

RETHINKING THE TAXONOMY OF PSYCHOLOGY WORKSHOP
Rotman Institute of Philosophy Annual Conference
April 15 – 17, 2016
London, Ontario
#RTPWorkshop

POSTER ABSTRACTS

Vincent Bergeron (University of Ottawa)
Cognitive Homology as Cognitive Ontology

A primary goal of cognitive neuroscience is to identify stable relationships between brain structures and cognitive functions using, for example, functional neuroimaging techniques. This knowledge is important not only because it provides an understanding of the organization of cognitive functions in the brain, but also because it allows researchers to make reliable predictions about the engagement of a particular cognitive function based on the observed activation of a particular brain structure (or network of structures). Aside from the many technical, theoretical, and methodological issues that accompany this kind of research, an important empirical challenge has begun to receive widespread attention. There is mounting evidence that a great many brain structures are recruited by different tasks across different cognitive domains (Anderson 2010; Poldrack 2006), which suggests that a given brain structure can participate in multiple different functions depending on the cognitive context.

One possible reason for this lack of one-to-one mapping between brain structures and cognitive functions is that a given brain structure might do something different (i.e. perform a different operation) for each, or at least some, of the different types of cognitive functions it participates in (Anderson 2014). Another possible reason is that our cognitive ontologies—i.e. our current descriptions of cognitive processes and their associated components—are either incorrect or too coarse (Poldrack 2010; Price & Friston 2005). Brain structures might each perform one kind of operation that can be recruited (or reused) by different types of cognitive functions in different cognitive contexts (one-to-one mapping between brain structures and these local operations), but these local operations might not correspond to anything in the cognitive vocabulary of our current theories about the structure of the mind. My aim in this paper is to explore this second possibility.

We know that the human brain shares many of its principles and functions with that of other species, and that for any human cognitive function, we can expect that (at least) some component(s) of it could be found in the cognitive repertoire of another species (de Wall and Ferrari 2010). What is less clear, however, is how best to exploit this evolutionary continuity in order to identify precisely the components of the human cognitive architecture that we share with other species and that have remained stable across extended evolutionary periods. Here, I argue that a useful way to think about these shared components is to think of them as cognitive homologies. In contrast with the well known concept of structural homology in biology—defined as the same structure in different animals regardless of form and function, where sameness is defined by common phylogenetic origin—the proposed notion of cognitive homology focuses on the

functional properties of homologous brain structures that tend to remain stable across extended evolutionary periods. I then argue, using recent findings from the cognitive neuroscience of social cognition, that the identification of cognitive homologies can greatly contribute to the identification of stable structure-function relationships which, in turn, can be used for the construction of new cognitive ontologies.

Daniel Burnston (Tulane University)

Functional Integration, Neural Coding, and Consumer Semantics

Psychological models of brain function are traditionally of the “box and arrow” sort. One part of the mind performs its function, generates a representation, and transmits that representation to another part of the mind. This latter part then uses the representation to fulfill its function. On its face, this kind of picture is deeply wedded to “consumer semantics,” the view that a mental representation should be characterized in terms of the information it conveys to and contributes to the function of bits of the mind to which it is transmitted. This intuitive picture has been extremely influential in philosophy, psychology, and neuroscience.

In this talk, I will articulate some underappreciated difficulties for consumer semantics arising from current developments in systems neuroscience. The difficulty is caused by two related and increasingly recognized facts—the widespread interconnectivity in the brain and the large amount of contextual variation in the function of many (perhaps all) brain areas. Consumer semantics is based on a privileged representational/semantic reading of a causal interaction between two parts of the brain. But if each part can represent distinct types of information and/or contribute to different tasks, depending on context and interaction with other areas, it looks difficult to isolate a privileged semantic influence conveyed by the “arrow” in a box-and-arrow diagram.

