *Cerebral Cortex, 25*(10), 3654–3672.