Précis of *After Phrenology: Neural Reuse and the Interactive Brain*

Michael L. Anderson

Department of Psychology, Franklin & Marshall College, Lancaster, PA 17604-3003

michael.anderson@fandm.edu
http://www.fandm.edu/michael-anderson

Abstract: Neural reuse is a form of neuroplasticity whereby neural elements originally developed for one purpose are put to multiple uses. A diverse behavioral repertoire is achieved by means of the creation of multiple, nested, and overlapping neural coalitions, in which each neural element is a member of multiple different coalitions and cooperates with a different set of partners at different times. Neural reuse has profound implications for how we think about our continuity with other species, for how we understand the similarities and differences between psychological processes, and for how best to pursue a unified science of the mind. *After Phrenology: Neural Reuse and the Interactive Brain* (Anderson 2014; henceforth *After Phrenology* in this Précis) surveys the terrain and advocates for a series of reforms in psychology and cognitive neuroscience. The book argues that, among other things, we should capture brain function in a multidimensional manner, develop a new, action-oriented vocabulary for psychology, and recognize that higher-order cognitive processes are built from complex configurations of already evolved circuitry.

Keywords: dynamic systems; embodied cognition; evolution; modularity; natural selection; neuroplasticity; pragmatism

1. Introduction

After Phrenology: Neural Reuse and the Interactive Brain (Anderson 2014; henceforth *After Phrenology* in this Précis) offers a framework for a science of psychology that harmonizes three aspects of the mind that are generally treated separately: its biological underpinnings, its situatedness in the environment, and its evolutionary history. The core of the framework is the theory of neural reuse, which posits that individual neural elements (at multiple spatial scales) are used and reused for multiple cognitive and behavioral ends. According to the principle of neural reuse, a diverse behavioral repertoire is achieved through the search for and consolidation of multiple, nested, and overlapping neural coalitions, in which each neural element is a member of multiple different coalitions and cooperates with a different set of partners at different times.

Such a neurofunctional architecture stands in stark contrast to the modularity assumption that has been a core tenet of most (although certainly not all) computational accounts of mind, and especially those derived from or influenced by evolutionary psychology (Barrett & Kurzban 2006; Carruthers 2006). Central to the overall project advanced in *After Phrenology* is a reconsideration of how best to unite psychological science and evolutionary biology. Because evolutionary psychology focuses its efforts on describing genetically encoded psychological solutions to the challenges posed in the so-called environment of selection, it expects the brain to be largely composed of special-purpose neural modules. Neither the neuroscientific nor the evolutionary evidence has borne out this expectation. What appears to better account for that evidence is a set of neurodevelopmental processes – including both Hebbian plasticity and neural reuse – that efficiently serve the adaptivity of the organism by marshaling the same limited pool of resources in different ways as tasks demand.

Neural reuse has three immediate implications. First and most obvious, newly acquired capacities are generally supported by mixing and matching the same neural elements in new ways. Second, and perhaps less obvious, neural reuse would appear to support and encourage *procedural* and *behavioral* reuse. That is, one reason neural reuse is an effective developmental strategy is that the cognitive processes that neural elements support and the behaviors they drive are in fact useful in multiple circumstances and can be marshaled to various ends. Hence, reuse has

both a biological and a behavioral aspect. And that brings us to the third implication, which follows neatly from the first two: Not every cognitive achievement – not even achievements as central to the life of a species as natural language is to ours – need be supported by a specific targeted adaptation. In fact, the principle of parsimony would appear to dictate that, ceteris paribus, we should prefer accounts that show how "higher-order" cognitive processes (such as language and mathematics) marshal existing neural resources and behavioral strategies in unique ways over accounts that posit unique adaptations.

After Phrenology outlines one such parsimonious approach to understanding mathematics and natural language. In order to understand math and language as developments of already existing resources, rather than as a particular species-specific cognitive adaptation, it is crucial to appreciate the ways in which cognition and overt behavior are mutually supporting and intertwined at multiple timescales. *After Phrenology* details these relationships at length. More pointedly: Thinking, calculating and speaking are adaptive behaviors and, as such, involve the whole organism acting in and with its environment. These capacities are not limited to, nor are they even primarily a matter of computation over, a set of mental symbol structures. Instead, thinking involves iterated interactions with elements of the environment. It leverages our highly developed and early-evolving capacities for acting in and manipulating the physical and social environment. Sociocultural cognitive achievements such as language and mathematics are extensions of – not radical departures from – these basic capacities.

All of the preceding together suggests that we may be on the cusp of a significant transformation in psychological science. The way we conceptualize cognitive function, the way we map these to supporting structures (and the range of structures that appear to be relevant supports), and the old distinctions between perception and action, action and cognition, cognition and emotion – all of that and more needs reexamination in light of emerging results. *After Phrenology* maps the terrain and charts an alternative path toward a unified biological, situated, sociocultural, evolutionary science of the mind.

2. Two kinds of neuroplasticity

The most familiar kind of neuroplasticity is Hebbian learning, also known as spike-timing dependent plasticity (Song et al. 2000). Hebbian learning is a crucial developmental process for tuning local neural interactions and helping determine the functional bias of local networks. Neural reuse, whereby individual neural elements are put to use for multiple cognitive and behavioral ends, involves an additional kind of neuroplasticity that I have called neural *search*. Neural search is a process that places neural elements into new functional partnerships with one another. During the course of learning and development, each element will come to be a member of multiple functional coalitions.

The first two chapters of *After Phrenology* are dedicated to marshaling the evidence for neural reuse in general and neural search in particular. I will just gesture at a few key pieces of evidence, here. If individual regions of the brain are in fact used and reused in multiple circumstances (as

posited by neural reuse), then they should be functionally *diverse*, active in support of multiple tasks from different task domains. If variety of function is a matter of putting the same neural elements into different functional coalitions (supported by neural search), then we should see different patterns of functional cooperation across the brain under different psychological circumstances.

In one recent study (Anderson et al. 2013), my coauthors and I borrowed a technique from ecology and measured *functional* diversity in various regions of the brain in much the same way as one measures *ecological* diversity. We asked: How many individual tasks (i.e., animals) in how many different task categories (i.e., species) does each region of the brain support? Using Shannon entropy (Shannon 1948) as our metric, and a large collection of more than 2,000 functional neuroimaging experiments, we measured functional diversity voxel-by-voxel using a spherical searchlight of various sizes. The unequivocality of the results surprised even us: Individual regions of the brain, even small regions, are highly diverse. Figure 1 reproduces the histogram of results from one run, using a 10 mm spherical searchlight and 11 task categories, and scaling the diversity metric from zero to one, such that zero diversity indicates that every observed activation is in a single category, and one indicates equal numbers of activations from each task category.

As can be easily observed from the leftward skew of the data, there are very few specialists in the brain, supporting only tasks from a single task category such as semantics or visual perception. Most regions of the brain are active during multiple tasks in different task categories. Regions nevertheless *achieve* their functional diversity in different ways. Figure 2 illustrates the *functional fingerprints* of three different voxels from the same data run described above: a voxel from left auditory cortex, a relative specialist with a diversity value of 0.41; a voxel from left anterior insula with a high diversity value of 0.88; and a voxel from left thalamus with a diversity value of 0.76, equal to the population median. Functional fingerprints display the relative degree of activity observed in each task category for the region pictured. Hence, the portion of auditory cortex shown is most frequently active during auditory tasks, and only occasionally in somesthesis, action, and language tasks. By contrast, the pictured region of anterior insula is active at least some of the time during tasks in
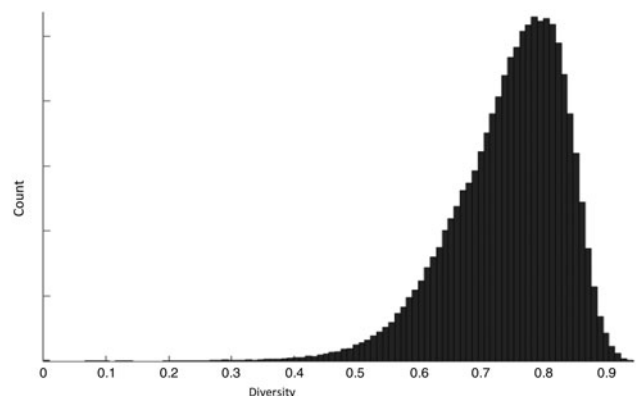

Figure 1. Histogram of whole-brain, voxel-wise functional diversity measurements.
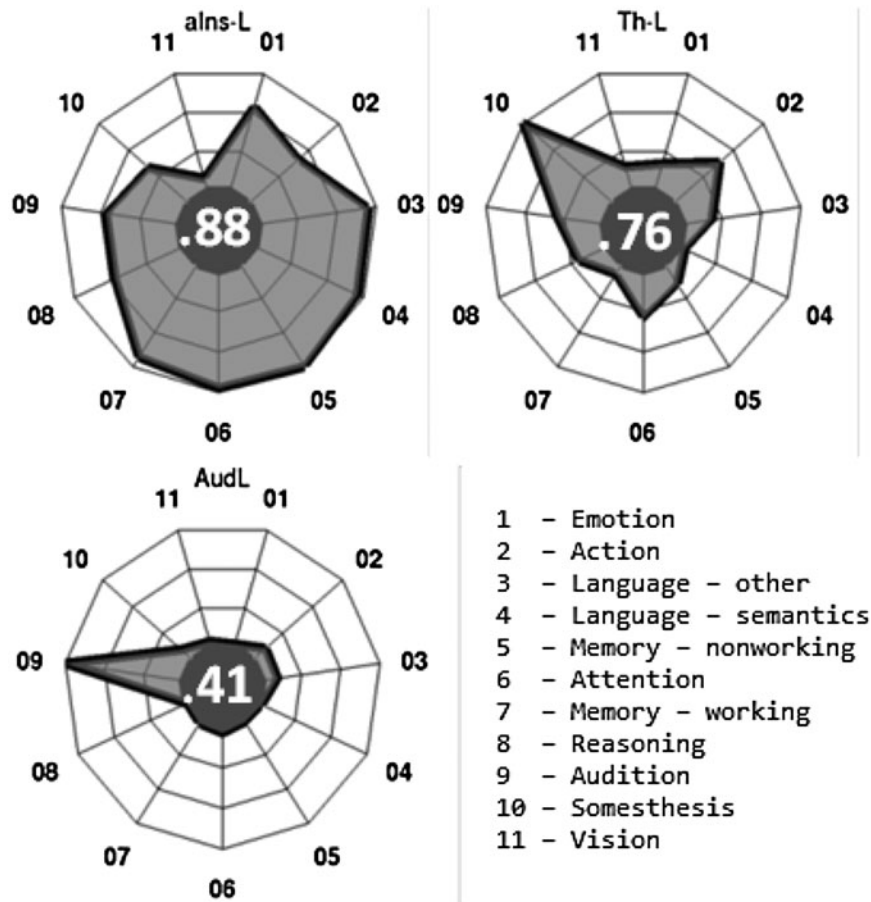
Figure 2. Functional fingerprints representing the relative amount of activity across 11 task categories for three voxels from left thalamus, left anterior insula, and left auditory cortex (counterclockwise from top right).

every category, albeit not uniformly. Functional fingerprints represent the likelihood that an active region is active during, or being activated by, a given type of task or stimulus, and thus offer a way to capture the different functional biases or underlying causal dispositions of individual regions.

This brings us to the question of whether we can observe regions of the brain cooperating with different partners under different circumstances. To illuminate this question, the technique of choice is a functional connectivity analysis. Using the same collection of neuroimaging experiments, we searched for deviations from statistical independence in the activity of individual regions. That is, we looked to see whether regions are more likely to be active during the same experimental task than would be predicted by chance. The results of such analyses can be represented as a graph, where the nodes of the graph represent regions of the brain, and edges between the nodes indicate that the connected regions are statistically likely to be co-active, and are therefore functionally connected. As was reported in a number of recent studies (Anderson 2008a; 2010; Anderson & Penner-Wilger 2013), it does indeed appear to be the case that regions of the brain – variously defined in the different analyses – have different functional partners during different functional circumstances. By way of illustration, Figure 3 depicts the functional connectivity graphs observed during emotion, attention, and semantics

tasks. The functional partners of left precentral gyrus are highlighted. As can be easily seen (and can be confirmed quantitatively), individual regions of the brain are active in multiple task circumstances, but have different functional partners in each.

These are just two pieces of suggestive evidence for neural reuse, both rooted in the neuroimaging literature, which is of course limited in various ways. *After Phrenology* also surveys electrophysiological studies of single neurons that highlight the importance and prevalence of mixed selectivity (Cisek 2007; Cisek & Kalaska 2005; 2010; Rigotti et al. 2013); cognitive interference and neural attenuation studies that demonstrate the activation of individual cells by multiple different tasks and stimuli (Glenberg & Kaschak 2002; Glenberg et al. 2008; Roux et al. 2003; Rusconi et al. 2005); work with sensory substitution devices that suggests that many regions of the brain are (and remain throughout life) capable of receiving and processing inputs from multiple sensory modalities (Merabet et al. 2008); and work demonstrating the importance and ubiquity of neuromodulation at multiple spatial scales (Bargmann 2012; Hermans et al. 2011). Overall, the evidence is far more consistent with neural reuse than with competing, modular accounts of brain organization.

The developmental framework advocated in *After Phrenology* is an extension of the Interactive Specialization framework (Johnson 2001; 2011). As with interactive
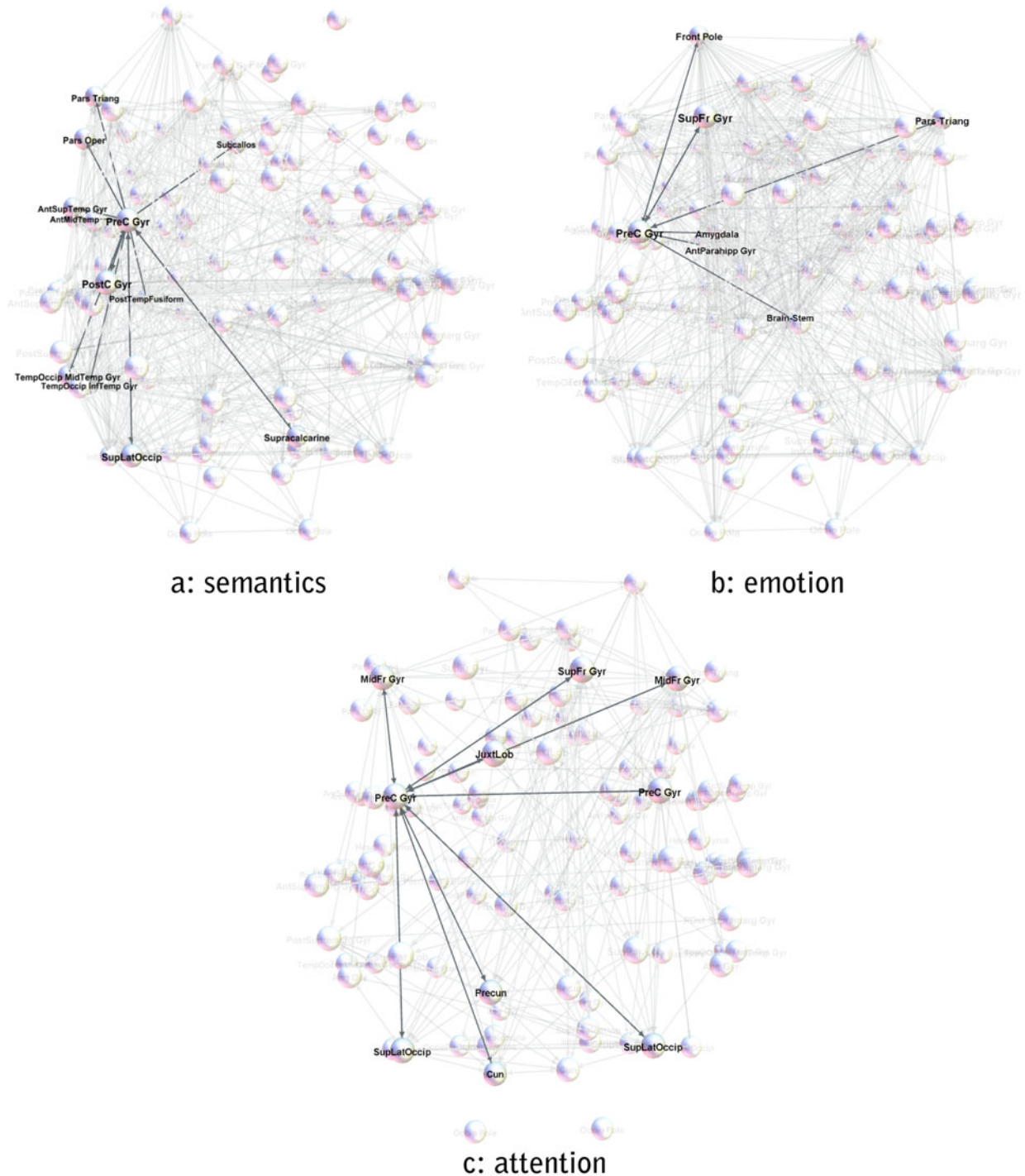
Figure 3. Functional connectivity graphs during semantics, emotion, and attention tasks. The functional partners of left precentral gyrus are highlighted. Nodes are placed in a projected three-dimensional space at the approximate center of each brain region from the Harvard-Oxford atlas; the figure shows the brain from above, front toward the top of the page.

specialization, and unlike the maturational viewpoint championed by Kanwisher (2010) and others (e.g., Atkinson 1984), neural reuse emphasizes the importance of experience in shaping the functional biases of local neural elements. It will only rarely, if ever, be the case that the functional properties of a region of the brain are shaped primarily by genetic factors. Similarly, neural reuse emphasizes that the functional properties of local regions both partly determine and are partly determined by the regions with which they interact. The multiple functional

coalitions that are set up during development and learning depend on the functional biases of their constituent regions, but these coalitions also help shape those functional biases as the behaviors the coalitions support are refined.

Neural reuse departs from interactive specialization by emphasizing the participation of neural elements in *multiple* coalitions. Consequently, it also departs from interactive specialization on the issue of whether and to what degree we should expect neural elements to be functionally *specialized*. That there is functional differentiation across

the brain is abundantly clear (and illustrated in [Fig. 2]). But there is apparently *not* functional specialization. Hence, the developmental framework advocated in *After Phrenology* is called *interactive differentiation and search*.

Although the evidence surveyed in *After Phrenology* does not appear to be consistent with the idea of functional specialization, it might nevertheless be the case that there exists some alternate taxonomy of function and level of description in terms of which brain regions could be assigned specific, dedicated functions (Price & Friston [2005]). In my own view, the apparent ubiquity of neuromodulation, and the prevalence of mixed selectivity in individual neurons, will make true functional specialization rare. But it is certainly an open question, one that is treated at length in *After Phrenology*.