I discuss two current frameworks for analyzing neural coding and representation, and argue that both are in conflict with consumer semantics. The first is “multi-variate” pattern analysis, which posits highly distributed neural representations characterized by their values along dimensions of variance between stimuli. This scheme is incompatible with consumer semantics because the dimensions that determine the structure of a representation are not themselves semantically evaluable—that is, they don’t correspond to particular features in the environment, but instead to projections across those features. As such, there are no signals with a specific representational content to be passed between areas; instead, specific representational content is an irreducibly distributed phenomenon. The second coding scheme is “multi-plexed information,” which posits that distinct representational aspects of a signal are parsed according to distinct temporal and/or spatial properties of the signal’s vehicle, such that these distinct codes might be “read off” by distinct receivers. Despite appearances, this framework is also incompatible with consumer semantics, because which aspect of a multi-plexed signal should be consumed is not determinable strictly in the relationship between two parts of the brain—other parts must intervene causally to distinguish which part of the signal is functionally relevant and when.

I then turn to evaluating what these arguments entail for thinking about neural representation and psychological decomposition. Consumer semantics attempts to employ the notion of representation to explain functional interaction between areas. I think this order of explanatory priority is wrong. Instead, we should view neural representation as the outcome of functional

interactions—a specific, task-oriented representation results from a kind of “tug-of-war” between mutually interconnected mechanisms in a specific context. I briefly discuss models of the psychological category of “decision” in leeches and mammals that illustrate this concept.

David Colaco (University of Pittsburgh)

The identification of scientific phenomena

I discuss the process by which researchers identify scientific phenomena. A cursory examination of publications in the life sciences suggests that researchers aim to relate the results of their experiments to phenomena in the world; however, while one need not look far to find a scientist talking about their results in relation to a phenomenon of interest, comparatively little research in the philosophy of science discusses, let alone analyzes, the research practice that makes this possible. With few exceptions, philosophical descriptions of science take the researchers' identification of a phenomenon as a given. Given that there are a number of instances in the fields of psychology and neuroscience where successful identification is considered an achievement in and of itself, I aim to discuss the process by which researchers produce these kinds of achievements. The identification process is a scientific activity that best addresses certain kinds of questions; as such, I investigate the strategies that researchers employ to answer these questions. While identification complements theorization, explanation, and model building, it should be thought of as distinct from each of them.

I proceed as follows. To begin, I give a review of the character of scientific phenomena, drawing on the distinction between them and data. I present two principal features of phenomena: the stability feature and the repeatability feature. Then, I discuss the general features of the identification process, and introduce my two principal case studies: research on long-term potentiation in cellular neuroscience, and the investigation of reinforcement learning in cognitive neuroscience and psychology. After introducing the cases, I outline the phases in the process, and describe the strategies that occur during each phase. In doing so, I describe how research starts, and how it can improve with additional tests. The interaction between the strategies and the products results in the process possessing a vindicatory character, which I relate to Hacking's self-vindicating laboratory sciences (1992). I further show that this interaction captures, in systematic terms, what William Bechtel and Robert Richardson call the reconstitution of phenomena (1993), and Carl Craver and Lindley Darden call the characterization of phenomena (2013). I conclude with a discussion of the outputs of the process, and contrast them to Nancy Cartwright's phenomenological models (1983). Throughout, I highlight the difference between the identification of phenomena and other scientific activities.

Joseph McCaffrey (University of Pittsburgh)

Neuroimaging and Cognitive Ontologies: Does Discriminability Warrant Kind Splitting?

Psychologists study cognition by positing a set of mental constructs such as "attention," "working memory," and "disgust." Collectively, these constructs constitute a "cognitive ontology"--i.e. a taxonomy of psychological kinds. Recently, many neuroscientists (e.g., Anderson 2010, Poldrack 2010, Lindquist et al. 2012) have argued that fMRI data compels a need for "cognitive ontology revision," a rethinking of the taxonomy of psychology. The basic idea is that brain data can be used

to assess the validity of psychological constructs. Proponents argue that fMRI can be used to: (1) identify novel cognitive constructs (e.g., Anderson 2014), (2) suggest that existing constructs, such as "anger," should be eliminated (e.g., Lindquist et al. 2012), and (3) lump or split existing categories (e.g., De Brigard 2014). Open questions include whether a significant degree of neural overlap between different constructs warrants kind lumping, or whether novel BOLD activation patterns reflect novel cognitive constructs.