3. Neural reuse, evolution, and modularity

As I hope is clear even from the brief discussion above, neural reuse is not consistent with the notion that the brain is composed largely of segregated, functionally dedicated, specialized neural modules. Different networks share parts, and the parts may do different things for each of the networks in which they participate, as a result of the constraints imposed by the network interactions (Anderson [2015]). The brain is functionally differentiated but also deeply integrated in ways that make modularity very unlikely. Yet, the modularity assumption remains pervasive, despite the mounting evidence for reuse in the cognitive neurosciences, and the scant evidence for mosaic evolution in evolutionary biology (Aboitiz [1996]; Finlay & Darlington [1995]; Finlay et al. [2001]; Stephan et al. [1988]; Yopak et al. [2010]). What accounts for this tenacity? In short: modularity appears to offer an answer to the paired questions of how behavior is heritable and how brains are evolvable. To break the hold of modularity, then, requires offering better answers to these questions.

Although a critique of evolutionary psychology is not central to *After Phrenology*, a few words about that approach to understanding the psychological and neural legacy of our evolutionary history will highlight some of the reasons modularity can seem attractive, and throw into relief the alternative account I am offering. Evolutionary psychology (Buss [2005]; Confer et al. [2010]) rests on two problematic assumptions. It assumes, first, that the environment of selection is different from the current environment and can be adequately described and, second, that the solutions to the adaptive challenges posed by that environment are individually genetically encoded. The first assumption is problematic not just because of the inherent uncertainty in identifying and accurately describing ancient environments, but also because of what might be called the evidentiary dilemma for evolutionary psychology. Insofar as the environment of selection is very different from our own environment, evidence for the persistence of psychological mechanisms optimized for that environment is always simultaneously evidence *for* (an) adaptation, but *against* adaptivity (because the mechanism is tuned to the "wrong" environment). Likewise, insofar as the environment is relevantly similar to our own, then the identification of psychological mechanisms appropriate to that environment is simply evidence for adaptivity, and *not* for an adaptation.

Hence, as it is currently conceived, evolutionary psychology is hard-pressed to do justice to both adaptation and

adaptivity, and a fully adequate evolutionary science of the mind must of course do both. This issue is related to the second assumption driving evolutionary psychology: because it assumes that solutions to environmental challenges must be encoded genetically (and result in dedicated neural modules), it is forced to conclude that the timescale of change will be quite long. I see little evidence for this latter assumption, but there is one important consideration that, at least on its face, seems to favor it. The idea is this: If psychological processes and the neural structures that support them are to be viewed as heritable adaptations, then they must be separately modifiable, for otherwise there is no available target for selection pressures. If this is correct, a nearly decomposable, modular brain consisting of separately modifiable subsystems appears to be *required* by evolution.

The mistake that evolutionary psychology makes here is subtle and twofold. The first mistake is to forget that not just genes but also *environments* are generally inherited, and the second is to suppose that a cognitive process is separately modifiable if, and only if, its supporting components are separately modifiable. In fact, the key to understanding how organisms inherit species-typical behaviors is seeing how genetics, environment, and developmental processes all work together (Anderson & Finlay [2014]). According to the interactive differentiation and search framework developed in *After Phrenology*, learning is a matter of finding and consolidating the right neural partnerships to support the acquisition of the target behaviors, where the "right" partners are those with the particular functional biases that together serve the behavioral ends. The functional biases are in turn shaped by learning and experience, all of the way back to and including very early experience. It is here that genetic and environmental factors have their most important initial impact. If we assume highly stereotyped projections from sensory afferents to specific regions of the developing brain, and an environment largely conserved between generations, then early experience will be sufficiently similar between individuals to induce neural structures with conserved, species-typical functional biases.

Given a similar stock of functional elements, and a species-typical developmental trajectory for skill acquisition, the processes of neural reuse – of the discovery and consolidation of functional coalitions – will tend to produce similar networks, and hence similar, species-typical solutions to the challenges posed by the largely conserved environment. On this model, selection pressures would tend to target not specific cognitive processes, but rather developmental mechanisms for ensuring the robust availability of neural elements with a wide range of functional biases. Note that this model also accounts for the possibility of psychological adaptations, *and* for the persistence of rapid adaptability to changing environments. One can inherit a psychological adaptation in virtue of inheriting the environmental challenge along with the neural elements that can be put together to meet it; and one can adjust a cognitive process to a new situation by changing the mix of elements in the neural coalition that implements it.

4. Networks of the brain

The brain is a network. So far, this is to say very little, for who would deny it? What is different about the neural reuse framework is not that it insists the brain is a network, but rather that it supposes the brain is a network

with some very important architectural and functional properties. These include multiscale dynamics, multidirectional feedback, noncomponentiality, and action-orientation. I will treat each of these properties in turn.

The brain is a dynamic network that remodels itself at multiple spatial and temporal scales. In addition to the two types of neuroplasticity detailed earlier (sect. 2) that co-operate to remodel the synaptic (or "wired") network, there are modulatory processes that change the effective connectivity of the synaptic network. Mechanisms include genetic expression that serves to activate and inactivate individual synapses, thereby changing the functional properties of local networks (Bargmann 2012); dendritic spine motility (Holtmaat & Svoboda 2009) that can make synaptic connections more or less reliable; extra-synaptic diffusion neurotransmission involving the release from non-synaptic sites of neurotransmitters that diffuse through the extracellular matrix and change the firing likelihood of the neurons to which the transmitters bind (Agnati et al. 2010); and various hormonal mechanisms and systems that modulate brain activity at long temporal and broad spatial scales (Bauer et al. 2001; Pfaff 2002). Hence, function in the brain depends upon, at least: a neural network, an underlying genetic network, and an overlaid chemical gradient. Each of these elements is only partially understood, and their dynamic interactions even less so.

At any given moment in a quiescent network, the current effective connectivity would dictate the evolution of any induced pattern of activity. But the brain is of course never quiescent. It is always active to some degree, whether as a result of the purposeful activity of the agent or the endogenous activity of the brain at "rest" (Raichle et al. 2001). The effect of externally induced (e.g., perceptual) neural activation will depend not just on the effective connectivity of the network, but also on the *ongoing* activity resulting from past patterns. The brain is decidedly *not* a primarily feed-forward system. Instead, interactions between feed-forward, feed-back, bottom-up, and top-down processes both determine how the activation patterns evolve and also induce further changes in the effective connectivity of the network (Cole et al. 2013). Moreover, in driving the ongoing behavior of the organism, these evolving patterns influence the nature of the externally induced activations; organisms are perception seeking, not passive recipients of environmental stimulation.

In a brain marked by such multidirectional feedback, understanding the interactions between parts becomes a significant challenge. Indeed, even defining the functional parts becomes difficult, as the relevant functional parts will themselves apparently change over time. For these reasons, we must move beyond componential computational models of the brain. Different neural patterns indexing different perceptual states, action choices, preferences, reward estimations, other predictions, and so forth, do not combine syntactically in the manner of compositional linguistic structures. Neither are the functional parts of the brain always best understood as components with stable, intrinsic input-output mappings and well-defined interfaces supporting the exchange of content-carrying symbols. Instead, patterns superpose in the brain and interact through the process of biased pattern competition (Cole et al. 2013; Desimone & Duncan 1995; Miller & Cohen 2001; Platt 2002). Ongoing perception and evolving reward estimates reinforce some patterns and disrupt others, changing the trajectory of the evolving neural state and thereby the behavioral (and perceptual) trajectory of the organism. Similarly, local function emerges from the complex, dynamic interactions between large- and small-scale structures in the brain. Sometimes the function of larger structures can be understood by understanding the intrinsic functions of its parts and the nature of their interaction (Craver 2007); but as is illustrated by the case of direction-selectivity in the dendrites of Starburst Amacrine Cells (SACs), other times the functions of the low-level parts appear to be determined by the constraints imposed by the larger structures with which they interact (Anderson 2015).

In the brain sciences, we need to develop models of explanation that allow for the possibility of top-down and bottom-up mutual constraint, in which both local and global function are synchronically co-determined by the dynamic coupling between elements at various spatial levels of organization. In *After Phrenology*, I therefore introduce the idea of Transiently Assembled Local Neural Subsystems (TALoNS). TALoNS are the temporary, reproducibly assembled functional parts (large- and small-scale networks and other elements) of the brain. TALoNS have intrinsic causal properties or dispositions determined by their internal structure and effective connectivity, but their functional selectivity (e.g., direction selectivity in SAC dendrites) emerges from the way these dispositions are constrained by the other functional structures with which they interact.

All of the above serve to underscore the following: The brain is a highly dynamic, adaptive system, in which structure and function are constantly adjusting to the changing circumstances of the organism. This is as it should be. The brain evolved to control action. It is a crucial mediator and modulator of the sensory-motor coupling that governs an organism's fit to its environment. Given this job, it *had* to be adaptive at multiple temporal scales, and capable of naturally managing the multiple simultaneous demands that are imposed by the complex interactions between an organism's needs and its perceived opportunities for action. The brain is an action-oriented, and not a perception-oriented, system. It is crucial to understand the implications of this fact for the nature of the brain and for the science that purports to study it. These matters are taken up in the next section.

5. Embodiment and cognitive processing

Traditional cognitive science is captured by a particular picture of our fundamental epistemic situation. According to that picture, sense organs are conduits for inputs called "sensations," on the basis of which the individual organism generates a representation of the causes of that input – internally reconstructing the objects and properties in the external world. Cognition, in this picture, consists of the targeted internal manipulations of this reconstruction in service of the agent's goals – ultimately, deciding what to do next. Perception is induction, and cognition is calculation.

Acceptance of this framework accounts for the fact that one of the fundamental jobs of cognitive neuroscience has been to discover what is represented where in the brain, and how each representation is transformed into or impacts the others. Acceptance of the framework accounts for the abiding interest in specifying the innate "knowledge" or stored assumptions that guide perceptual reconstruction, whether that involves solving the (otherwise

apparently intractable) problem of inverse optics (Edelman 2008; Marr 1982) or inducing the grammar of natural language (Chomsky 1957), for it is readily apparent that "sensations" are impoverished and unreliable – and need to be supplemented. Acceptance of this framework even accounts indirectly for the componential assumption that is built into most theories of the functional structure of the brain, for insofar as the challenges of perceptual reconstruction and cognitive calculation require specialized knowledge, it is natural to imagine specialized neural systems for solving those problems. Moreover, insofar as cognition is a matter of *representation transformation*, it must also involve *information communication* among these systems, which requires conduits and interfaces, and naturally leads to a modular architecture of stable, specialized, relatively isolated, nearly decomposable, message-passing components (van Gelder 1995). The framework, this is to say, is deeply embedded in the cognitive sciences. But it is time to abandon it.

Perhaps the most fundamental problem is with the very concept of a sensation. To make a point that is at least as old as James (1890): "Sensation" is a theoretical construct, an abstraction away from actual experience. As with some fundamental particles of physics, sensations do theoretical work, but no sensation has yet been observed. If perception is reconstructive, then it needs building blocks, and sensations are the hypothesized blocks. If perception is reconstructive, it needs a starting point, and sensations are the hypothesized points. If perception is reconstructive, there is a definite order of events: sense, think, act. But perception is not reconstructive; representing the environment is not what our brains evolved to do. Our brains evolved to control action. Experience is not composed of atomic units, nor does it have a definite starting point; it is a continuous stream. Action does not come after thinking, which comes after perceiving; thinking, perceiving, and acting are synchronous and co-determining.

The alternative, action-oriented framework developed in *After Phrenology* consists of the following tenets: perception is active; perception is relational; the brain is a control system. Thinking – cognition – involves harnessing the mechanisms of sensory-motor coordination and environmental interaction to more abstract ends, but the character of the underlying mechanism remains what it has always been.

Perceiving is always acting because to know the world is to move about in it. Consider the case of olfaction, which is largely useless without the ability to move. All of the useful information about chemicals lies in the *distribution* in the environment, and picking up this information requires moving around. Put differently, chemical detection is not chemical perception unless and until it is chemotaxis. Touch, too, is quintessentially active: We feel the support offered by a surface, or the hardness of a material by *pressing*, the heft of a thing by *lifting*, and texture and shape by brushing and grasping. Naturally, one can be touched, just as one can subject to a chemical impingement, and such events may well convey information without movement, but these are degenerate cases for perceptual systems that normally function via movement. The same is true of vision: The passive reception of reflected light is the degenerate case for what is an active perceptual system. The data of visual perception are not the momentary impacts of reflected light in the retina, but rather the changes in the retinal projection as our posture and position changes.

The problem of visual perception is *not* one of constructing a three-dimensional model of the world from passive two-dimensional stills; it is rather a matter of picking up on the world-specifying information available in the actively gathered stream of experience (Gibson 1966; 1979). The processes whereby we do this are of course still poorly understood; the point is that perception poses a *different problem* from what has been traditionally supposed. And it is this latter problem that our brains evolved to solve.

Because perception is both active and in the service of action, much of the information to which organisms are attuned is not objective information of the sort one might need for model-building, but rather *relational* information that is more immediately useful for guiding action in the world. It is the overall job of the organism's brain and nervous system to manage various organism–environment relationships. Perceptual systems keep the organism in touch with the values of these relationships: the closeness of the obstacle, the support of the surface, the passability of the gap. When we think otherwise, we can make scientific errors of an interesting sort, underestimating the accuracy of our perceptual systems. Consider the matter of weight perception. Humans are notoriously poor weight estimators and are liable to such errors as the size–weight illusion: given two objects of the same weight but different sizes, the smaller object will be judged heavier (Murray et al. 1999). On the traditional view, this fallibility is unsurprising. After all, the torque imposed on our arm as we hold an object in the hand will depend on the length of one's arm, the angle of the shoulder and the elbow, and other variables, and will change as we move about. Extracting any stable, objective property of the object would naturally be very difficult in light of such variation. But this is not how perception works. In fact, the information *is in the variation*, and the relational property that the information specifies in this case appears to be the throwability of the object. Humans turn out to be *very* accurate estimators of throwability (Zhu & Bingham 2011). The position defended in *After Phrenology* is that most of perception should be understood on this relational model.

The last tenet that makes up the embodied framework outlined here is that the brain evolved to be the control system for an active, environmentally situated organism. The fundamental cognitive problem facing the organism – deciding what to do next – is best understood not as choosing the right response to a given stimulus, but rather as choosing the right stimulus – the right experience to seek – in light of a goal. Knowledge of sensorimotor contingencies (Noë 2004) – of how perceptions change with action – and the perception of affordances (relationships between an organism's abilities and objects in the world that indicate opportunities for action) work together to allow an organism to follow chains or sequences of experiences to achieve its ends, whether that be a feeling of satiety, the experience of safety, or the perception of a finished nest. As Paul Cisek (1999) has pointed out, all living things have homeostatic mechanisms that keep biologically relevant variables such as temperature, pH, or chemical concentrations within some acceptable range. Some of these mechanisms are metabolic or physiological, but others are behavioral: moving, eating, manipulating, and so on. The fundamental function of behavior, then, is to maintain organism-relevant variables within some desired range, and the fundamental function of the brain is to manage such behavior. The

brain is a dynamic control system that modulates the sensorimotor coupling at multiple spatial and temporal scales.

In *After Phrenology*, I follow Cisek (2007; Cisek & Kalaska 2010) in arguing that the biased pattern competition observed in the brain should be understood *psychologically* as biased affordance competition. What an organism's brain is fundamentally doing is managing the relationship between the organism and the environment, and its perceptual apparatus is specially suited for facilitating that task. An organism perceives the values of salient organism–environment relationships and, in light of some goal(s), acts so as to perceive the right changes in those relationships. The brain that manages this behavior is organized in such a way that its various parts have different dispositions to manage the values of the perceived relationships. Interaction with an environment offering multiple affordances causes regions of the brain to be differentially activated in accordance with their functional biases. A situation posing several possible courses of action will cause multiple distributed patterns of neural activation across the brain, and the behavior of the organism in this situation will be ultimately determined by competition among the patterns. I argue that this competition should be understood to reflect tension among the various behavioral control loops that could be enacted; loosely speaking: Pattern competition in the brain is affordance competition in the mind. The summed cooperation and competition among the active dispositions in the brain both determines the course of action and structures the control loop that facilitates the required behavior.

6. Function–structure mapping in an interactive brain

Over the past several sections, I have been advocating for a picture of the functional structure of the brain that illuminates its evolutionary and developmental origins, and does justice to the significant functional complexity of its individual working parts. I also advocated for functional fingerprinting as an appropriate tool for capturing and quantifying functional complexity. In fact, functional fingerprints and the style of thinking they promote may help point the sciences of the mind in a new and fruitful direction.

To see how and why, we need to appreciate the epistemic situation we are in. A scientific experiment is a deliberate intervention into the causal structure of the world. We intentionally manipulate – vary the value of – some physical condition and record the value of another. The signal that this intervention produces is generally mixed – that is, dependent on numerous causal factors that we would ideally like to disentangle. A simple example is the measurement of weight (or force more generally), which physics teaches us is in fact the product of two more fundamental properties, mass and acceleration. This realization gave us a better purchase on the underlying causal structure of physical reality. Similarly, the varying price of a stock over time is a mixed signal driven by multiple economic factors including the money supply, corporate profits, and perceived innovation, whereas the price of 100 stocks is a set of mixed signals all being driven by the same causal factors but to different degrees. The price of a tech stock might be relatively less sensitive to earnings and more to innovation than the price of an energy stock, for example.

Given this situation, we need to ask: When we measure the activity of 1 or 100 or 1,000 different parts of the brain, what is the underlying nature of this set of mixed signals? What are the psychological factors that contribute to the changing values we record from brain and behavior? In *After Phrenology*, I argue that the central guiding scientific quest for the cognitive neurosciences should *not* be determining what the basic cognitive operations implemented in individual regions of the brain are. The functional complexity of the brain suggests that this approach will offer at best an incomplete and at worst deeply misleading account of brain function. There should nevertheless be detectable regularities in the patterns we record from brain and behavior; there should be some underlying *structure* in the signal. Hence, I advocate for a science that asks: What are the psychological factors that best capture and account for the differential activity of the brain in various circumstances?

One reason functional fingerprinting can be so powerful is that it offers an avenue toward an answer. In the same way that analysis of people's responses to a variety of interventions can reveal a common set of factors defining individual personalities, so too the analysis of multidimensional functional fingerprints of brain regions and networks may reveal a set of primitive psychological factors (Barrett & Satpute 2013; Gold et al. 2011; Lindquist & Barrett 2012; Lindquist et al. 2012; Poldrack 2010; Poldrack et al. 2009). I call these neuroscientifically relevant psychological (NRP) factors. According to this approach, psychological states such as anger and fear, as well as processes such as attention and cognitive control, involve different mixtures of many of the same domain-general ingredients. These factors would map to the brain such that more than one part of the brain would support each factor, and more than one factor would load on each part. That is, brain regions and networks will differ not necessarily in terms of their component operations, but rather according to their loadings on a set of primitive NRP factors.

This scientific approach appears to better respect three organizational features of the brain emphasized here: (1) the functional diversity of individual regions of the brain, (2) the functional differentiation of individual regions of the brain, and (3) the frequent functional overlap between the constituents of different networks. It will also help us think our way beyond the functional model of linearly interacting components that we inherited from seventeenth-century mechanism and nineteenth-century engineering practices. In the brain, function emerges from structure in ways more complex than that model can capture. But we are developing tools adequate to the task.

What exactly *are* NRP factors? What is their best psychological construal? That is, of course, an open question, one that will be answered as part of *doing* the science described in *After Phrenology*, not in advance of it. In my view, NRP factors index basic dispositions to help manage the value of some organism-relevant environmental variable or relationship (see sect. 5). Because neural reuse has both an anatomical and a behavioral aspect, we should expect to see these dispositions manifest in multiple circumstances. Hence, there might be basic factors for managing closeness and warmth, and these might manifest in both physical and interpersonal contexts (Bargh & Shalev 2012; Xiao & Van Bavel 2012). Therefore, we would also expect the regions of the brain that load on the relevant factor to be active across these different contexts.

It is of course an implication of the approach that the fundamental NRP factors that we are seeking have generally *not* been already identified, and will cross-cut the current taxonomy of psychology. In *After Phrenology*, I marshal the evidence for this claim; here, I will simply note the following: given that cognitive neuroscience (and, indeed, psychology more generally) has yet to be deeply influenced by evolutionary biology, and that it adopted wholesale the psychological taxonomy of cognitive psychology that, as I have argued above, is organized around a faulty framework (and was initially devised to be a science *autonomous* from the neurosciences besides), then it would be something of a miracle if the right set of concepts had already been formulated. I believe that following the path laid out in *After Phrenology* will lead to a new and better vocabulary for understanding mind, brain, and behavior. Moreover, I argue that this vocabulary will better reflect the evolutionary history of human beings, and the action-orientation of cognition, if it is organized not around the concepts of sensation and representation, but rather around the notion of an affordance.

7. Reuse, interaction, and "higher-order" cognition

As I noted in the introduction, thinking and acting are mutually supporting and intertwined at multiple timescales. We think with and through our interactions with objects and one another. We routinely act to help us see and think: we spin puzzle pieces to make their fit easier to perceive, rearrange playing cards and Scrabble tiles to make patterns easier to detect, and label our environments with signs to aid memory and ease navigation (Clark 1997). And just as we create physical tools such as hammers, knives, and levers to augment our physical capacities, so too we invent cognitive artifacts to augment our mental ones. Among the most important of these are the cultural practices of speaking, writing, and calculating, and the symbol systems that support them. And what is deeply fascinating, and helps illuminate the true nature of human intelligence, is that we treat these cognitive artifacts just like physical ones, reusing our finely honed abilities for interacting with objects in the service of improving our thinking.

Consider mathematical symbols: People point at them, gesture over them, move them, and strike them out. These actions serve myriad purposes: They direct spatial attention, they aid memory, they keep one's place in the problem-solving procedure, and they make a solution easier to reach. These actions are not peripheral to knowing and doing mathematics, but part and parcel of it. Mathematical symbols have the character that they do *so that* perception–action loops can be brought to bear on – be harnessed to – the practice of calculating. Equations have affordances that invite us to act on and with them to achieve the task they were designed for. To learn algebra is to acquire a sensorimotor skill, and acting in accord with the rules of algebra is a matter of learning to see and act in accord with the transformations that the equations afford (Landy & Goldstone 2009).

Does this mean that doing math is mindless and noncognitive, that it does not involve *thinking*? Of course not! I hope it is clear by now that the rigid distinctions between sensing and thinking and doing are among the many bad ideas that need to be jettisoned in our reformed science of the mind. Seeing and touching and interacting with and manipulating things are partly *constitutive* of thinking. We have achieved our cognitive capacities in part because we have found ways to reuse our physical capacities to augment our mental ones; in a process supported by *neural* reuse, we repurpose our behavioral routines in multiple circumstances for myriad cognitive ends.

To drive this point home, and to preempt the argument that the embodied, embedded, evolutionary developmental account of cognition that is developed in *After Phrenology* can never account for our capacity for natural language, I outline a theory of language according to which language is an *interactive social practice*. It is both a form of joint action (Clark 1996; Sebanz et al. 2006) and a coordinating structure for facilitating cognitive and social interactions (Tomasello 1999). Language works by presenting and manipulating cultural affordances that will cause one's dialog partner(s) to see and do what the speaker intends to be seen and done. Language works *because* it has developed to take advantage of and is fitted to our interactive sociality (and *not* because we evolved specialized, dedicated, modular neural machinery to support it). Like all successful artifacts – physical and cognitive both – it has the right two-way fit: It suits both our abilities and its purpose. And like all successful cognitive artifacts, it enhances our capacities in various ways: It aids memory, improves self-control, biases attention, and more. There is, of course, much more to the argument, and much more to the story, and for that I hope you will turn to *After Phrenology*.

8. Psychology after phrenology

As I hope is clear in this précis, and as I hope is compelling in the book it introduces, I am calling for the development of a new functionalism as the basis of a unified science of mind that respects its biological bases, its evolutionary history, and its environmental and cultural embeddedness. Among its tenets are the claims that the functional architecture of the brain has been established by natural selection through a process marked by both differentiation and continuity, that our complex and diverse behavioral repertoire is supported primarily by the ability to dynamically establish multiple different functional coalitions coordinating both neural partnerships and extra-neural resources, and that the brain is fundamentally action-oriented, with its primary purpose to coordinate the organism's ongoing interactions with the world and adjustments to external circumstances. What might psychology and neuroscience look like if the framework I advocate in *After Phrenology* were widely adopted? In an appendix to the book, I lay out the theoretical challenges and a specific research agenda. Here, I will end with a broad-strokes characterization of the science to come.

1. We will represent the functional activity of the brain in a multidimensional manner that captures the underlying functional and dispositional properties, and we will give up the notion that the neural responses we observe and measure must reflect the engagement of a single unified function.

2. We will expect not just local, but also distributed contributions to overall function, determined by the interactions between top-down and bottom-up, feed-forward and feedback processes. Structurally, we will attend to the interactions between regions – how these change and how they

map onto changes in behavior. We will develop better non-componential models of functional integration that can capture the myriad ways that function emerges from interacting structure. Developmentally, we will work to establish the mechanisms whereby potential functional partnerships in the brain are discovered, tested, and maintained. Evolutionarily, we will seek to capture the adaptivity of the organism in all its forms and to understand that natural selection targets not just structures but also processes.

3. We will deeply rethink the vocabulary of cognition, ideally giving the brain a voice in the process. In discerning what the brain cares about, we will remember that it evolved to be an action-control system, specializing in managing the values of salient organism-environment relationships. Hence, many of the properties to which the brain is attuned will be action-relevant and relational; throwability and climbability will likely be more important to the brain than weight and slope.

4. We will recognize that cognition does not take place in the brain alone. We think with and through artifacts and one another. Although it will always be tempting (and occasionally necessary) to bracket off the natural and social worlds to focus on the brain in isolation, we will work to develop experimental paradigms that include robust social and environmental interactions, and we will develop techniques for measuring the details of the interactions among brain, body, and world.

5. We will embrace the empirical tools offered us by machine learning, graph theory, independent component analysis, multidimensional scaling, linear algebra, dynamic systems theory, and so forth, that promise to help us do justice to the dynamic complexity of the brain. We will realize that the focus on local, linear correlations between brain activity and simple stimuli will never be by itself sufficient to capture the complexity of the brain and its interacting parts. We will turn to empirical tools better suited to measuring distributed information and able to disentangle the psychological mixtures that brain activity reflects.

I believe that this is the most exciting time in the history of the neurosciences. We have at our disposal phenomenal technological tools allowing us to measure and analyze function in ways unimaginable even just a few short years ago. If we can manage to match the quality of our conceptual and experimental tools to the quality of our technology, the scientific future is very bright. I hope *After Phrenology* can help illuminate the path.

Open Peer Commentary

After phrenology: Time for a paradigm shift in cognitive science

Paul Benjamin Badcock,[a,b,c] Annemie Ploeger,[d] and Nicholas Brian Allen[e]

[a]*Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia, 3010;* [b]*Centre for Youth Mental Health, The University of Melbourne, Melbourne, Australia, 3052;* [c]*Orygen, The National Centre of Excellence in Youth Mental Health, Melbourne, Australia, 3052;* [d]*Department of Psychology, University of Amsterdam, 1018 XA Amsterdam, The Netherlands;* [e]*Department of Psychology, University of Oregon, Eugene, OR 97403.*

pbadcock@unimelb.edu.au   a.ploeger@uva.nl
nallen3@uoregon.edu
http://www.uva.nl/profiel/a.ploeger
http://psychology.uoregon.edu/profile/nallen3/

Abstract: Anderson (2014) uses an impressive, consolidating review of the literature to argue for major changes in cognitive science. Arguably, however, much of what he proposes is not particularly new. He also neglects important predictive coding approaches that call his perspective of the brain into question, and his misconstrual of evolutionary psychology devalues an influential paradigm that promises to complement his own.

Anderson's (2014) *After Phrenology* makes an authoritative and timely contribution to the literature by describing two key developmental mechanisms (i.e., neural reuse and, more speculatively, interactive differentiation and search) that have important implications for our understanding of the brain. In particular, we applaud his provision of a rigorous, empirically informed treatment of the functional diversity and overlap of different brain regions that also explains how higher-order, domain-general cognitive capacities emerge from evolved neural circuitry. Arguably, these contributions call massive modularity into serious question. Like many others, we agree with Anderson that in light of the evidence, this view can no longer be reasonably sustained. As such, we think that one of the chief virtues of Anderson's work is that it brings us one step closer to the abandonment of massive modularity as a model of the brain. Naturally, any appeal for greater collaboration across the cognitive sciences should also be lauded (particularly one accompanied by such clearly specified and promising research directions), and Anderson's broader treatment of the functional properties of the brain is comprehensive, scholarly, and elucidating.

That being said, Anderson also seems to tell us little new. Buller (2005), for example, forwarded neural plasticity as an important developmental mechanism that accounts for domain-general capacities in his critique of massive modularity a decade ago, and as Anderson readily admits, he is not the first to focus on neural reuse to explain how evolved neural mechanisms come to perform new cognitive functions. Similarly, cognitive scientists have been emphasizing the complexity of the brain and its development, the functional integration of neural regions, the need to use sophisticated dynamical methods to analyze such complexities, and the pitfalls of modularity for many years now (e.g., Fotopoulou 2014; Karmiloff-Smith 1992; Kelso 1995; McIntosh 2000; Mesulam 1990; Meunier et al. 2010; Pfeifer & Allen 2012; Price & Friston 2002). As a consolidating synopsis of disparate research programs, *After Phrenology* is undoubtedly impressive, but it arguably falls short of constituting the inception of a new paradigm – Anderson suggests that we use his treatise as an impetus to "get to work," but cognitive scientists have already been undertaking such work for quite some time.

More critically, although we certainly agree that the brain can be seen as an action control system, we are nonetheless skeptical of Anderson's claim that perception is not reconstructive. In particular, highly influential predictive coding approaches in neuroscience call his action-oriented framework into question by placing perception center stage (or at least on an equal footing with action). Take, for example, Friston's (2005; 2010) free-energy principle, which asserts that the brain instantiates a biological imperative to model the world. According to this perspective, perception and action operate synergistically to minimize prediction errors and optimize an individual's internal representations of the environment. A key corollary of this hypothesis is active inference: the idea that all behavior can be understood in terms of the selective sampling of sensory data so that we experience what we expect to (in order to avoid surprises). That is, we act upon the world to ensure that our predictions are self-fulfilling (Friston et al. 2009; 2010). Of course, one way to interpret this view is to

suppose that action subserves perception, but at the very least, it cautions against definitively attributing causal primacy to it.

Notably, there are times when Anderson also appeals to more of a synergistic relationship between action and perception, and he discusses ideas that are highly reminiscent of predictive coding such as perceptual hypothesis generation (e.g., Gregory 1980). As such, we imagine that Anderson has a clear idea about how predictive coding accounts align with his own view, or – considering his professed aversion both to global brain theories and the notion of reconstructive perception – he might be inclined to reject them outright. In any case, given the widely recognized explanatory power of these models (e.g., Clark 2013b; Hohwy 2013) – and their capacity to accommodate the sort of phenomena that Anderson calls upon to substantiate his own view (e.g., the size–weight illusion; see Clark 2013a; Lupyan 2015) – his failure to explicitly address them renders his conclusion that we should privilege hypotheses of action control somewhat questionable and premature. For our part, we wonder why Anderson feels the need to advocate such a polarizing view – the way we see it, given insufficient evidence to the contrary, one should assume more of a dialectical, reciprocally causal relationship between action and perception.

This brings us to our greatest reservation about Anderson's work – its cursory and misleading construal of evolutionary psychology. Of particular note, his assertion that species-typical cognitive patterns emerge from the repeated assembly of reliably recurrent developmental resources was proposed by an evolutionary psychologist at the turn of the century (Caporael 1997; 2001) and has been echoed by many others since (e.g., Badcock 2012; Frankenhuis et al. 2013; Geary & Bjorklund 2000; Kenrick 2001; Lickliter & Honeycutt 2003). Given his critique of massive modularity – which he appears to erroneously equate with the broader paradigm of evolutionary psychology itself – we think it would only be fair to acknowledge the theoretical heterogeneity of the field, not to mention its provision of a similar view of the evolved brain (and the development of its functional properties) long before him.

Anderson's underestimation of evolutionary psychology is particularly apparent in his treatment of massive modularity. Like most critics of this view, he fails to distinguish between massive modularity as an explanatory account of the properties of the brain and its methodological value as a research heuristic. Regardless of the veracity of massive modularity, evolutionary computational theories continue to guide research in a systematic and highly productive way, providing a wealth of insights into the adaptive significance of mental processes and behavior (e.g., Buss 2012; Crawford & Krebs 2008; Dunbar & Barrett 2007; Ellis & Bjorklund 2005; Pinker 1997). Indeed, the convenience of the phenotypic gambit is that it can produce substantive, testable hypotheses of behavior without demanding recourse to a mechanistic explanation of how it occurs. In other words, evolutionary psychologists chiefly concentrate on ultimate or functional analyses of cognitive and behavioral patterns, not their neurobiological instantiation (Klasios 2014). With this in mind, the pitfalls of massive modularity by no means vitiate evolutionary psychology.

Consider the illustrative case of attentional biases toward threatening stimuli. For example, research on snake detection shows that both humans (Masataka et al. 2010) and other primates (Shibasaki & Kawai 2009) display faster reaction times when detecting a snake in an array of neutral pictures, relative to a neutral picture in an array of snakes. Such findings resonate with a wealth of studies suggesting that humans exhibit a perceptual bias toward evolutionarily recurrent threats; biases that occur early in infancy – well before the acquisition of threat-relevant fears – and are likely to play a causal role in privileging fear learning for certain stimuli (LoBue & Rakison 2013). Note, too, that this need not denote a dedicated "threat-detection" or "snake-detection" module. Although some researchers have certainly suggested as much, others have left the mechanisms responsible for this bias open to question – this has not, however, prevented them from emphasizing its adaptive significance (see LeBou &

Rakison 2013). Indeed, to an evolutionary psychologist, such phenomena make clear sense. If, however, we were to follow Anderson's suggestion by abandoning this approach, we would be hard pressed to conjure an equally parsimonious explanation for them.

By contrast, Anderson's own suggestions for research seem overly neurocentric – advocating a rather nonsubstantive, descriptive approach to cognitive science. He tells us what cognitive scientists should look for – and provides excellent advice on how to go about it – but says very little about what, exactly, we should expect to find. The idea that distinct patterns of functional neural activity load on to "neuroscientifically relevant psychological (NRP) factors" (sect. 6, para. 4) is self-evidently true, but it also provides few insights into what these factors actually are. We therefore question the capacity of his approach to provide substantive hypotheses about overt behavior – precisely what one would expect from an action-oriented psychology. For this reason, although Anderson's work will no doubt prove useful to cognitive neuroscientists, we strongly suspect that most psychological scientists will find evolutionary psychology far more fruitful than his own scheme.

Ultimately, although Anderson's contribution is clearly important, some of his central claims warrant skepticism, and it is far from revolutionary. In particular, he recapitulates a perspective of the evolved brain that originated from the very paradigm he tries to refute, suggesting that evolutionary psychology and his own view might be a lot more compatible than he thinks. By our estimation, once we set a massively modular view of functional neuroanatomy aside, evolutionary psychology and Anderson's approach have the potential to be highly complementary. The former addresses the ultimate "why" questions of psychological science by elucidating the adaptive significance of cognition and behavior; the latter, mechanistic account can help us explain "how" these adaptive patterns are neurobiologically realized. As evolutionary psychologists have long been telling us (e.g., Badcock 2012; Barrett 2008; Dewsbury 2009; Frankenhuis et al. 2013; Kenrick et al. 2002; Ploeger et al. 2008; Scott-Phillips et al. 2011), these distinct levels of explanation can be mutually informative and reinforcing. In short, although we entirely agree that cognitive neuroscientists should hunt for NRP factors, we think that evolutionary psychology is one of the most promising places for them to look.

Why a developmental perspective is critical for understanding human cognition

Dean D'Souza and Annette Karmiloff-Smith
Centre for Brain and Cognitive Development, Birkbeck, University of London, London WC1E 7JL, United Kingdom.
dean.dsouza@ymail.com   a.karmiloff-smith@bbk.ac.uk
http://www.cbcd.bbk.ac.uk/people/affiliated/Dean
http://www.bbk.ac.uk/psychology/our-staff/academic/annette-karmiloff-smith/karmiloff-smith

Abstract: The evidence that Anderson (2014) marshals in support of his theory of neural reuse is persuasive. However, his theoretical framework currently lacks a developmental dimension. We argue that an account of the fundamental aspects of *developmental change*, as well as the *temporal context* within which change occurs, would greatly enhance Anderson's theory.

A number of theories have been put forward with the aim of unifying and providing a framework for the neural and psychological sciences. Do we need another one? According to Anderson (2014), yes, we do. Anderson argues that many current theories are based on suppositions that are demonstrably false, such as the assumption that the mind is modular (Barkow et al. 1992; Sternberg 2011) or that neurons and neuronal populations reach a state in which they can respond only to a specific and restricted

set of stimuli (Johnson 2001; 2011). On the contrary, according to Anderson, the brain is a continuously self-organising system; neurons are adaptively recruited in real time into partnerships to form hierarchies at multiple levels and timescales, but these partnerships are never truly *fixed* ("neural reuse").

We agree with Anderson's view that neural circuits can be put to multiple uses. His theory of neural reuse can indeed explain a range of phenomena, from synaesthesia to cross-modal plasticity (where a cortical area for one function [e.g., vision] is used for a different one [e.g., hearing in blind persons]). However, although Anderson's framework was motivated "first and foremost by functional, *developmental*, and evolutionary considerations" (p. 95, italics added), he actually focuses on activity at multiple *physical* – but not temporal – scales. Yet, understanding developmental processes – at multiple timescales – is crucial to understanding brain function. Therefore, we believe that an account of the fundamental aspects of developmental change, as well as the *temporal context* within which change occurs, would greatly enhance Anderson's theory.

The reason an emphasis on *development* is critical for understanding brain function is that (embodied, social) brains are adaptive "complex" systems, and complex systems are *history dependent* (Buzsaki 2006; Kelso 1995). In other words, the brain does not comprise neural and cognitive modules that passively process sensory information in order to prepare a behavioural response. Rather, much of the brain's activity is self-generated; sensory inputs help the brain to adjust its "internal connectivity and computations to the spatial and temporal metrics of the external world" (Buzsaki 2006, p. 11). That is, sensory information "gets embedded into a context, an important part of which is time" (Buzsaki 2006, p. 11). For example, auditory perception depends both on the perceiver's brain state and past experience with physical inputs, as well as on the pattern of sound vibrations reaching the cochlea. Hence, as Buzsaki (2006) argues, "each time the same stimulus is presented, it generates a somewhat different and unique trajectory in the neuronal space" (p. 13).