My poster examines the claim that some fMRI studies warrant splitting psychological kind categories, focusing on recent studies using multi-voxel pattern analysis (MVPA) or other pattern classification techniques to analyze fMRI data. Many neuroscientists infer from pattern discriminability to the distinctness of the underlying mental constructs. For example, Lenartowicz et al. (2010) argue that since a pattern classifier can routinely distinguish neural signatures related to "working memory," from those related to "response inhibition," these labels likely reflect distinct cognitive control components. Similarly, Woo et al. (2014) argue that BOLD responses to bodily pain are distinguishable from responses to social rejection; therefore, social rejection does not involve somatic pain as some (e.g., McDonald and Leary 2005) argue. These studies rely on the assumption that different cognitive constructs elicit different BOLD patterns, and suggest (e.g., Poldrack 2010) that one's cognitive ontology is valid when each construct is associated with a selective, specific pattern of brain activation.

However, different tasks will usually elicit *somewhat* different brain activation patterns; the challenge is determining what count as *relevant* differences from a cognitive standpoint (e.g., Machery 2012). I claim that the ability of a pattern classifier to discriminate BOLD signals from one another may often depend on differences that are irrelevant to the relationships between members of one's cognitive ontology one is hoping to test. First, the same neural process operating on different inputs might result in discriminable BOLD activation patterns--I call this the problem of "implementation differences." Haxby et al. (2001) used a pattern classifier to distinguish participants looking at shoes vs. bottles, but "seeing a shoe" and "seeing a bottle" may be different manifestations of the same basic capacity for object recognition. Second, two tasks may recruit some of the same constructs, and some different ones. Thus the ability to discriminate brain scans from those tasks may depend not on the construct of interest, but on "additional constructs" recruited by the tasks. I conclude that discriminability alone should not motivate kind splitting, and explore how neuroscientists (e.g., Woo et al. 2014) can move beyond mere discriminability when using pattern classification techniques to motivate kind splitting.

Jorge Morales (Columbia University)

When Behavior Is Not Enough: Reading Metacognition From The Brain

In my poster presentation I will argue that we should use nonhuman animal neurophysiological experimental results of higher cognitive functions as a means to refine our own concepts of human higher cognition.

The study of animal psychology often has two goals. The first goal is to determine whether other animals have mental states and cognitive functions similar to ours. The second goal is to use animals as models of our own neuropsychological mechanisms. When carrying out the latter task, basic psychological functions (e.g., perception) are the usual target.

A problem with recent neuroscientific attempts for testing higher cognitive functions in other animals is that these two main goals are confounded. When testing for basic functions (e.g., vision), it is assumed that the studied animal indeed has such capacity. With higher cognitive functions, a double pass, as it were, is attempted. At the same time researchers try to determine whether the capacity is present (first goal) and also to learn from the animal's neurophysiology to understand better the mechanisms that underlie their and our own cognitive functions (second goal).

I take metacognition as a case study. Philosophers often argue that metacognition can only be carried out by having an explicit representation of other mental states (i.e., by means of a metarepresentation). Even though animal metacognition researchers have designed sophisticated experiments to show that animals can pass metacognitive tasks (Smith, 2009), alternative explanations that do not require metarepresentation could suggest the behavioral evidence is not sufficient for attributing metacognitive capacities to other animals (Carruthers & Ritchie, 2012). Hence, the behavioral evidence to date is, at least, ambiguous as to whether other animals have metacognition or not.