Anderson alludes to the importance of the temporal dimension throughout his book, but rarely does he actually take time into account. For example, how might neural reuse explain the finding that, after brain insult, recovery of function is massively dependent on *time*-related variables such as *age* at insult (e.g., Anderson et al. 2010)? Yet, such a discussion might actually strengthen Anderson's theory. For instance, it is widely assumed that early insults are distinct from those occurring in adulthood, and consequently that the infant brain is more plastic than the adult one. For example, whereas children with perinatal left-hemisphere (LH) damage acquire age-appropriate language skills, adults with similar LH lesions typically present with aphasia (Bates et al. 2001; Ricci et al. 2008). According to Johnson's (2001; 2011) Interactive Specialization framework (and the broader neuroconstructivist approach; Elman et al. 1996; Mareschal et al. 2007), this reduction in plasticity is a consequence of neural commitment, because functional brain development is characterised by a process of increased tuning or selectivity of functions (see discussion in Thomas 2003). However, recent evidence has yielded a more complex, nonlinear relationship between (for example) age of insult and recovery (Anderson et al. 2010; see Anderson et al. 2011, for review). For example, Vicki Anderson and colleagues (2010) found that children who had sustained brain injury prior to 2 years of age performed *worse* on a range of standardised neurobehavioural measures than those who had acquired brain insults after 7 years of age, indicating "greater early vulnerability" rather than "greater early plasticity". How can our current theories explain such findings? Could the early brain insults have had cascading effects on the developing system? Could the fact that plasticity occurs on different timescales – from spine plasticity (minutes) to the reorganization of large functional networks (days, months) – help explain differences in recovery outcome? Indeed, a review by Vicki Anderson et al. (2011) suggests that neither plasticity nor vulnerability theories explain the wide range of functional outcomes after brain insult (see also Giza & Prins 2006; Mosch et al. 2005). On the contrary, many different (time-related) developmental and environmental factors are involved (Anderson et al. 2011). A theory that takes development seriously (e.g., dynamic systems theory, neuroconstructivism) may have greater success in explaining these time-related phenomena.

Timing is also important in understanding the *typically developing* (TD) brain. Take learning, for example. The acquisition of new skills over the lifespan is constrained by *age* of acquisition; early learning constrains the development of later cognitive and neural structures (Hernandez & Li 2007). For example, in the language domain, young children gain much deeper grammatical sensitivity when learning new languages than (cognitively superior, even fluent-speaking) adults (Johnson & Newport 1989). Why do we not learn equally well over the lifespan? How does neural reuse help explain, for example, why adult Japanese speakers who learn English as a second language often have great difficulty in discriminating /r/ from /l/ accurately? Takagi and Mann (1995) examined the perception of /r/ and /l/ in adult Japanese learners of English as a function of their exposure to the English language. They found that, although persistent exposure (12 or more years in an English-speaking environment) improved /r/-/l/ identification accuracy, the learners never attained perfect perceptual mastery (see also Flege et al 1999; Takagi 2002; for reviews, see Birdsong 2006; Hernandez & Li 2007). Yet, infants worldwide can initially perceive all of the phonetic and phonemic differences across languages (e.g., whatever their mother tongue, early on all can discriminate between English /r/ and /l/) and have the potential to master multiple languages (Eimas 1975). Nevertheless, infants' perception of *nonnative* distinctions declines during the second half of the first year of life (Werker & Tees 1984), and their ability to discriminate *native* speech sounds increases (Kuhl et al. 2006). This developmental process (known as *perceptual narrowing*) is clearly important, because it correlates with greater language and reading skills later in life (Kuhl et al. 2008). But it also suggests that the neural commitment arising from learning a language early in development decreases the ability to learn other languages later on in development (Kuhl 2004).

Neuroconstructivism can explain perceptual narrowing. Again, this is because such accounts involve "neural commitment" (Kuhl 2004) or "restriction-of-fate" (Mareschal et al. 2007), whereby the neurocognitive functions involved in human perception gradually "specialize" over developmental time for socially relevant information – in faces, voices, music, language, and action (e.g., Lewkowicz & Ghazanfar 2009; Scott et al. 2007; Werker & Tees 2005; see Pascalis et al. 2014, for discussion). Anderson's framework can also account for these developmental data, by arguing that the early formation of functional networks constrains the emergence of future ones. Moreover, there is evidence that perceptual narrowing can be reversed in 9-month-old infants under certain conditions (Kuhl et al. 2003), which fits well with the concept of neural reuse. But a focus on developmental change of this nature is missing from Anderson's book. Precisely *what* is the difference between infants and adults in these kinds of studies? *Why* are neurons or neuronal populations not being reused to help Japanese adult learners of English (i.e., perceptual narrowing has not been reversed in children older than 9 months)? *How* are later emerging networks constrained by the existence of earlier developing ones?

Developmental change is also crucial for understanding *individual differences*. As Anderson acknowledges, the emergent behaviour of a system (e.g., social anxiety in an individual) is not reducible to the microstates of the system's parts (e.g., genes, neurons). Rather, emergent properties arise as a result of additional laws that apply at higher levels. These additional laws, which include *developmental* principles, as well as principles of functional brain organisation and social context, constrain the arrangement and behaviour of the components at levels below, and help explain why particular properties (e.g., social anxiety) are present in some individuals but not in others. Moreover, because complex systems tend to gravitate towards certain

stable patterns of activity (Kitano 2004; Thelen & Smith 2006), understanding developmental processes can also help explain why qualitatively new and unexpected functional states emerge (e.g., psychosis). Furthermore, it can help explain why some states (e.g., psychosis) are relatively common. Some states can induce changes in the system that make it more likely for them to re-emerge in the future. For example, there is evidence that normal homeostatic processes in dopaminergic signalling pathways can, under certain circumstances, compound the effects of perinatal deficits in the brain – and subsequently drive the system into psychosis and (thereafter) schizophrenia (e.g., Lisman et al. 2008). Hence, brain structure and function are deeply yoked to developmental processes, and therefore we must consider developmental trajectories when striving to account for individual differences and especially when trying to understand how various neurodevelopmental disorders arise (Karmiloff-Smith 1998). This is especially important if, as Anderson contends, relations between neurons (or populations of neurons) are never fixed but vary over time.

Therefore, thinking *developmentally* can help the scientist to understand a range of phenomena that are related to brain structure and function. Important developmental questions include: Is there more neural search early on in life? How and when ontogenetically does neural reuse actually occur? Some groups of neurons (e.g., those in anterior insula) are more functionally diverse than others (e.g., those in primary visual cortex) (Anderson 2014): Do all groups of neurons start off functionally diverse, with some becoming more specialised than others over developmental time? Do the (low-level) functional biases of local networks also change over time? As coalitions of neurons emerge and grow in strength, does the brain become progressively less plastic? If neurons used for processing objects become co-opted into learning a language, does that affect their role in object processing? Or can neurons be "reused" without any cost (see Liu et al. 2015, for evidence of neural trade-offs)? Does the brain develop from a relatively labile state to one that falls in between labile and stable, or is it always labile? The neural reuse theory accounts for rapid online adaptivity to changing environments. But is this plasticity unconstrained – i.e., might it produce a brain that is *too* malleable (Bavelier et al. 2010; Hensch 2005; Karmiloff-Smith 2015)? If neurons can be co-opted into different coalitions without any cost, then how would neural reuse explain phenomena such as the developmental "rise and fall" of synaptic density (Goddings & Giedd 2014; Huttenlocher 1990; 1994), "U-shaped" curves in development (e.g., Casey et al. 2010; Giedd & Rapoport 2010; see Rogers et al. 2004, for discussion), or perceptual narrowing in infants (Werker & Tees 1984)? The preceding are all fundamentally developmental questions. A developmental perspective is, in fact, crucial because the *timing* of developmental events is likely to be an important constraining factor.

In conclusion, Anderson's idea that neurons are recruited into partnerships that are never fixed is a compelling one. However, we argue that at every level of an adaptive, complex system, one cannot ignore its developmental history and environmental context. Hence, an understanding of the human mind/brain requires a truly developmental perspective that, in our view, would greatly enhance Anderson's theoretical framework.

Becoming an expert: Ontogeny of expertise as an example of neural reuse

Alessandro Guida,[a] Guillermo Campitelli,[b] and Fernand Gobet[c]

[a]Department of Psychology, Université Rennes 2, Rennes, 35043 Rennes Cedex, France; [b]School of Psychology and Social Science, Edith Cowan University, Perth WA6027, Australia; [c]Institute of Psychology, Health and Society, University of Liverpool, Liverpool L69 7ZA, United Kingdom.
alessandro.guida@univ-rennes2.fr   g.campitelli@ecu.edu.au
fernand.gobet@liverpool.ac.uk
http://alessandro-guida.blogspot.fr/
http://gcampitelli.com
http://www.chrest.info/fg/home.htm

Abstract: In this commentary, we discuss an important pattern of results in the literature on the neural basis of expertise: (a) decrease of cerebral activation at the beginning of acquisition of expertise and (b) functional cerebral reorganization as a consequence of years of practice. We show how these two results can be integrated with the neural reuse framework.

In *After Phrenology*, Anderson (2014) presented the *neural reuse* framework, which opposes both modular and holistic views of brain architecture. In evolutionary terms, neural reuse claims that the brain evolves, not by adding new specialized modules, but by acquiring new functions by recombining local areas in new ways. Ontogenetically, neural reuse involves the processes of *interactive differentiation* (local areas have different profiles and interact in different ways as a function of development) and *neural search* (the active testing of multiple neuronal combinations until finding the most appropriate one for a specific skill, i.e., the neural niche of that skill).

One testable hypothesis of the neural reuse framework is that novices in a domain of expertise show widespread brain activation when performing a domain-specific task, whereas experts show a more focused pattern of brain activity. This is because, at the beginning of acquiring a skill, the brain is searching for an appropriate combination of areas, whereas the experts' brains have already settled in a specific network of brain areas to perform domain-specific tasks. Indeed, Anderson (2014) presents data supporting this hypothesis (e.g., Merabet et al. 2008; Petersen et al. 1998; Petersson et al. 1997; Poldrack et al. 1998).

The field of research on the neural implementation of expertise has been very prolific in the last 15 years; therefore, it seems pertinent to evaluate Anderson's hypothesis exhaustively in light of new data. In fact, Guida et al. (2012; see also Guida et al. 2013) reviewed the literature on neural implementation of expertise in tasks related to working memory. They found two effects: (a) studies investigating individuals who receive training in working memory-related tasks (from two hours up to five weeks) show mainly a decrease of cerebral activity in prefrontal and parietal working memory areas after training, whereas (b) studies using experts and novices in different fields performing domain-specific working memory-related tasks tend to show that the brain areas activated to perform those tasks differ between novices and experts (Guida et al. [2012] referred to this effect as *functional neural reorganization*). As suggested by Anderson (2014), the first effect (i.e., the reduction of brain activity as a result of a number of hours of training) is consistent with the process of neural search. Even though the second effect was not envisaged by Anderson (2014), we propose that it is also compatible with neural search. The first effect reflects the fact that a developed skill finds a neural niche within the network of brain areas used at the beginning of skill acquisition, whereas functional neural reorganization reflects a more radical type of neural search: the skill finds its neural niche in a different set of brain areas. In the rest of this commentary, we explain in more detail these two effects, which are connected by Guida and colleagues (Guida et al. 2012; 2013) in a two-stage framework, and link them to the three implications of the neural reuse framework that Anderson put forward in his précis (sect. 1, para. 3).

Experts: Reusing the mediotemporal lobe. (1) *"First and most obvious, newly acquired capacities are generally supported by mixing and matching the same neural elements in new ways"* (sect. 1, para. 3). This first implication is in accordance with the expertise literature. As a consequence of their extended practice, experts develop domain-specific knowledge structures (i.e., chunks and more sophisticated knowledge structures; see an

explanation in the next section). These new knowledge structures allow experts to reuse the mediotemporal lobe in a completely different way compared to novices (Campitelli et al. 2007; Guida et al. 2012; 2013). Whereas novices typically use episodic long-term memory areas (e.g., the mediotemporal lobe) for performing long-term memory tasks, experts are able to (re)use these areas also for performing working-memory tasks.

As theorized by Ericsson and Kintsch (1995) and Gobet and Simon (1996), this reuse occurs when there is a tight connection between working memory and long-term memory through retrieval cues or slotted schemas, which allow a fast transfer of information between these two types of memory (see also Gobet 2000a; 2000b). Therefore, the interaction between working memory and long-term memory is crucial for functional neural reorganization to take place. This certainly echoes Anderson's (2014, p. 40) view that "function depends much more on the *interactions between* parts than on the *actions of* parts," and relates to Anderson's second principle of a functionalist neuroscience: "Our complex and diverse behavioral repertoire is supported primarily by the brain's ability to dynamically establish multiple different functional coalitions" (2014, p. 296).

The relation between working memory and episodic long-term memory also relates to another interesting effect described by Anderson: unmasking. The basic idea is that brain regions are supposed to be specialized to process one type of input (e.g., the occipital cortex is supposed to be specialized in processing visual input). However, under special conditions in which the source of dominant input is disrupted (e.g., injury, sensory deprivation), this brain area can process a different type of input, unmasking a new processing capacity for this area. Moreover, Anderson argues that the disruption of normal input is not a necessary condition and that unmasking can be observed also under more typical conditions. The development of expertise, as postulated by the two-stage framework, offers a good example of this. It is indeed possible that the decrease of activity in the first stage may help the unmasking and thus reuse of the mediotemporal lobe, allowing the occurrence of the second stage. We also agree with Anderson when he proposes that unmasking must not be seen as passive. From our point of view, functional neural reorganization occurs through the use of knowledge structures. The efficiency of these structures is a necessary condition. However, the biological reasons that undergird such processes are unknown.

Experts: Re-using spatial processes. (2) "*Second, and perhaps less obvious, neural reuse would appear to support and encourage* procedural *and* behavioral *reuse*" (sect. 1, para. 3). Guida et al. (2012) provided an explanation of the cognitive processes that both cause and are the consequence of the two identified patterns of brain activity in expertise studies. The first stage – decrease of cerebral activity – has been linked to chunking (Chase & Simon 1973; Cowan 2001; Gobet et al. 2001). When practice begins, individuals start binding various domain-specific patterns (e.g., in chess, configurations of pieces) together, which ultimately result in a compression (Mathy & Feldman 2012) of the elements into one structure, a chunk. Once chunks are built, separate domain-specific patterns can be processed as one element, which means that less cognitive resources are needed, and this is reflected as a reduction of brain activation to perform a domain-specific task. As a consequence of practice and expertise, chunks grow in size (e.g., Cowan et al. 2004; Chen & Cowan 2005) and complexity (e.g., Chase & Simon 1973; Gobet & Simon 1996), and with years of training, they become high hierarchical chunks: knowledge structures. These structures (Ericsson & Kintsch 1995; Gobet & Simon 1996) allow experts to encode information in episodic long-term memory in a fast and reliable fashion even in conditions typical of working-memory tasks (rapid presentation of several elements), which is not possible for novices using similar cell assemblies in the mediotemporal lobe.

As pointed out above, to be able to use episodic long-term memory in a fast and reliable fashion, individuals must develop specific knowledge structures. This illustrates how human beings, "repurpose our behavioral routines in multiple circumstances for myriad cognitive ends" (sect. 7, para. 3). A well-known example in the domain of expertise is the *method of loci*, which is thought to be the first (internal) mnemonic (Worthen & Hunt 2011; Yates 1966), initially proposed by Simonides of Ceos more than two millennia ago. In ancient Greece, orators would visualize a sequence of familiar locations (in their house or a familiar route with salient locations) before a speech and use them to mentally store important words. Subsequently, during their speech, they would take a mental tour and retrieve each word through the familiar locations. This technique is still in use among expert mnemonists (Pridmore 2013). Maguire et al. (2003) revealed the functional cerebral pattern of these mental walks by comparing mnemonists with all-comers. They found comparable activations in both groups in prefrontal working memory areas, but specific activations for the mnemonists, in the left medial superior parietal cortex, in the bilateral retrosplenial cortex, and in the right posterior hippocampus (for a similar pattern, see also Pesenti et al. 2001); these regions are important for episodic memory and crucial for spatial memory and navigation (e.g., Burgess et al. 2002). Therefore, it seems that with hours of training, mnemonists are capable of using the mental image of their house (or of a route) as a slotted schema and transfer the incoming information from working memory to long-term memory by associating the new information with each slot (the familiar locations), instead of simply storing the information in working memory. The consequence of the use of the method of loci is an increased memory capacity and cerebral functional reorganization.

This example clearly shows how experts reuse spatial cognitive processes to encode verbal information. These elements are linked to section 7 of Anderson's précis, "Reuse, interaction, and 'higher-order' cognition," and perfectly illustrate the fact that "we have found ways to reuse our physical capacities to augment our mental ones; in a process supported by *neural* reuse" (sect. 7, para. 3). Interestingly, a similar process has also been found in "all-comers." Van Dijck and Fias (2011, see also Guida et al. 2015) showed that verbal information processed in immediate memory was mentally organized from left to right based on the order of presentation. This suggests that order in working memory could be coded through spatial positional tagging (Abrahamse et al. 2014; Guida & Lavielle-Guida 2014). Based on this idea, Guida and Lavielle-Guida (2014) proposed that spatial positional tags in all-comers were comparable to the spatial locations of expert mnemonists, proposing the generic term of spatialization. The final twist that links all-comers to expertise is that the left to right spatial positional tagging observed in all-comers could be the result of expertise in reading and writing.

Experts versus novices: Same working-memory tasks but different processes and cerebral substratum. (3) "*The third implication … not every cognitive achievement … need be supported by a specific targeted adaptation*" (sect. 1, para. 3). As emphasized by Anderson, the last implication follows neatly from the previous points. Indeed, we believe that the assemblies of neurons that code for location did not evolve for encoding and retrieving words like in the method of loci. This example enters in a much wider picture when taking into account working memory. It is well established that prefronto-parietal areas are crucial for working memory in all-comers (Cowan 2011; Postle et al. 1999; Postle & D'Esposito 1999; Todd & Marois 2004; Vogel & Machizawa 2004). However, as highlighted above (see also Guida et al. 2012; 2013), when experts execute working memory related-tasks within their domain of expertise, completely different brains areas are activated (e.g., the mediotemporal lobe). In this case, expertise via new assemblies of cells allows experts to circumvent the limits of working memory by using a part of episodic long-term memory. This shows that a same cognitive achievement (here, working memory task) needs not to be supported by a specific targeted adaptation.

Conclusion. The neural reuse framework proposes that the same assemblies of cells can be used for different cognitive functions or tasks in different contexts. Research into the neural implementations of expertise supports Anderson's (2014) hypothesis that brain activity decreases and becomes more focalized when a skill is learned. However, another pattern of results in the expertise literature was not envisaged by Anderson – a change of the set of networks used to perform a working-memory task as a function of expertise. We presented an explanation of how this effect can be explained by the concept of neural reuse.

A registration problem for functional fingerprinting

David M. Kaplan[a] and Carl F. Craver[b]

[a]*Department of Cognitive Science, Macquarie University, Sydney NSW 2109, Australia;* [b]*Philosophy-Neuroscience-Psychology Program, Washington University in St. Louis, St. Louis, MO 63105.*
david.kaplan@mq.edu.au   ccraver@wustl.edu
http://www.davidmichaelkaplan.org/
https://pages.wustl.edu/cfcraver

Abstract: Functional fingerprints aggregate over heterogeneous tasks, protocols, and controls. The appearance of functional diversity might be explained by task heterogeneity and conceptual imprecision.

Anderson (2014) promises to move neuroscience beyond phrenology by rejecting strict functional localization, the idea that the brain is composed of highly selective and functionally specialized areas connected along developmentally and evolutionarily dedicated pathways. Anderson proposes a competitor idealization, the *neural reuse hypothesis*, according to which the activities of different brain regions flexibly recombine to support performance across many different task domains. Anderson supports this hypothesis in part by appeal to *functional fingerprints*, a novel methodological contribution for representing and analyzing functional diversity in the brain.

Functional fingerprinting is a data-driven tool that relies on meta-analyses of neuroimaging studies to characterize which task domains preferentially engage a given brain region. Anderson borrows his task domains from the BrainMap database (Fox et al. 2005; Laird et al. 2005). They are defined by two features: (a) a cognitive *construct* (such as working memory) and (b) a collection of tasks (and, more specifically, a set of studies) unified by the fact that they are commonly accepted ways of studying that construct. Domains include several emotions, action, attention, working memory, reasoning, vision, and others (see Fig. 2 in the target article). Fingerprints are designed to capture the functional diversity of a given brain region or network. For a brain area to be functionally involved in a task domain (for a given construct) is for it to be active during tasks that neuroscientists accept as valid for studying that construct. The functional fingerprint for that brain area is a polar plot in which vertices represent different task domains. Distances along each vertex represent the number of activations at a given site for a particular task domain expressed as a percentage of the total activations reported at that site across all sampled task domains. Anderson extends this idea to explore the functional diversity of brain networks, but this extension relies fundamentally on the more basic project of constructing the fingerprint itself.

This method is prone to the *problem of functional registration*. Anderson's fingerprinting method aggregates findings obtained in fMRI studies using diverse experimental task conditions, distinct subtraction conditions (controls), and distinct experimental protocols. Given the diversity of tasks, controls, and protocols, one would expect to observe activation in regions that are nonspecific to the domain-defining psychological construct under investigation. Performance across different experimental task and control conditions will often rely on different cognitive capacities and will therefore recruit different underlying neural mechanisms, leading to differences detectable in neuroimaging experiments (e.g., Owen et al. 2005; Price et al. 2005). Fed into Anderson's method, such nonspecific activations will functionally implicate a region in a task domain simply because it was not controlled for in the task in question. As a result, failures to register differences between tasks, controls, and protocols within a given task domain will contaminate one's measurements of functional diversity with extraneous and ancillary activations tied to aspects of the comparison that were either irrelevant or simply uncontrolled for in the context of the original studies. Our suspicion is that Anderson's method glosses over heterogeneity in task and control conditions to a degree that could explain the functional diversity he reports.

To illustrate this, suppose for the moment (as Anderson does) that we accept the BrainMap taxonomy as a more or less correct taxonomy of cognitive capacities or functions. Anderson does not characterize precisely how task-relevant activations are sorted from task-irrelevant activations, but it is difficult to envision how this could be done systematically for all studies subsumed within a given meta-analysis in a way that avoids the perils of simply associating activations with tasks and tasks with constructs. Consider the task domain of working memory, for example.

Owen et al.'s (2005) recent meta-analysis of working memory activations focuses specifically on 24 studies employing the so-called *n*-back task (just *one* type of task associated with the working memory task domain in BrainMap). Although all of these studies nominally employ the same task, Owen et al.'s (2005) systematic cataloging of different parameters used in the *n*-back task reveals considerable task diversity. In particular, they identify four major categories of *n*-back task (location monitoring, identity monitoring, verbal stimuli, and nonverbal stimuli), which can be further subdivided along a number of finer-grained dimensions including how many trials back subjects are matching (n = 1-, 2-, 3-back). These *n*-back studies also differ substantially in the chosen contrast (i.e., the control condition used). For example, a task subtraction might subtract activation observed in the $n = 3$ condition from the activation observed in the $n = 2$ condition, it might subtract activation in $n = 2$ from that in $n = 0$, it might subtract activation during matching of Korean words from that of English words, it might subtract activation in response to letters from that in response to shapes, or it might reflect monotonic increases in task difficulty.

Surprisingly, Owen et al. (2005) report that despite this task diversity, some frontal and parietal activations are consistent across these different task conditions. This result is surprising and valuable precisely because it reveals the signal in the noise. One does not expect such tidy results emerging from such a motley collection of experimental paradigms. Yet, critically, Owen et al. (2005) also show that there are differences in activations depending on whether the material is presented visually or aurally and on whether the task involves identity or location monitoring. No task is "pure" in the sense that it requires all and only the mechanisms responsible for a given task domain. When one pools data across different tasks that are "impure" in different ways, one is likely to aggregate over ancillary activations resulting from aspects of the task not specific to the construct in question: in other words, the false appearance of functional diversity. And this is the primary point: there will be many regions showing nonspecific activation that do not overlap between these task presentations. Although these diverse regions of nonoverlap are not the focus in Owen et al.'s (2005) meta-analysis, they are central to interpreting Anderson's findings because they are the data points for his functional fingerprints. The appearance of functional diversity could hence result from the incautious pooling of data from heterogeneous tasks and protocols employing distinct control conditions.

Anderson's fingerprints are a kind of aggregate "reverse inference" (from activation during a task to functional involvement in

the construct/task domain), but without the careful attention to task construction and control required in each case to make the reverse inference convincing. Traditional problems with reverse inference in neuroimaging (such as the existence of nonspecific activations during task performance) are thus both multiplied and obscured in Anderson's functional fingerprints. Indeed, given the diversity of protocols with which the analysis begins, one would expect evidence of functional diversity even if localization were broadly true. The challenge going forward is to devise methods that can successfully establish functional diversity as a real feature of brain organization rather than as a reflection of the heterogeneity and imprecision in our methods.

Performing an informative meta-analysis about the functional diversity of a brain region will require precisely the kind of work that should have been, and in some quarters has been, driving task-based fMRI all along: to devise task-control pairs in such a way that they isolate the areas involved in the construct under investigation independently of other ancillary activations. Anderson does not explain how tasks and controls are chosen, related to one another, or grouped into task domains in his meta-analytic method. Without this information, attempts to read off "functional involvement" directly from activation profiles each involve a separate, incautious reverse inference for each activating task hidden behind the veil of a meta-analysis.

The problem of functional registration is just a specific application of a more general challenge facing any meta-analytic approach to functional diversity such as Anderson's – to distinguish the signal of functional diversity from the inevitable and expected noise produced by experimental heterogeneity. Variability in task and control conditions is just the tip of the iceberg. Other sources of experimental "noise" in fMRI meta-analyses include differences in subject population, spatial normalization, scanner strength, and essentially any other uncontrolled variables capable of affecting experimental outcomes (for further discussion, see Brett et al. 2002; Costafreda 2009). Within the localizationist framework, the rules are clear: search for a task (or task domain) that preferentially drives the area in question. In the context of neuroimaging meta-analyses, the primary objective is to identify the consistently activated regions (if any exist) across a set of studies that are assumed to probe the same psychological state or capacity using similar or identical experimental tasks (Fox et al. 2014).

Anderson urges us to abandon (or at least, relax) these localizationist assumptions and to think instead of brain regions multitasking and recombining across different task domains. Anderson's framework predicts that brain activation patterns will tend not to show sharp functional specialization, but will instead fan out broadly across the polar graph. One limit of this framework, as it is currently developed, is that it makes no specific predictions (comparable to those made by localization), except that one will not see the functional specialization predicted by the localizationist. But if functional diversity is the expected outcome when pooling fMRI data across different experimental tasks (regardless of whether the hypothesis of localization, reuse, or some other hypothesis is correct), then the data reported in functional fingerprints fail to decide between localization and reuse. Anderson's proposed method currently lacks a principled way to sort the noise introduced by experimental heterogeneity from the signal reflecting real functional diversity in the brain. Perhaps more specific, risky predictions about the kinds of diversity one is or is not likely to see would be more compelling.

Despite these criticisms, we think that Anderson's critical perspective on classical localization is commendable. The very idea of functional diversity enjoins us to think more broadly about how functions might be localized in the brain. However, we do not think that Anderson has succeeded entirely in sketching a way to do cognitive neuroscience "without the analysis, decomposition, and localization of component cognitive operations" (Anderson 2014, p. 117). In the first place, Anderson relies on the BrainMap taxonomy of task domains and so simply embraces the dominant ideas in contemporary cognitive science concerning how brain

systems should be functionally analyzed and decomposed. (Notably, Gall [1835], one of the original phrenologists, promoted radical revision in our taxonomy of cognitive functions.) Whether a given brain region turns out to have a narrow or broad orientation around the polar graph is highly sensitive to how the vertices of the graph are defined. What appears as functional diversity through the lens of one particular taxonomy of task domains could appear as functional unity through the lens of another.

The fact that Anderson's method implicitly reifies the task domains of BrainMap brings to mind a warning issued long ago by Petersen and Fiez (1993). They counsel against assuming that the function of a brain region can be identified with the tasks used to activate it; as they prosaically remark, there is no tennis forehand area in the human brain. There is no such area, first, because the tennis forehand likely involves contributions from many distinct and dissociable cognitive processes (i.e., recruits many different task domains). Again, this is why the problem of functional registration is a difficult one to solve. Second, there is no such area because any particular experimental task (including performing a tennis forehand) is at best a proxy for or representative of some broader class of behavioral or cognitive phenomena that is the real target of explanation. The functions that ultimately get localized in the brain might therefore be very distant from the tasks that are paradigmatically used in our experimental investigations. The general lesson here is that the conceptual relationships between tasks, task domains, and cognitive constructs is complex and dynamic, and cannot be taken for granted without costs.

With the above points taken into consideration, Anderson's neural reuse hypothesis might be understood, not as a complete rejection of localization, but rather as a form of localization consistent with dominant attitudes in the contemporary neuroimaging community (Petersen & Fiez 1993). According to this approach, elementary operations, not tasks, are functionally localized to brain regions. Recent work on so-called canonical neural computations – i.e., standard computational operations applied across different brain areas – reinforces this idea (Carandini & Heeger 2012). According to this view, elementary operations might be rather task-general and might be flexibly recombined in many different task domains. The picture is still localizationist, but the localized functions are conceptually distant from traditional task domains and psychological constructs. These areas will be functionally diverse from the point of view of the BrainMap task domains, but functionally unitary once the correct elementary operation has been identified. Regardless, we will continue to face the challenge of separating diversity in the brain from messiness in our cognitive categories and from imprecision and heterogeneity in our experimental tasks.

The reification objection to bottom-up cognitive ontology revision

Joseph B. McCaffrey and Edouard Machery

Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA 15217.
jbm48@pitt.edu   machery@pitt.edu
http://www.josephbmccaffrey.com
http://www.hps.pitt.edu/profile/machery.php

Abstract: Anderson (2014) proposes a bottom-up approach to cognitive ontology revision: Neuroscientists should revise their taxonomies of cognitive constructs on the basis of brain activation patterns across many tasks. We argue that such bottom-up proposal is bound to commit a mistake of reification: It treats the abstract mathematical entities uncovered by dimension reduction techniques as if they were real psychological entities.

Reverse inference consists in inferring that a task recruits a psychological process (*P*) on the grounds that a brain structure (*S*) is activated during this task (as observed by, e.g., fMRI). It is often assumed that reverse inference is valid only if activation is *selective*, that is, if the ratio *P*(*activation of S/P is recruited*)/*P*(*activation of S/P is not recruited*) is high (Poldrack 2006). Because brain areas are typically multifunctional, cognitive neuroscientists have grown skeptical of area-based reverse inference. Anderson endorses this pessimistic conclusion – "It should go without saying that we must also curtail the common practice of reverse inference" (Anderson 2014, p. 113) – and the first two chapters of *After Phrenology* (Anderson 2014) extensively review the multifunctionality, hence low selectivity, of brain regions.

One can address the problem raised by multifunctionality in three different ways. First, reverse inference can be reformulated to depend on *diagnosticity* instead of selectivity (Machery 2014). In this approach, reverse inference is valid only if the activation discriminates between the recruitment of a first psychological process, *P*, and of a second psychological process, *P′*, that is, only if the ratio *P*(*activation of S/P is recruited*)/*P*(*activation of S/P′ is recruited*) is high. Second, one can increase the selectivity of brain activation by revising cognitive neuroscientists' *brain ontology*: Instead of focusing on regional activation, one can reverse infer on the basis of activation in other brain structures (e.g., networks) that may be selectively associated with psychological processes (e.g., Glymour & Hanson, forthcoming). In chapter 4 of *After Phrenology*, Anderson rejects this second approach on the grounds that brain networks too can be multifunctional. Anderson's concern here is speculative, and more evidence is needed before discrediting brain ontology revision. Large-scale brain networks (e.g., effective connectivity networks), or activation patterns within those networks (e.g., as measured by MVPA), may be far more selective or diagnostic than individual regions. Third, one can increase the selectivity of brain activation by revising cognitive neuroscientists' *cognitive ontology*: On this approach, activation of brain structures is not selective because cognitive neuroscientists lack the right set of cognitive constructs for describing the functions or computations that these structures perform (e.g., Poldrack 2010).

This third approach has led to a lively debate about cognitive ontology revision (Klein 2012; Lenartowicz et al. 2010; McCaffrey 2015; Poldrack et al. 2009; Price & Friston 2005). As Anderson perspicuously notes, most "revisionists" have a *conservative* goal: Taking current cognitive ontology as their starting point, they attempt to validate cognitive constructs by investigating whether they can be selectively associated with brain activation patterns (e.g., Lenartowicz et al. 2010). By contrast, chapter 4 of *After Phrenology* advocates a *revolutionary* goal. Anderson's project is not to determine which members of current cognitive ontologies are valid and which are invalid, but rather to propose entirely new cognitive constructs by mining fMRI datasets. Before describing and assessing Anderson's proposal, we note that it is unclear whether his goal is to revolutionize the constructs *psychologists* are working with (e.g., recommending they stop using the construct of working memory) or, less ambitiously, whether he is proposing a new cognitive ontology for *cognitive neuroscientists*: In this case, the idea would be to develop novel ways of characterizing what neural structures do.

Anderson's central idea is that cognitive neuroscientists should not characterize *the intrinsic function* of each brain region – that is, the operation the region performs independently of its neural context (e.g., its computational function); instead, they should quantitatively characterize each region's *disposition* to be involved in a given set of tasks. Anderson calls such dispositions "neural personalities." Neural personalities allegedly vary with respect to some fundamental psychological dimensions (or "neuroscientifically relevant psychological (NRP) factors"), exactly as personality varies with respect to a few dimensions (e.g., extraversion). The dimensions of neural personality need not correspond to existing cognitive constructs, and they must be discovered by examining brain activation across many tasks (more on this below).

Several points about Anderson's proposal are noteworthy. First, the focus on neural personalities instead of intrinsic functions is a radical change of heart for Anderson, who previously advocated characterizing regions' *workings* – roughly, their context-insensitive computational functions (Anderson 2010). Second, it is not clear whether Anderson denies that brain regions have intrinsic functions or merely thinks the best strategy for cognitive neuroscientists is to characterize their neural personalities, while conceding that future efforts could identify their intrinsic functions. The anti-computationalist rhetoric in *After Phrenology* suggests the former, but more guarded remarks support the latter. Third, Anderson mainly resists the call to revise brain ontology, focusing mostly on the brain structures – that is, individual regions – that cognitive neuroscientists have traditionally studied. In this respect, *After Phrenology* is surprisingly conservative. Fourth, Anderson's focus on neural personalities implies that, in contrast to Poldrack's approach, the search for selective activation plays no role in cognitive ontology revision: A "central point of this book is not just that we don't *get* selectivity in the brain but that *we don't need it*. *We can stop looking for it*" (2014, p. 141, emphasis in the original). Fifth, Anderson proposes to identify the dimensions of neural personalities (the NRP factors) in a strictly *bottom-up* manner: The proposal is to infer these new cognitive constructs from the brain's "behavior" – its activation patterns – across many tasks. In this respect, *After Phrenology* is surprisingly radical. Cognitive neuroscientists typically impose existing cognitive constructs onto the brain to interpret task-related activation. Instead, Anderson proposes using brain activation patterns across tasks to determine their psychological nature – what the tasks have in common and how they differ from a psychological point of view: "[O]ne can (…) use these data [i.e., the data from imaging experiments] to let the brain tell us something about these experiments – to reveal the underlying attributes of the task situation to which the brain differentially responds" (2014, p. 138).

How should researchers interpret NRP factors (the dimensions along which neural personalities vary) and neural personalities themselves? There are two ways of interpreting them: an *instrumentalist* or a *realist* interpretation. According to the instrumentalist interpretation, these dimensions (NRP factors) are just a way of *summarizing* how similar the brain activation patterns elicited by the tasks under consideration are, and ascribing a neural personality to a brain area is just nothing more than a way of summarizing the data showing how this area is differentially active in a set of tasks. According to the realist interpretation, the dimensions of neural personality are *real* psychological constructs: That is, they can feature in causal explanations. *After Phrenology* is unclear about which of these two interpretations is correct, but Anderson appears to view NRP factors as explanatory and causal: "NRP factors should be understood as a region's disposition to help shape an organism's behavior in a situation, to help determine the character of the organism's interaction with its environment, or to manage some aspect of the organism-environment relationship" (2014, p. 151). These two interpretations of neural personalities should be familiar to readers acquainted with the history of psychology: Psychologists have long debated whether traits such as IQ or personality dimensions should be interpreted instrumentally or realistically.

Our main contention is that, just like other attempts at revising cognitive ontologies in a strictly bottom-up manner, Anderson's revolutionary endeavor to develop new cognitive constructs – the NRP factors and the neural personalities – can be interpreted only *instrumentally*, and that this is in tension with his goal of developing a new set of causally explanatory cognitive constructs. To characterize brain areas' dispositions, Anderson first appeals to the notion of a *functional fingerprint* developed by Passingham et al. (2002) (Anderson 2014, sect. 4.2; Anderson et al. 2013; Uddin et al. 2014). Identifying a region's functional fingerprint begins with categorizing the tasks in the fMRI literature on this area as recruiting one of several psychological processes. Anderson and

colleagues (2013) typically use a coarse-grained categorization scheme, distinguishing about 20 processes such as vision, attention, phonology, semantics, learning, or working memory. This scheme allows them to represent quantitatively how often, according to a given literature, a given area is activated when one of these 20 processes is recruited by an experimental task, for example how often articles studying working memory report activation in the dorsal anterior insula. The pattern of recruitment of a given area, given a particular set of fMRI articles and a categorization scheme, is its functional fingerprint. Although, unsurprisingly, areas tend to be activated by many processes, their functional fingerprints vary. Importantly, a functional fingerprint is a mere *summary* of a data set: It does not explain why the area is activated the way it is.