Despite this ambiguity, I argue that single-cell neural-recording results are useful tools for establishing functional and morphological homologies between our and other animals' psychology (Kepecs et al., 2008; Kiani & Shadlen, 2009; Rilling, 2014). By focusing on the main traits of human metacognition I argue that these can be captured by non-metarepresentational confidence judgments about perceptual and cognitive processes. Concretely, I propose that formal models of perceptual and cognitive confidence allow us to link metacognition with behavior and neural activity. To this end, I suggest that either signal detection theoretic models or evidence accumulation models of confidence allow us to link the basic traits of (human and non-human animal) metacognition with behavior and physiology.

I end with a discussion regarding the relevance of morphological and functional homology for making interspecies comparative judgments about cognitive capacities (García Aguilar, 2010). In particular, we cannot expect that primate neurophysiological findings leave our concepts of human higher cognitive functions unrevised. A potential result of achieving the second goal is, indeed, the revision of folk concepts of our own psychology (Anderson, 2015).

Sofia Ortiz-Hinojosa (MIT)

The Place and Unity of Imaginings

Imaginings are certainly one of the bugbears of mental taxonomy. Previously, scientists and authors have argued that they are either too heterogenous to comprise a single class of mental entities, that they are reducible to other mental processes that are already well understood, such as memory or inference, and even that imaginings as such do not exist. In "The Place and Unity of Imaginings," I argue that imaginings are a unified set of mental entities, none of which can be reduced to some special operation of memory, perception, or inferential reasoning. One problem for my view is that imagination is supposed to explain a number of quite dissimilar capacities, such as predicting others' mental states, engaging in fiction, playing pretend, and finding out what is possible. I show that imaginings can all be unified by a simple family resemblance theory based on familiar epistemic concepts, like truth-orientation, etiology, and method, which together yield

necessary (although not sufficient) conditions for the class of mental entities we typically call imaginings. The necessary conditions are: (1) imaginings need not aim at truth, (2) imaginings originate in states of the cognizer or actor rather than states of the world, and (3) imaginings often employ combinatorial processes. I give an argument for why it should be permissible for normative concepts to influence our taxonomy, such that these epistemic concepts and family of traits justify the taxonomic separation of imaginative states and imagining processes from other mental states and mental processes. I also show how these features are sufficient to explain all of the diverse ways we characterize and make use of distinct imaginative capacities.

Luka Ruzic, Lawrence Ngo, Greg Appelbaum, Felipe De Brigard and Scott Huettel (Duke University)

The neural components of cognition

Mapping neural processes to cognitive operations has proven particularly challenging. This occurs in part because (presumably) identical neural processes tend to be redeployed in a variety of cognitive operations. This poster reports a promising analytic strategy to help to understand the precise ways in which different basic neural components are combined to give rise to diverse cognitive operations. First, a thesaurus of non-equivalent/non-synonymous neural terms from the Neurosynth database was produced, resulting in 414 independent terms. Next, an independent component analysis (ICA) conducted across the reverse inference maps of the 414 terms yielded 65 components where each component is a pairing of a vector describing variation in the spatial domain (i.e. a “spatial map” of weights on brain voxels) and a vector describing variation in the term domain (weights indicating the relevance of the spatial map to each Neurosynth term). Term weights for each component were then sorted by relevance, the difference between terms was calculated, and a maximal differentiation threshold was employed as a cutoff for inclusion in the list of terms used to describe each component.

Components were discarded if their spatial map corresponded to white-gray matter brain junctions, if they were primarily related to individual difference terms (e.g., “women”, “adult”, “autism”, etc.), or if they were characterized as relevant to all terms except some small set (e.g., “not time, image, or response”). The remaining 50 components were categorized into general groups. The first group comprised 13 components related to sensory experience. The second group comprised 5 components related to terms such as emotion and reward. A third group comprised 4 components related to terms such as default network, mentalizing and theory or mind. A fourth group comprising 4 components loaded onto terms related to semantic and language processing. A fifth group comprised 10 components related to terms like working memory and task demands. A sixth group comprised 6 components related to motor control and somatosensory processing. A seventh group comprised 4 components related to autobiographical and episodic memory. And an eighth group comprised 2 components related to decision and learning. To corroborate this grouping, a hierarchical clustering analysis using a Nearest Point Algorithm was performed on a distance matrix of the ICA components according to their term domain weights. The resultant clustering confirmed some of the groupings and contradicted none of them.

A proposed visualization of these resultant groups is suggested, and the implications of these results for our understanding of the relationship between brain activation and cognitive processes is further discussed.

Serife Tekin (Daemen College)

Mental Disorders and Natural Kinds

Scientific investigation on mental disorders has the pragmatic goals of identifying their causes and developing strategies to effectively treat them. Philosophers of psychiatry have contemplated on mental disorders, predominantly, by debating whether psychopathology is a legitimate target of scientific inquiry and, if so, how mental disorders should be explained, predicted, and intervened on. However, as I show in this paper, these discussions have mostly neglected the actual state of inquiry in psychiatry and relevant disciplines, as well as the first person experiences of those with psychopathology.

A prominent view in the literature is that whether something can be empirically investigated depends on whether or not it is a natural kind (Cooper 2004; Samuels 2009). Members of a particular natural kind are thought to share a large number of logically unrelated and scientifically important properties that enable the formulation of successful scientific generalizations. Accordingly, if mental disorders are natural kinds, the discovery of the scientifically relevant properties that their members share should yield successful explanations, reliable predictions, and effective interventions. Thus, the scientific legitimacy of mental disorders has hinged on their status as natural kinds.

Proponents and opponents of the view that mental disorders are natural kinds participate in the Looping Debate. They inquire into the scientifically relevant properties that will facilitate explanations, predictions, and interventions in psychiatry by adopting what I call the comparative strategy. They evaluate whether mental disorders, such as schizophrenia, are “fundamentally like or unlike kinds such as chemical elements, fundamental particles and biological species” (Cooper 2004), and compare mental disorders to the paradigm examples of the natural kinds studied by other sciences to draw inferences about their empirical investigability (Hacking 1995; Cooper 2004; Zachar 2001). Proponents of the Looping Argument (LA) argue that mental disorders are not natural kinds because members of psychiatric kinds change in response to their scientific labels; whereas natural kinds are not subject to such looping effects. Being subject to looping effects make the properties of psychiatric kinds unstable, thereby unfit for developing successful explanations, reliable predictions, and effective interventions in psychiatry. Proponents of the Parity Argument (PA) object to LA, and argue that some natural kinds are also subject to looping effects, and this is not an obstacle for their scientific investigation. Thus, having unstable properties does not rule psychiatric kinds out as legitimate targets of scientific inquiry.

The target of my criticism is the claim that the comparative strategy will take us to a set of properties that achieve three tasks: explanation, prediction, and intervention in psychiatry. I show that this is false; the Looping Debate's utilization of the comparative strategy fails to take us to the properties relevant for intervention in psychiatry. I propose a trilateral strategy that examines the scientific and clinical work on mental disorders as well as first person accounts of those with such disorders. Using schizophrenia as an example, I illustrate that the trilateral strategy takes us to

properties instrumental for developing effective interventions on mental disorders –without having to settle their natural kind status.

Matthieu de Wit¹, John van der Kamp² and Rob Withagen³ (1. Moss Rehabilitation Research Institute, 2. VU Amsterdam, 3. University of Groningen and University Medical Center Groningen)