Following Poldrack et al. (2009), Anderson (2014, sects. 4.3 and 4.4) proposes to use dimension reduction techniques (factor analysis, MDS, PCA, etc.) to identify a few dimensions explaining why an area has its functional fingerprint. Instead of merely summarizing the involvement of a given area in a set of tasks, as functional fingerprints do, neural personalities *explain* this involvement: They allow cognitive neuroscientists to claim that *because* an area has a given neural personality (its score is i on NRP factor 1, j on NRP factor 2, etc.), it is involved more in some tasks than in others.

However, dimension reduction techniques are ill suited for discovering new cognitive constructs (Glymour 2001; Gould 1996). These statistical techniques project high-dimensional spaces onto spaces with fewer dimensions. On their own, the resulting dimensions cannot be interpreted realistically; they merely provide convenient ways of summarizing high-dimensional data. Three main arguments support this deflationary understanding of dimension reduction techniques. First, the outcome of these techniques is *underdetermined*. A given set of vectors in a high-dimensional space can be projected onto different spaces with different dimensions. To highlight merely three issues, there are many nonequivalent dimension reduction techniques, the number of dimensions is typically arbitrarily chosen, and these dimensions can be oriented in different manners. None of the possible spaces should be interpreted realistically because it would be arbitrary to treat one of them as real to the detriment of the others. Second, just like causally-based correlations, *accidental* correlations can be projected onto a lower-dimensional space, resulting in meaningless dimensions (e.g., Gould 1996, p. 280). Hence, that a high-dimensional space can be projected onto a lower-dimensional space does not justify interpreting the resulting dimensions realistically. Finally, the capacity of dimension reduction techniques such as factor analysis to identify causes has not been validated (Glymour 2001, Ch. 14). These three arguments bear on Anderson's project, exactly as they bear on IQ and personality research: On their own, dimension reduction techniques do not justify interpreting the dimensions of neural personalities realistically. Forgetting their limitations is committing the error of *reification* – namely, presuming that the abstract mathematical entities uncovered by dimension reduction analyses correspond to real psychological entities.

Naturally, the products of dimension reduction techniques can sometimes be interpreted realistically instead of as mere instruments for summarizing high-dimensional data. To do so scientists need to bring their broader empirical knowledge to bear on the interpretation of the dimensions of the lower-dimensional space. In the present context, this means that a purely bottom-up approach to cognitive ontology revision is unlikely to succeed: Some other information beyond the activation of brain areas across a range of tasks and their dimension reduction is needed to interpret the resulting dimensions realistically. Perhaps it is also worth noting that establishing the predictive validity of neural personalities does not justify understanding them realistically.

Anderson's approach to cognitive ontology revision is not the only one to fall prey to this *reification objection*; in fact, we speculate that in general purely bottom-up cognitive ontology revisions

commit the error of reification (e.g., Poldrack et al. 2009). Such approaches must reduce the very high-dimensional space defined by the number of voxels considered in order to identify cognitive constructs defined solely by brain activation patterns. Doing so probably requires using techniques whose product cannot be interpreted realistically. In our opinion, the reification objection reveals a fundamental shortcoming of bottom-up cognitive ontology revision.

Reason for optimism: How a shifting focus on neural population codes is moving cognitive neuroscience beyond phrenology

Carolyn Parkinson[a] and Thalia Wheatley[b]

[a]Department of Psychology, University of California, Los Angeles (UCLA), Los Angeles, CA 90095; [b]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755.
cparkinson@ucla.edu   thalia.p.wheatley@dartmouth.edu
http://csnlab.org
http://www.wheatlab.com

Abstract: Multivariate pattern analysis can address many of the challenges for cognitive neuroscience highlighted in *After Phrenology* (Anderson 2014) by illuminating the information content of brain regions and by providing insight into whether functional overlap reflects the recruitment of common or distinct computational mechanisms. Further, failing to consider submaximal but reliable population responses can lead to an overly modular account of brain function.

There is much to like in Michael Anderson's new book, *After Phrenology* (2014). First and foremost, he synthesizes a large and diverse literature to support a sophisticated and evolutionarily grounded approach to brain science. He describes the brain as a system in constant flux, transiently coalescing competing and cooperating neural assemblies in the service of action. The brain, he reasons, evolved to engage successfully with the environment rather than to represent the world accurately. He argues convincingly that this difference is underappreciated by cognitive neuroscience, which instead tends to look for neural activity that encodes objective properties in the environment as if they are context- and perceiver-free.

Anderson's description of the brain as a dynamic information processing device that responds to structured signals from the environment to guide action dovetails with his and other theories of neural reuse (Anderson 2010; Chang et al. 2013; Dehaene & Cohen 2007; Marcus 2006; Parkinson & Wheatley 2015). The basic idea here is that the brain solves new problems by repurposing preexisting neural architecture that solved structurally similar problems. This evolutionary and action-focused account of the brain is interwoven with the "extended mind" thesis, that intelligence is subserved by iterative interaction with the environment. These ideas form a solid, well-articulated, and reasonable book.

Less convincingly, Anderson argues that cognitive neuroscience is held in thrall to a "massively modular" account of the brain. From Anderson's vantage point, the field's modus operandi is to test for and report the single unified function of each brain area, as if the brain is a "collection of organs" (p. 292). As he points out, a brain composed mostly of special-purpose modules would be inefficient, prohibitively large, and inconsistent with the bulk of currently available evidence. We agree with Anderson that massive modularity is untenable, but so too does all but a vanishing minority of neuroscientists. The near-consensus view among contemporary neuroscientists is that most cortical regions exhibit considerable functional heterogeneity. Hence, the pleas to jettison modularity feel dated.

Anderson correctly asserts that the interconnected complexity and flexibility of neural activity presents immense challenges for brain science. The categories that we use to reflect on our own cognition are unlikely to match the functional organization of the brain, and a given brain region likely serves multiple functional roles, which can differ depending on the current environmental context and the networks in which the brain region is currently participating. However, we are somewhat more optimistic that the field can make significant headway in uncovering the structure of operations underlying human thought and behavior. This optimism is fueled in large part by more than a decade of cognitive neuroscience research that has honed techniques for multivariate pattern analysis of functional neuroimaging data. *After Phrenology* briefly alludes to this approach (Kriegeskorte et al. 2008a), but we believe that it is of particular relevance to Anderson's proposed framework and to his suggested ways forward for the field, and hence, merits further emphasis.

We now know that the brain encodes many kinds of information (e.g., memories, tastes, reach trajectories) in neuronal population codes, rather than in the magnitude of single units' responses (Georgopoulos et al. 1988; Lin et al. 2006; Pouget et al. 2000; Simon et al. 2006). Fortunately, much of the information contained in neuronal population codes can be gleaned from the indirect and spatially coarse measurements of brain activity that are currently available to researchers studying human brain function (e.g., functional magnetic resonance imaging, fMRI; Kriegeskorte et al. 2008b).

However, much of the information present in neural population codes is lost when analyzing the magnitude of responses of a given functional unit in isolation. For example, one of the first studies that analyzed fMRI data in terms of population response vectors demonstrated that brain regions that are typically characterized in terms of the categories of stimuli eliciting maximal responses (e.g., the fusiform face area) also differentiate between stimulus categories that evoke submaximal responses, such as chairs, shoes, cats, and houses (Haxby et al. 2001). This information was carried in unique and consistent topographic arrangements of responses for each stimulus category, and such information had previously been missed by analyzing data in terms of response magnitude only. Hence, analyzing neural responses only in terms of overall response magnitude can lead to an exaggeratedly modular conceptualization of brain function, as it can underestimate the number of brain regions involved in processing a stimulus category, as well as the number of stimuli a given brain region is involved in processing. Therefore, the brain may be even less modular than the functional fingerprinting method that Anderson employs would suggest, given that this method appears to only take into account response magnitude.

Analyzing data in terms of population response vectors, rather than overall response magnitude, can also provide insight into how information is organized within brain regions. To this end, researchers have applied both supervised and unsupervised machine learning techniques when analyzing the distributed topographies of activity within brain regions in order to characterize the distinctions that those brain regions make about various classes of stimuli. For example, by characterizing the similarity structure of population response vectors to visual stimuli within brain regions of the ventral visual pathway, researchers have gained insight into the distinctions that brain areas within that pathway make about sets of stimuli, and hence, the operations that they may contribute to visual processing (Connolly et al. 2012). Studies using unsupervised learning techniques (e.g., clustering, multidimensional scaling) have the notable benefit of being wholly data-driven in the structure that is revealed. Hence, such approaches will be useful in following Anderson's call for a science focused on illuminating the psychological factors that best account for observed differences in neural activity, rather than on attempts to map psychological constructs onto the brain.

Considering the information content of multivoxel response patterns, rather than focusing solely on response magnitude,

also allows researchers to probe the functional significance of overlap in the brain regions that are recruited to process diverse domains of contents. Univariate, subtraction-based approaches to analyzing brain activity are relatively well suited to asking questions such as, "Is this brain area involved in both Task X and Task Y?" and "Is this brain area recruited 'more' for Task X than Task Y?" In contrast, analyzing neural population responses is an approach that is relatively better suited to uncovering *how* a brain region might contribute to various tasks. Functional overlap can reflect the recruitment of shared or distinct mechanisms, and these possibilities can sometimes be disentangled by studying neural population response vectors. For example, although univariate, magnitude-based analyses of fMRI data suggested that most voxels in a region of lateral occipital-temporal cortex responded to both static pictures of bodies and movement displays, subsequent analysis of the population responses in that region demonstrated that the response patterns to these two sets of stimuli, though spatially overlapping, were unrelated (Peelen & Downing 2007). Observations such as these are consistent with Anderson's suggestion that a brain region's computational role can differ depending on the circumstances surrounding its recruitment (e.g., the processing demands of a stimulus, the other brain regions with which it is interacting). Hence, analyzing population response vectors can help determine whether spatially overlapping functional activity for two different tasks reflects the recruitment of the same or dissociable underlying mechanisms.

The analysis of population response vectors can also reveal when functional overlap reflects the recruitment of a common underlying neural operation across multiple experimental contexts. For example, a machine learning classifier trained to distinguish between population response vectors within the superior parietal lobule (SPL) associated with leftward and rightward eye movements can correctly distinguish between population response vectors in the SPL corresponding to mental subtraction and mental arithmetic (Knops et al. 2009). This suggests that the SPL contributes shared operations when shifting spatial attention within the external environment (during saccades) and within internal mental representations (when performing arithmetic). This likely reflects a case of neural reuse through which functional architecture with a preexisting role in shifting attention in the external environment is exploited in order to perform analogous operations on mental representations, and hence, to perform the culturally learned skill of arithmetic.

We recently used classification and similarity structure-based analyses of neural population response vectors to test for a common neural encoding of egocentric distance in spatial, temporal, and social frames of reference (Parkinson et al. 2014). Cross-domain decoding analyses, as well as analyses of the similarity structure of population responses, revealed that areas such as the right anterior inferior parietal lobule, which has a long-established role in representing physical space, organize information not based on distance category (e.g., "Is this a social or temporal distance?"), but based on distance from the self (i.e., "Is this relatively close to or far away from me?"). These results are consistent with Anderson's suggestion that the brain is largely concerned with computing behaviorally relevant information, such as distance from oneself (and hence, relevance for action). Further, these results are consistent with the possibility that neural mechanisms for encoding information about physical space may have been reused in order to operate on increasingly abstract contents, such as temporal and social relationships (Parkinson & Wheatley 2013; 2015).

Understanding how the brain gives rise to a given facet of cognition or behavior requires not only identifying the neural substrates involved, but also uncovering the kinds of operations performed by those substrates. We are optimistic that advances in neuroimaging methods, such as the analysis of neural population response vectors, will be useful in moving toward a deeper and more veridical understanding of functional brain organization and of the flexible and adaptive computations that underlie

everyday thought and behavior. In so doing, cognitive neuroscience has the potential to elucidate the particular functional capacities that may be reused over the course of evolution and development to solve novel problems.

Multisensory integration substantiates distributed and overlapping neural networks

Achille Pasqualotto

Faculty of Arts and Social Sciences, Sabanci University, Tuzla 34956, Istanbul, Turkey.
achille@sabanciuniv.edu

Abstract: The hypothesis that highly overlapping networks underlie brain functions (neural reuse) is decisively supported by three decades of multisensory research. Multisensory areas process information from more than one sensory modality and therefore represent the best examples of neural reuse. Recent evidence of multisensory processing in primary visual cortices further indicates that neural reuse is a basic feature of the brain.

Cognitive and perceptive functions are supported by highly overlapping neural networks distributed throughout the brain, and this phenomenon can be referred to as "neural reuse" (Anderson 2010; 2014; Pessoa 2012). To use a metaphor, we might imagine the brain as a choir and neurons as the singers composing the choir; when the choir sings Song 1, some singers do not sing at all while others sing vigorously. "Active" singers represent the portion of the choir performing Song 1 (see Fig. 1a); in terms of the brain, those singers represent the neural network (parts of the brain) activated during a given cognitive or perceptive process (e.g., visual perception). When the choir sings Song 2, a slightly different, but highly overlapping portion of the singers might be active (see Fig. 1b). This exemplifies how different tasks (Song 1 and Song 2) recruit highly overlapping, but not identical, portions of the choir (see Fig. 1c); in the brain, a given cognitive or perceptive task (e.g., visual imagery) recruits a neural network highly overlapping, but not identical, to the one associated with visual perception (Ganis et al. 2004).

One of the hot topics for debate in experimental psychology and neuroscience is on the extent of *specialisation* or *distribution* of the brain functions (Anderson 2010; Driver & Noesselt 2008; Liang et al. 2013). Empirical evidence seems to favour the idea that the brain operates in a distributed, rather than specific, manner. For example, brain areas that were considered to be specialised for specific tasks (e.g., the fusiform face area for face

recognition) have been found to be activated during performance of other tasks (e.g., recognition of cars; Gauthier et al. 2000). Research has also shown that cognitive and perceptive processes usually involve networks of brain areas rather than one specific area; for example, a very specific process such as face recognition recruits a network of brain areas including the occipital, temporal, and frontal lobes rather than one specific portion of the brain (Goldstein 2009). This *distribution* of brain functions is not only seen for the processing of faces, but also for other perceptual and cognitive functions (if not all; Van Dijk et al. 2010; Yeo et al. 2011) such as perception (Takahashi et al. 2013; Uesaki & Ashida 2015), attention (Posner & Rothbart 2007; Shulman et al. 1999), memory (Alain et al. 1998; Desgranges et al. 1998), language (Duffau 2008; Horwitz & Braun 2004), spatial cognition (Burgess et al. 2001; Vallar et al. 1999), and body representation (Filippetti et al. 2015; Longo et al. 2010).

This phenomenon is not limited to the macroscopic level (i.e., brain areas), but has also been observed on the microscopic level. Some neurons have been found to respond to several types of stimuli, rather than to a specific type of stimulus, with different response patterns (e.g., firing frequencies). The difference in response patterns is the feature that distinguishes the stimuli represented by the neurons (Decharms & Zador 2000; Gerstner et al. 1997). In a simplistic example, during the processing of Stimulus A, Neuron 1 fires at a high frequency, Neuron 2 at a low frequency, and Neuron 3 at a medium frequency. Whereas, during the processing of Stimulus B, Neuron 1 fires at a medium frequency, Neuron 2 at a high frequency, and Neuron 3 does not fire at all. To reuse the above-mentioned metaphor, the same singers might participate differently in performance of different songs.

The notion that brain functions are based on distributed and overlapping neural networks is convincingly supported by the findings that input from different sensory modalities activate distributed and overlapping networks of brain areas – namely, multisensory processing (Ricciardi et al. 2014; Stein & Stanford 2008; Stein et al. 1988). Multisensory areas are portions of the brain processing input from different sensory modalities. In the last decades, an increasing number of multisensory areas have been identified (Amad et al. 2014; Gallese et al. 1996; Gobbelé et al. 2003; Sereno & Huang 2006), suggesting that the brain is more engaged in multisensory processing than was initially believed. Classic multisensory (or *associative*) areas are activated by visual, auditory, and somatosensory input and consist of prefrontal (Fuster 1988; Öngür & Price 2000), posterior parietal (Andersen et al. 1985; Serino et al. 2011), and superior temporal (Beauchamp et al. 2008; Bruce et al. 1981) cortices. More recently, new multisensory areas have been identified, in the posterior frontal (Grafton et al. 1997), temporoparietal (Matsuhashi et al. 2004), and occipitotemporal (Beauchamp 2005) cortices.
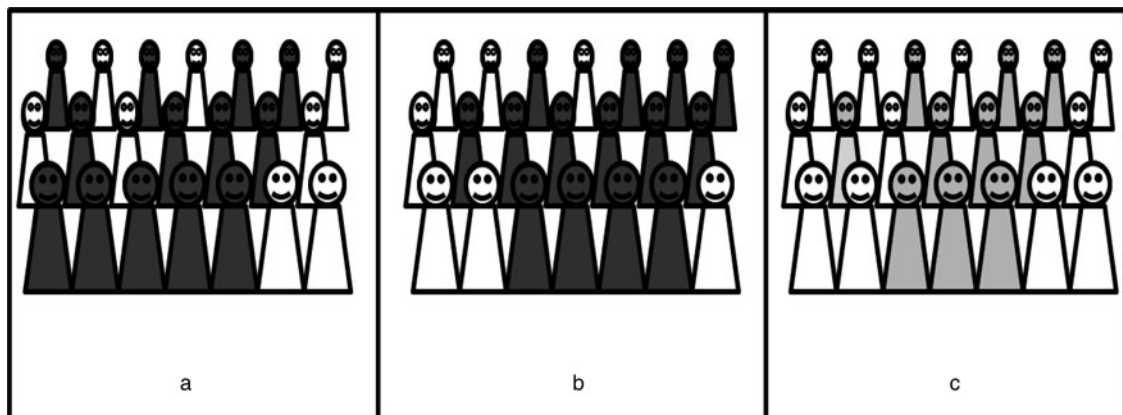

Figure 1 (Pasqualotto).   Singers of the choir singing (a) Song 1 (dark-grey); (b) Song 2 (dark-grey); (c) both Songs 1 and 2 (light-grey).

Additionally, some subcortical structures have been associated with multisensory processing, such as the superior colliculus (Jiang et al. 2001), hippocampus (Ravassard et al. 2013), and amygdala (De Gelder & Bertelson 2003). Widespread multisensory processing in the brain is responsible for the well-documented interaction of the senses during perceptual (Kawachi et al. 2014; Vidal & Barrès 2014) and cognitive tasks (Lawson et al. 2014; Pasqualotto et al. 2013a). Therefore, the hypothesis that largely overlapping brain networks underlie cognitive and perceptive functions is strongly supported by the findings that "overlapping" brain areas process input from different modalities.