Framing neuroscientific questions through the lens of ecological psychology

Historically, ecological psychology has largely ignored findings from neuroscience, and generally with good reason. Most neuroscientific studies are guided by the mechanistic assumptions that ecological psychology disputes, which typically renders them unsuited to inform ecological theoretical development. Recent work by for example Anderson (2014) on neural reuse, which refers to the involvement of the same local regions in the brain in multiple functions across multiple behavioral domains, provides a notable exception and could provide the starting point for an ecological or Gibsonian neuroscience. We consider the concept of neural reuse as an elaboration of the notion of “vicarious functioning” or multiple realizability central to Gibson’s theory of perceptual systems and Reed’s (1982) theory of action systems. Gibson (1966) claimed that, “[t]he individual nerve or neuron changes function completely when incorporated in a different system or subsystem” (p. 56) and that, “[t]he same incoming nerve fiber makes a different contribution to the pickup of information from one moment to the next. ... The individual sensory units have to function vicariously, to borrow a term from Lashley, a neuropsychologist” (pp. 4–5). To advance our understanding of the brain and the behavior and cognitive processes to which it gives rise, we suggest consideration of a taxonomy built on Reed’s (1982) concept of action systems, which we take to be dispositional properties of organisms that enable them to establish different functional relationships with their environments—causally supported by, but not identical to, anatomical substrates (Withagen & Michaels, 2005). This is to say that an organism can typically establish the same functional relationship with its environment by means of several substrates, and that the same substrate can be used to establish several functional relationships. Initial neuroscientific evidence for this contention exists in the form of degeneracy and neural reuse in the brain.

Anderson, M. L. (2014). *After phrenology: Neural reuse and the interactive brain*. Cambridge, MA: The MIT Press.

Gibson, J.J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.

Reed, E.S. (1982). An outline of a theory of action systems. *Journal of Motor Behavior*, 14, 98–134.

Withagen, R., & Michaels, C. F. (2005). On ecological conceptualizations of perceptual systems and action systems. *Theory & Psychology*, 15(5), 603-620.

Jessey Wright (Western University)

Categorizing cognitive processes with brain data

The validity of psychological constructs is assured by verifying that the construct of interest is discriminable from other constructs using a variety of methods and measurements. Multivoxel

pattern analysis (MVPA) involves the use of machine learning classifiers to identify the behavioural and cognitive processes that can be discriminated between on the basis of neuroimaging data. The analysis procedure involves building a classifier that is trained to label neuroimaging data according to the task, behavioural or cognitive conditions of interest. Once trained, the classifier is provided with unlabelled data, and its accuracy at predicting the correct label is assessed. When a classifier can accurately predict the labels of novel neuroimaging data, patterns in the data can be used to discriminate between the conditions associated with the labels. These techniques provide a way to evaluate if constructs can be discriminated between on the basis of patterns of brain activity, and so are well positioned to provide critical feedback on the validity of psychological constructs from the perspective of the brain. However, while MVPA provides a measure of discriminability, it is difficult to interpret the analysis results.

Classifiers will leverage any features of the data that permit reliable discrimination. Thus, successful classification does not necessarily entail that the constructs are discriminable by 'the brain'. In order to interpret the results in this way one must assume that the classifier is accessing the content of the neural activity, that is the content of the signal itself. Verifying this assumption requires providing evidence that the classifier is leveraging information in the activity pattern that is available to the brain. In the context of a neuroimaging experiment, I show that this can be done by analyzing the relations between classification accuracy and a subject's performance in a behavioural task. I demonstrate this with an example from research on memory in the hippocampus.

Then, I raise this as a challenge for meta-analyses aimed at evaluating the validity of, and establishing new, psychological constructs using machine learning classifiers. These analyses, because they rely on the aggregation of neuroimaging data produced in a variety of lab contexts, cannot necessarily use correlations with behavioural performance to verify that the classifier is leveraging only information available to the brain. Absent such controls, interpreting the analysis as revealing the psychological categories preferred by the activity and organization of the brain is suspect. I conclude by briefly considering how contrasts between the training and testing phases of the analysis influence interpretations of MVPA results with respect assessments of the validity of psychological constructs.

This research was supported by the Social Sciences and Humanities Research Council of Canada. Further support for this workshop provided by a generous grant from the Canadian Institute for Advanced Research (CIFAR) Mind & Consciousness Program, and from Research Western.



Social Sciences and Humanities
Research Council of Canada

Conseil de recherches en
sciences humaines du Canada

Canada

CIFAR
CANADIAN
INSTITUTE
FOR
ADVANCED
RESEARCH



Western
The Brain and
Mind Institute

Research Western



**Rotman Institute
of Philosophy**
ENGAGING SCIENCE.