Decisive evidence for the hypothesis that overlapping networks are the basis of brain functions comes from the surprising findings that multisensory processing also occurs in areas that were considered strictly unisensory, such as the primary visual (Borra & Rockland 2011; Zangaladze et al. 1999) and primary auditory (Lakatos et al. 2007; Murray et al. 2005) cortices. Until recently, the notion that primary sensory cortices were concerned with processing input only from the corresponding sensory modalities was one of the core assumptions in neuroscience. However, this theory was challenged when studies with visually impaired participants surprisingly showed that the visual areas of these individuals were not "silent," but they were active while visually impaired individuals performed a variety of tasks (Amedi et al. 2003; Poirier et al. 2006; Sadato et al. 1996). The activation of "visual" areas during nonvisual tasks performed by visually impaired individuals is a clear example of neural reuse, where a neural substrate deprived of its typical input is reused to process input from another modality (Guerreiro et al. 2015; Iachini et al. 2014; Pasqualotto et al. 2013b; Saenz et al. 2008). Additionally, sensory substitution, principally consisting of devices that convert visual information into its auditory or tactile equivalent (Bach-Y-Rita et al. 1969; Proulx et al. 2014), provides theoretical and practical insights into the ability of the brain (including the primary sensory cortices) to respond to environmental pressures (in this case sensory loss) by altering its functions.

Findings of studies with visually impaired participants demonstrate that even the primary visual cortex is involved in multisensory processing; but is this an effect of blindness? Studies wherein sighted adult participants underwent blindfolding showed that this is not the case. In fact, it is reported that participants who had been blindfolded for a few days subsequently exhibited activation of the primary visual cortex during performance of tactile tasks (Pascual-Leone & Hamilton 2001), and that this activation was necessary for successful completion of those tactile tasks (Kauffman et al. 2000). Such a rapid effect of sensory deprivation on the brain function is incompatible with establishment of new brain connections and therefore suggests that multisensory processing "naturally" occurs in the primary sensory cortices (see also Hagen et al. 2002; Kayser et al. 2008; Sathian & Zangaladze 2002). Multisensory processing across distant parts of the brain is supported by preexisting brain connections that recent tracking techniques have started to uncover (Beer et al. 2011; 2013; Kim et al. 2006). In sum, mounting evidence indicates that distributed and overlapping neural networks encompassing both multisensory and "unisensory" (primary sensory) areas are underlying the brain functions, hence substantiating the idea that neural reuse is a ubiquitous phenomenon.

If highly overlapping parts of the brain are responsible for processing much of the information, how can we consciously undergo different experiences such as perceiving the smell of coffee or remembering the events of the past weekend? The answer is that highly overlapping parts of the brain are activated in a particular manner according to the content they process (Burgess et al. 2001; Horwitz & Braun 2004; Rolls & Tovee 1995). Specific patterns of activation amongst overlapping neural population determine the types of neural processing and, ultimately, of the "mind content" (Shinkareva et al. 2008). To recall the initial metaphor, different performance by each of the singers composing the same choir is responsible for the execution of a potentially unlimited number of songs; some up-beat, some down-beat, of different genres and styles.

Scientific intuitions about the mind are wrong, misled by consciousness

Leonid Perlovsky

Athinoula A. Martinos Center for Biomedical Imaging, Harvard University, Charlestown, MA 02129.
lperl@rcn.com
http://www.leonid-perlovsky.com/

Abstract: Logic is a fundamental reason why computational accounts of the mind have failed. Combinatorial complexity preventing computational accounts is equivalent to the Gödelian incompleteness of logic. The mind is not logical, but only logical states and processes in the mind are accessible to subjective consciousness. For this reason, intuitions of psychologists, cognitive scientists, and mathematicians modeling the mind are biased toward logic. This is also true about the changes proposed in *After Phrenology* (Anderson 2014).

In this commentary, I discuss the general fundamental reason evolutionary psychology and other attempts to design computational accounts of the mind have not been successful, and suggest that architectural changes considered in *After Phrenology* (Anderson 2014) would likely be insufficient for computational modeling of the mind. Anderson appropriately challenges some of the accepted wisdom, but the challenges, as well as the suggested modifications, are insufficient. The three suggested principles – "mixing and matching" neural elements, "*procedural* and *behavioral* reuse" (emphasis in original), and that "not every cognitive achievement … need be supported by a specific targeted adaptation (sect. 1, para. 3) – are useful for overcoming some misconceptions about the working of the mind, but they are not sufficient.

Recent cognitive-mathematical results have connected the Gödelian breakthrough in logic incompleteness (Gödel 2001) to difficulties of the computational modeling of the mind. Specifically, whereas Gödel had demonstrated the insufficiency of logic in the 1930s, to this very day logical models of the mind and artificial intelligence dominate computational accounts of the mind and psychological explanations. Logical bias affects even mathematical approaches specifically designed for overcoming logical difficulties, such as "neural networks" and "fuzzy logic" (Perlovsky 1998; 2001; 2006). Whereas every computational paradigm has its own set of mathematical "reasons" for the failings of the resulting algorithms (Perlovsky 2001), a more general reason is the reliance on logic at some algorithmic steps. For example, all learning algorithms have to use an operation that is mathematically equivalent to "this is food," which is a logical statement. The Gödelian incompleteness of logic was shown to be equivalent to the combinatorial complexity of algorithms, which has plagued artificial intelligence and neural networks since the 1960s (Perlovsky 2001; Perlovsky et al. 2011).

For a very long time, logic was considered the best way of deducing scientific truths. But this is a misconception. Aristotle taught his students not to use logic for explaining the mind (Barnes 1984; Perlovsky 2007b). According to Aristotle, the fundamental operation of the mind is a process, in which illogical forms as potentialities meet matter and become logical forms as actualities. Today we describe this process as the interaction of top-down with bottom-up signals. A mathematical description of the Aristotelian process in which vague-fuzzy representations interact with bottom-up signals and turn into crisp representations is given by dynamic logic, or DL (Perlovsky 2006; Vityaev et al. 2013). The experimental confirmation of this process to be an adequate model of perception in the brain is given in Bar and colleagues (2006). The publication also demonstrated that illogical states and processes in the brain-mind are not accessible to consciousness. This point explains why thinking is biased toward logic: Subjective consciousness accesses only logical states in the brain. This selective access to logical states has fundamental consequences for cognitive science, psychology, and artificial intelligence: Intuitions of cognitive scientists, psychologists, and

mathematicians modeling the mind are mostly about logical states and processes in the mind. Correspondingly, cognitive models mostly describe logical mechanisms.

The challenges to widely accepted views about workings of the mind in *After Phrenology* assail some logical foundations of the accepted models (among which modularity is a most obvious consequence of the logical bias) and hence are highly desirable. But these challenges and the proposed modifications are not sufficient: Logical bias is too strongly ingrained in our thinking, and overcoming it would require conscious examinations of the unconscious fundamentals of existing models, analyses of experimental data, and approaches to designing experiments. It is likely that mathematics can help in these analyses and might be highly desirable. In this commentary, I suggest several reasons for the insufficiency of the current ideas and briefly discuss future directions.

Let me repeat: The suggested principles are useful for understanding the working of the mind, but they are not sufficient. For example, mixing and matching the same neural elements in new ways mathematically is similar to the problem of tracking multiple targets with multiple models. In this problem models have to be mixed and matched to signals; the problem is unsolvable; it faces Gödelian complexity except if using DL – the Aristotelian process "from vague to crisp" (Perlovsky et al. 2011); but using of the DL-Aristotelian process requires a new type of intuition about mind mechanisms. It is for the same reason that Hebbian adaptation faces Gödelian complexity: An algorithm would have to augment Hebbian adaptation by searching for synapses belonging to specific "sub-networks." Similar problems would be faced by "neural search" (sect. 2.4) as well as by mechanisms to perceive "external symbols" (p. 190) as different from just objects. The logic-based system, ACT-R, cannot serve as a computable cognitive model even with modifications.

I would like to mention a few topics that are missing from *After Phrenology* that are closely related to logical complexity and are important to a computable theory of the mind (Perlovsky 2006; 2007a; 2007b; 2009; 2010a; 2010b; 2013a; 2013b; 2013c; 2014; 2015):

1. The role of emotions as fundamental to cognition;
2. Actions, specifically human actions, mostly occur inside the brain-mind;
3. Differences between symbols and signs; because *symbol* is a loaded word in psychology, computational cognition must differentiate between signs as conventional marks and symbols as cognitive processes connecting concepts, emotions, conscious, and unconscious (like "simulators," Barsalou 1999). It was the mixing of symbols and signs that destroyed the attempt to construct "symbolic AI" in the 1960s;
4. The difference between higher cognition and perception;
5. Interaction of language and cognition;
6. Grammar is important for language emotionality; it is not an "inessential" system of logical rules in a textbook;
7. Cultural affordance is a beautiful idea, but it is inexorably logical and therefore would likely be incomputable.

I appreciate *After Phrenology*'s emphasis that the componential computational theory of mind often misleads psychological intuition; yet, adding environmental and bodily structures will not be sufficient to overcome its deficiencies.

Beyond disjoint brain networks: Overlapping networks for cognition and emotion

Luiz Pessoa

Department of Psychology, University of Maryland, College Park, MD 20742.
pessoa@umd.edu
http://www.emotioncognition.org

Abstract: It is taken for granted that the brain is composed of a set of networks. But what is a brain network? How should we characterize them? Adding to Anderson's (2014) persuasive anti-modular framework, I propose that (1) networks do not have a single, unique function; (2) networks are highly overlapping and not disjoint; and (3) tasks reconfigure networks found at rest.

In *After Phrenology*, Anderson (2014) argues persuasively against modular frameworks of brain function. The book has much to offer on how distributed brain networks provide a much richer framework to clarify the mind-brain. Here, I would like to provide additional thoughts concerning the organization of brain networks.

Brain networks do not have a single, unique function. Human neuroimaging work, with its clean activation maps showing a few isolated peaks, has been an important contributor to a quasi-phrenological view of the brain. It certainly does not help that many review papers, when summarizing varied results, often include dreaded arrows indicating where some function putatively resides – say, self-referential processing in the medial prefrontal cortex.

Nevertheless, with the explosion of network-related developments in the sciences in general, neuroimaging has finally shifted away from the region-centric approach to one that embraces networks. Hence, network analysis of human neuroimaging data has contributed to a view of brain function that focuses on how groups of brain regions participate in mental functions, and less on how particular regions operate in isolation. *But what are networks?* How should we understand and study them?

Functional MRI data during the so-called resting-state have been extensively investigated in order to characterize *network structure*. A central finding is that, at rest, brain regions can be grouped into a relatively small number of stable communities, also called clusters (or simply networks). For example, Yeo and colleagues (2011) described a seven-community parcellation of cortical areas that was based on a very large sample of participants. Based on anatomical and functional considerations, the communities were given labels such as "visual," "frontoparietal," "default," and so on.

The large body of work employing modern network methods to study brain community structure and other network measures makes a fundamental assumption that appears reasonable at first: The set of nodes (i.e., brain regions) should be partitioned into a *disjoint* set of clusters, such that each node belongs to one and exactly one community. Unfortunately, this assumption, which is adopted frequently in network science (Newman 2010) more generally, too, is problematic. Why should brain regions belong to only a single network?

The first reason is methodological. The ready availability of *disjoint* community-detection algorithms has certainly been a driving force behind this trend. But the methodological aspect as the sole explanation misses a key conceptual point.

I contend that the choice of using disjoint communities to understand brain networks is, in part, linked to the idea that brain regions perform specific functions; that is, the structure-function mapping is, more or less, one-to-one (Pessoa 2014). In this view, a given brain region, which has a specific function, belongs to a single network. Importantly, a network is viewed as having a single, though more general, function, too. For example, Menon and Uddin (2010) have described a "salience network," whose nodes include the anterior insula and anterior cingulate cortex. The "salience network" is suggested to detect salient events and initiate switches between networks involved in self-related, internally oriented processing and those involved in goal-directed, externally oriented processing. But how can the idea that networks implement a single, coherent function be validated beyond that of a suggestion? Bressler and Menon (2010, p. 285) admit that "to determine whether this network indeed specifically performs this function will require testing and validation of a sequence of putative network mechanisms."

But there are good reasons to believe that the mapping between networks and functions is not a simple one-to-one relationship. I suggest that the attempt to map structure to function

in a one-to-one manner in terms of networks will be fraught, as is the one based on brain regions – the problem is simply passed along to a higher level. Hence, two distinct networks may generate similar behavioral profiles (many-to-one); and a given network will also participate in several functions (one-to-many). Broadly speaking, a network's operation will depend on several more global variables – namely, an extended context that includes the state of several neurotransmitter systems, arousal, slow wave potentials, and so forth. In other words, a network that is solely defined as a collection of regions is insufficient to eliminate the one-to-many problem. What if we extend the concept of a network with these additional variables? For example, Cacioppo and Tassinary (1990) suggest that psychological events can be mapped to physiological ones in a more regular manner by considering a spatiotemporal pattern of physiological events. The notion of a network is hence extended to incorporate other physiological events – for example, the state of a given neurotransmitter. How extensive does this state need to be? Clearly, the usefulness of this strategy in reducing the difficulties entailed by many-to-many mappings will depend on how broad the context must be.

In sum, an implicit motivation for the idea of disjoint networks – that networks have stable, unique functions – does not stand scrutiny.

Overlapping brain networks. Let us return to the issue of disjoint versus overlapping networks more specifically. Based on studies of perception, cognition, emotion, and motivation, I have proposed that brain networks are highly *interdigitated* (Pessoa 2013; 2014). At the broadest level, cognition and emotion are not instantiated by separate brain regions; regions important for cognition participate in many emotional processes and vice versa. In a nutshell, cognition and emotion rely on interdigitated networks.

Consider this idea from the point of view of a single brain region, such as the amygdala. Even a rather simplified view of its anatomical connectivity shows that, minimally, it belongs to three networks. The first is a "visual network"; the amygdala receives fibers from anterior parts of the temporal cortex and influences visual processing via a set of projections that reach most of the ventral occipitotemporal cortex. The second is the well-known "autonomic network," and via connections with the hypothalamus and periaqueductal gray (among many others), the amygdala participates in the coordination of many complex autonomic mechanisms. The third is a "value network," as evidenced by its connectivity with orbitofrontal cortex and medial PFC. Hence, the amygdala affiliates with different sets of regions ("networks") in a highly flexible and context-dependent manner.

I propose that brain networks should be considered as highly overlapping; the example of the amygdala being a simple case of three networks overlapping at this structure. In having overlapping networks, the brain is not special. For example, in early work on overlapping community structure, Palla and colleagues (2005, p. 814) suggest that, in all likelihood, "actual networks are made of the interwoven sets of overlapping communities." More generally, the importance of understanding and characterizing overlapping

structure has been discussed by sociologists, as well as biologists, for some time. For example, in a large-scale analysis of the yeast proteome, Gavin and colleagues (2002) showed that a considerable proportion of the proteins studied belonged to multiple networks.

Overall, very little is known about the overlapping community structure of brain networks (but see Mesulam 1998; Yeo et al. 2014) (Fig. 1). But, as in the case of proteins, a large fraction of brain regions may belong to several neural circuits *simultaneously* (see Cocchi et al. 2013; Cole et al. 2013; Hilgetag et al. 1996; Pessoa 2014). It is hence likely that the focus on disjoint clusters has precluded the discovery of important structure in large-scale brain networks (Fig. 1C).

The "flexible hub theory" by Cole and colleagues (2013) supports the idea that brain networks are interdigitated. The framework predicts that "some brain regions flexibly shift their functional connectivity patterns with multiple brain networks across a wide variety of tasks" (Cole et al. 2013, p. 1). The "dynamic cooperation and competition" framework by Cocchi and colleagues (2013) also speaks to the issues at hand. They argue against the idea of segregated systems supporting cognitive control and suggest, instead, that complex control functions are supported by anatomically distributed brain networks that share information in a dynamic manner.

Taken together, time has come to study the overlapping structure of brain networks.

Tasks reconfigure brain networks found during rest. Whereas the large-scale structure of brain networks has been studied extensively at rest, less is known about the large-scale structure during task performance. The central question in this regard is the following: Is the structure seen at rest the same observed during tasks? In particular, are the seven or so networks observed at rest (labeled "visual," "somatosensory," "default," etc.) also present during tasks? Or Buckner and colleagues (2013) ask: Do networks studied during the resting state capture fundamental units of organization or should "rest" be considered just another arbitrary task state? Some have argued strongly that functional connectivity (and hence, network structure) at rest is affected in minor ways by tasks (Cole et al. 2014). In this view, the activity covariation at rest forms a "backbone" that is only mildly influenced by task execution. An alternative proposal is that tasks alter patterns of functional connectivity more substantially (e.g., Buckner et al. 2013).

The claim that networks are largely the same during rest and tasks brings us back to the idea that networks are relatively fixed units of brain function. I suggest that considerable *reorganization* is observed during specific tasks, and that it is therefore better to consider "rest" as a particular task state. Functional MRI is a particularly suitable technique to investigate activity covariation among brain parts that are broadly distributed in space; in other words, functional connectivity patterns which, by definition, are independent of direct (and strong) anatomical connections. And, as revealed in the literature, functional interactions can be strong even when strong anatomical connectivity is absent.
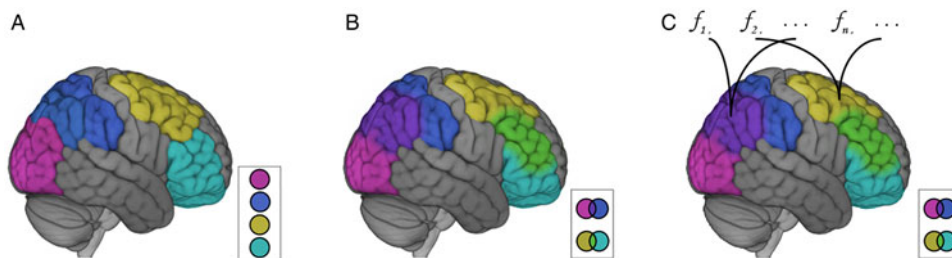

Figure 1 (Pessoa). Brain networks. (A) Standard networks are disjoint (inset: colors indicate separate communities). (B) Overlapping networks are interdigitated, such that brain regions belong to multiple communities (inset: community overlap indicated by mixed colors). (C) Networks (same as in panel B) do not have single, unique functions, but instead participate in multiple functions, f.

We know that the functional connectivity between two regions can increase or decrease as a function of several variables, including task performance (Rissman et al. 2004), motivation (Padmala & Pessoa 2011), and emotion (Pessoa et al. 2002). In one study, the functional connectivity pattern between early visual areas was investigated during affective and neutral contexts (Damaraju et al. 2009). During the affective context, participants viewed faces that were surrounded by a ring whose color signaled the possibility of mild shock. During the neutral context, faces appeared surrounded by a ring whose color signaled safety. A measure of functional connectivity was strengthened during the affective relative to a neutral context. Hence, the affective context not only changed the magnitude of evoked responses, but also altered the *pattern* of responses across early visual cortex (and beyond).

The dual competition model (Pessoa 2009) proposes that the effects of reward during perception and cognition depend in part on interactions between valuation regions and frontoparietal regions important for attention and executive control. Such interactions lead to the up-regulation of control and improve behavioral performance during challenging task conditions (and higher likelihood of reward) (Padmala & Pessoa 2011). Notably, in one study, comparison of the pattern of connectivity between reward and no-reward contexts revealed increases during the former (Kinnison et al. 2012). The increases were observed mostly between two communities of brain regions, reflecting increased *integration* with reward. In particular, the caudate and the nucleus accumbens showed increases in functional connectivity to nearly all cortical regions that were driven by reward.

Large-scale changes in functional connectivity were also found during an emotional manipulation in which a cue indicating the possibility of a mild shock was shown prior to a response-conflict task (Choi et al. 2012). In this case, we observed enhanced functional integration between subcortical regions (such as the bed nucleus of the stria terminalis and thalamus) and cortical regions (including the insula and medial PFC) (Kinnison et al. 2012). Whereas in the case of reward, functional connectivity increased within cortex, in the case of threat, functional connectivity *decreased* within cortex for several pairs of regions.

Together, our findings revealed several ways in which both emotional and motivational processing altered functional connectivity, including increased global efficiency and reduced decomposability (Kinnison et al. 2012). In all, I therefore suggest that network structure is dynamically reconfigured by task states, as also advanced by Cocchi and colleagues (2013) and Cole and colleagues (2013).

In conclusion, I believe that the ideas described above are in close alignment with those advanced by Anderson (2014) and, hopefully, will help us advance the field of brain research "after phrenology."

Toward mechanistic models of action-oriented and detached cognition

Giovanni Pezzulo

Institute of Cognitive Sciences and Technologies, National Research Council, 00185 Rome, Italy.
giovanni.pezzulo@istc.cnr.it
http://www.istc.cnr.it/people/giovanni-pezzulo

Abstract: To be successful, the research agenda for a novel control view of cognition should foresee more detailed, computationally specified process models of cognitive operations including higher cognition. These models should cover all domains of cognition, including those cognitive abilities that can be characterized as online interactive loops and detached forms of cognition that depend on internally generated neuronal processing.

There is increasing consensus in cognitive science and neuroscience that we need a novel view of the brain that is more action-oriented; in this perspective, it has been argued that we might be facing a "pragmatic turn" in cognitive science (Engel et al. 2013). One central proposal of action-oriented views is that the brain evolved to be a control system for the interaction with the external environment. We should look at cognitive (and brain) functions, including the most advanced (or "higher-cognitive") functions, within an interacted and control-theoretic framework – as activities that an organism performs in interaction with its environment, rather than in terms of modular computational operations over discrete symbols independent from perception and action systems. This idea can be traced back to early theories in cybernetics, pragmatism, and ecological psychology (Ashby 1952; Craik 1943; Gibson 1977; Wiener 1948) and has been often reposed, in slightly different forms, in disciplines such as cognitive science, neuroscience, robotics, and philosophy (Cisek 1999; Cisek & Kalaska 2010; Clark 1998; Engel et al. 2013; Pezzulo 2011; Pezzulo & Castelfranchi 2009; Pezzulo et al. 2015; Pfeifer & Scheier 1999; Scott 2012; Stoianov et al. 2016; Varela et al. 1992). Anderson (2014) contributes to this debate both theoretically and empirically. He proposes a view of brain organization in which regions form functional coalitions dynamically and implement control loops for agent–environment interactions. He also discusses how neuronal coalitions could be reused to implement higher cognitive skills, forming possible control loops to support mathematical and language processing.

Despite its appeal, the impact of the action-oriented view is mitigated by various factors. One reason is the level of ambition of the novel proposal, which requires a unitary perspective on brain function and behavior. It is easier to study the brain if one assumes that each area or network is the locus of a single (or a small set of) cognitive operation(s) such as perception, memory, language, and so forth, and that each of these functions can be largely studied in isolation. But this perspective is not fully compatible with an action-oriented view, which assumes that the function of the brain *as a whole* is engaging in interactive control loops. Because these loops extend beyond the brain, we should also consider the contribution of the body and the external environment, including other agents, to cognition.

Given this level of ambition, advocating a generic "control metaphor" for cognition that replaces the older metaphors of the brain – the brain as a computer or as a serial transducer from stimuli to behavior – might be insufficient. What might be more effective is the realization of a coherent set of computationally specified process models of action-oriented cognitive processing – from sensorimotor transformations and situated choice to higher cognitive skills such as mathematics, language processing, and problem solving – that can guide empirical research and provide a more compelling and unifying explanation of empirical findings.

Indeed, despite noteworthy empirical demonstrations of the relevance of action-based concepts to understand cognition, including higher cognition (Barsalou 2008; Glenberg 1997; Jeannerod 2006), we still largely lack detailed process models. For example, several studies illustrate the importance of bodily processes and visuomotor strategies for decision making or problem solving (Spivey 2007), but the (relative) lack of mechanistic models hinders a complete understanding of these phenomena. Progress in the field will benefit from a "new alliance" between proponents of embodied and action-oriented views cognition and computational modelers including roboticists (Pezzulo et al. 2011; 2013). The idea that cognitive robotics can shed light on brain and cognition is relatively new, but it should not be too surprising if one considers that robots are well suited to implement the kind of brain-body-environment (and social) interactions that are considered essential for action-based cognition (Verschure et al. 2014). Furthermore, developing complete models (robotic or nonrobotic) of cognitive operations can counteract an excess of specialization in the field – say, the focus on the functioning of just one brain area.

If we consider the theoretical debate within action-based theories, it often revolves around (the need for) internal representation. Various competing proposals include: abandoning the notion of internal representation to see cognition as online control and interaction with the environment (Chemero 2009; Gallagher 2005); externalizing representation (Clark & Chalmers 1998; Kirsh 2010); replacing it with other constructs such as sensorimotor contingencies (O'Regan & Noe 2001) or dispositions (Engel et al. 2013); amending it in a more embodied and action-oriented view, in terms of, for example, perceptual symbols (Barsalou 1999), action-oriented representations, or emulators that support control structures (Clark & Grush 1999; Grush 2004), to mention just a few.

Although important, this debate might be limited if not accompanied by the realization of mechanistic models, which answer specific questions such as: What kind of control system is the brain? What control loops are instantiated between the brain, the body, the environment, and with other agents, during the implementation of tasks of various complexity, such as walking, solving a puzzle, or planning a trip together? How can these and other tasks be mechanistically described in terms of agent–environment interactions, and which variables, if any, are technically "controlled" in these tasks? What is the contribution of different brain areas to each of the control loops required by different tasks, how are their contributions dynamically assembled, and which are their neuronal signatures? Which loops require a continuous engagement with external variables and which ones their internal, endogenous generation? Which aspects of the structure of the agent–environment interactions can be exploited online, which can be offloaded externally, and which, if any, need to be internalized? How can we better simulate, measure, and empirically study these interactive loops? And so on.

Several researchers are already addressing these and other relevant questions. However, the space of possibilities has not been systematically explored, with many interesting constructs from control theory, cybernetics, and related fields that remain untested. These include early proposals such as hierarchical perceptual control (Powers 1973), the test-operate-test-exit architecture (Rosenblueth et al. 1943), and several other constructs from early cybernetics (Ashby 1952; Wiener 1948), as well as more recent developments such as, for example, optimal feedback control, model-based control, and risk-sensitive control, which are often mentioned in theoretical debates but not systematically tested (Shadmehr & Mussa-Ivaldi 2012).

Incorporating these ideas into specific process models could guide progress in psychology and neuroscience and illustrate where exactly an action-based theory differs from traditional ones. One interesting success case is the *affordance competition hypothesis* (Cisek 2006), in which action specification and selection proceed in parallel and compete – on the basis of the various sensory, contextual, affective, and motivational biases they continuously receive – until one action pattern is selected, hence dispensing from a central decision module that processes abstract values prior to action. Another recent example is the *active inference* framework (Friston et al. 2010; 2015). It can be considered a modern development of cybernetic ideas (Seth 2014), although it casts control as a (probabilistic) inference problem (see also Attias 2003; Toussaint 2009). It challenges traditional models of perception and action control in several ways – for example, by natively incorporating top-down processes and by highlighting the importance of predictive processing and error correction over and above stimulus-response associations – and it paves the way to the realization of models of more elaborated cognitive operations, suggesting that they might be based on the same principles (Clark 2013b; Pezzulo 2012).

Developing mechanistic models can help advance our theoretical understanding of action-based cognition, too, because some of the aforementioned models cut across dichotomies (e.g., interactivity vs. internal modeling) that are often debated, including in Anderson (2014). For example, model-based control and active inference offer a dynamical view of online action control and interaction but include internal generative and predictive modeling loops to make those interactions more effective. These systems learn the statistical structure of the environment and of agent–environment interactions, and they perform state estimation. However, structure learning and estimation are functional to effective control rather than having as a final goal a veridical representation of the external environment. These models remain to be fully tested. However, these examples suggest that the development and empirical testing of specific process models might contribute to the theoretical debate on action-based cognition – for example, by clarifying the possible contribution of predictive processing and structure learning to control and cognition.

These arguments are also important to meet another challenge of action-oriented cognition: developing process models of the brain's ability to temporarily *detach* (or *disengage*) from the here-and-now of the overt sensorimotor loop, as in the case of imagining or planning the future, or mental time travel. These and other detached forms of cognition have been traditionally considered difficult to explain from an action-based perspective and especially in terms of online control loops. In many parts of *After Phrenology*, Anderson (2014) suggests that cognition consists in online interaction with the external environment, with no need for neuronal tissue to internalize the structure of such interaction. He also describes aspects of cognition as the manipulation of *external symbols* (Clark & Chalmers 1998) and briefly alludes to the possibility that, if external symbols are unavailable, internal resources such as memory and imagination might be deployed in their stead. This latter possibility, and in general the contribution of internally generated neuronal processing to cognition, would require much more attention to fully understand some aspects of higher cognition as detached cognition.

One domain where internally generated neuronal processing has been studied in great detail is rodent navigation. It has been consistently reported that hippocampal place cells, whose firing is normally associated with the animal's spatial position, can also fire when the animal is outside its standard "place field," especially during periods of rest or sleep, and at decision points. This "out-of-field" neuronal activity cannot be driven by external stimuli (consider the case of the sleeping animal) but needs to be *internally generated* based on intrinsic network dynamics. Still, it unfolds in highly organized forms and is neurophysiologically similar to the activity observed during overt action. One example is *internally generated sequences* of place cell activations observed when the animal rests or sleeps. These sequences can, for example, "replay" spatial trajectories that the animal has experienced (or recombined trajectories), forward or backward, in a time-compressed manner (e.g., during sharp wave ripple complexes, at about 140–200 Hz). These and other forms of internally generated sequences (e.g., preplays) have been associated with various functions such as memory consolidation and planning (Diba & Buzsáki 2007; Dragoi & Tonegawa 2011; Gupta et al. 2010; Pezzulo et al. 2014; Pfeiffer & Foster 2013). Another kind of internally generated neuronal sequence is expressed in the theta rhythm (8–12 Hz), when the animals are engaged in decision tasks, and has been associated with "what-if" loops and the anticipation of the consequences of possible choices in order to select, say, one of two arms of a maze (Johnson & Redish 2007; Wikenheiser & Redish 2015).

These covert phenomena are not unique to the hippocampus but have been reported in several other brain areas and can at least potentially support a variety of detached or covert cognitive operations (Buzsáki et al. 2015; Lisman 2015). Importantly, the covert activity recruits the same (or closely related) neurophysiological mechanisms and neuronal resources (e.g., neural assembly sequences) as those implied in overt action, but through internally generated processing. This raises the possibility that detached forms of cognition such as planning, but also potentially others such as imagination and mental time travel, can be explained within an action-based framework rather than require a distinct

ontology of neural constructs – if we allow control loops to extend beyond online interactions to also cover internally generated neuronal dynamics.

In this vein, the possibility that the same brain networks can operate and realize control loops in two distinct modes – one stimulus-driven and another (detached) based on internally generated dynamics – has been raised by several researchers and is often associated with the functioning of predictive and generative models (Grush 2004; Maye & Engel 2011; Pezzulo & Castelfranchi 2009) and to theories of reuse (Anderson 2014; Dehaene 2005). However, these hypotheses remain to be systematically tested with the help of mechanistic models. If these theories are on the right track, (some) detached forms of cognition might be based on internalized neuronal loops that recapitulate the same dynamics as those learned during overt action (e.g., replays of spatial trajectories in hippocampal ripples) and hence remain fundamentally action-oriented, albeit in an indirect or covert form. The progressive internalization of overt loops, possibly linked to the functioning of internal predictive models, might be one way the neuronal architecture for action control of our ancient evolutionary ancestors progressively developed more sophisticated cognitive abilities in continuity with its more basic sensorimotor skills (e.g., planning and mental time travel abilities on top of the systems that supported spatial navigation, Buzsáki and Moser 2013). This hypothesis would also explain why these apparently disconnected abilities have partially shared neuronal circuits (Schacter et al. 2007; 2012). Clearly, more specific predictions on the similarities and differences between these abilities are required that can be derived from mechanistic models, which could possibly be tested within the neurophysiological framework for neural coalitions and neural reuse discussed in Anderson (2014).

Computational specificity in the human brain

James M. Shine, Ian Eisenberg, and Russell A. Poldrack
School of Psychology, Stanford University, Stanford, CA 94305.
macshine@stanford.edu   ianeisenberg90@gmail.com   poldrack@stanford.edu
https://poldracklab.stanford.edu/

Abstract: Although meta-analytic neuroimaging studies demonstrate a relative lack of specificity in the brain, this evidence may be the result of limits inherent to these types of studies. From this perspective, we review recent findings that suggest that brain function is most appropriately categorized according to the computational capacity of each brain system, rather than the specific task states that elicit its activity.

Introduction. The human brain is capable of coordinating a wide variety of complex adaptive behaviors. When appropriately wired over development, the same collection of cell bodies and synapses can subtly alter their patterns of activity and communication to coordinate a range of vastly different behaviors. In *After Phrenology*, Michael Anderson (2014) presents a unique and elegant thesis that attempts to explain how, over the course of evolutionary and developmental time, the human brain has come to afford such a range of flexible behaviors. Specifically, it is proposed that the mechanisms of natural selection allow existing brain regions to recombine into novel architectures capable of executing new behavioral patterns. Although Anderson's attempt to synthesize a wide variety of literature across multiple fields is commendable, some of the evidence that he uses to defend this "neural reuse" hypothesis may be overstated.

In this commentary, we will provide evidence that suggests potential challenges to the thesis presented in *After Phrenology*. We will first discuss a major result provided in support of the "neural reuse" hypothesis – namely that meta-analytic evidence from

human functional neuroimaging experiments demonstrate a dearth of specificity in the human brain, and suggest that the conclusions of this experiment may reflect limits inherent to the use of meta-analytic methods within cognitive neuroscience, rather than the re-instantiation of existing neural regions for novel ends per se. Second, we will suggest that brain function is most appropriately categorized according to the functional capacity of each brain system, rather than the specific task states that elicit its activity. Indeed, in considering the computational capacities of independent brain regions, we will make the argument that computational specialization is not only abundant in the brain, but also that it would be difficult to imagine a working brain that did not contain such specialization.

The limits of meta-analysis. One of the key lines of evidence presented in support of the "neural reuse" hypothesis comes from meta-analytic studies of functional neuroimaging data, such as BrainMap (Laird et al. 2005) or NeuroSynth (Yarkoni et al. 2011). Using the BrainMap database, the author describes an experiment that aimed to determine the amount of functional diversity associated with each region of the brain (Anderson et al. 2013). Specifically, the experiment was designed to determine whether given brain regions (here defined using a voxel-wise searchlight approach) were associated with a range of different tasks in a number of unique task categories. The results of this experiment suggested that there was almost no evidence of functional specialization in any region of the brain. Indeed, the author goes as far as to conclude that all brain regions are involved in multiple functions and then takes this as evidence for his hypothesis that all "individual neural elements are put to use for multiple cognitive and behavioral ends" (sect. 2, para. 1).

However, the underlying assumption is that the tasks used in these databases measure "cognitive and behavioral ends" that directly map onto neural computation, which they do not. Instead, these tasks measure surrogate patterns of brain activity elicited by particular task demands, which themselves may rely on a multitude of fundamental neural computations. For example, a neuroscientific experiment might be designed to interrogate the neural mechanism of language production using alterations in the BOLD response, as is measured by fMRI. Assuming appropriate analysis and power, the experimenter will most likely discover an increase in the BOLD response associated with the language task in the left lateral frontal lobe, a region commonly associated with language production (Price 2010). However, it would be problematic to compare the resultant statistical map with a similar pattern derived from a different, yet related experiment (e.g., an experiment measuring working memory demands associated with word formation) and, after finding extensive overlap between the two, conclude that this convergence represented the functional diversity of the left lateral frontal lobe. Indeed, it is much more plausible that this region was performing a similar computation in both experiments (e.g., the top-down manipulation of language-related information). Put another way, even if the brain is perfectly modular in the computations it performs, we would expect a mixture at the level of behavior that would mimic flexible recruitment. That is, "absence of evidence is not evidence of absence" when it comes to modularity.

Importantly, few studies in cognitive neuroscience have been designed to accurately describe the taxonomy of neural computations across the brain. Instead, cognitive neuroscientists propose hypotheses at the level of psychological processes and/or behaviors, which may or may not reflect the true dimensions of neural organization. As such, the maps that are created (and hence populate meta-analytic databases) will not necessarily reflect an accurate taxonomy of brain function. In addition, there are likely to be a number of neural computations that are simply not amenable to analysis using fMRI. For example, Dubois and colleagues (2015) recently showed that, during a face identification task, the identity of individual faces could be effectively decoded from posterior regions of ventral occipital cortex using both fMRI and direct neuronal recordings, whereas in more anterior regions the

information was only able to be decoded from the neuronal signals. Together, this suggests that the data within these meta-analytic repositories are necessarily limited and, hence, conclusions deriving from these studies should be tempered with an appropriate level of caution.

Specialization as a computational entity. Given the nature of the computational demands placed on the brain, it is quite possible that a number of functional units are inherently modular. For example, there are cells within the retina that fire rapidly when exposed to a particular pattern of light, such as a thin, dark horizontal line on a white background,that falls within their receptive field. The information from these cells, when appropriately combined and fed forward to thalamic and occipital cortical regions, underpins the very building blocks of visual perception. However, should the pattern of light exposed to this receptive field alter in some small way (e.g., the light hitting the retina may shift its angle of approach by some small degree; or the contrast-to-noise ratio along the outlines of the image might change), the retinal cells will quickly decrease their firing rate. Although there is no doubt that these same retinal cells will be used in the completion of a wide variety of behavioral tasks and for very different ends (e.g., tracking the flight of a ball that one is attempting to catch versus deciphering the presence or absence of a particular letter that may be associated with a cognitive task), the computations that they perform are specialized. That is, these retinal cells display a specialized computational architecture.

In contrast, other brain regions display a much more flexible computational repertoire. For example, there are neurons within the prefrontal cortex that can switch the target of their receptive fields according to whichever cue is currently of motivational value to an animal (Barbas & Zikopoulos 2007). Similarly, experiments using direct cell recordings in non-human primates have shown that prefrontal cortical neurons can change their sensitivity to patterns of information based on shifting goals (Mante et al. 2013) and can also alter the direction of information flow within frontal circuits that control sensorimotor decision making (Siegel et al. 2015). In humans, distinct regions of the prefrontal cortex modulate their functional connectivity as a function of task-set (Sakai & Passingham 2003). Importantly, although these neural regions will presumably be among the most commonly reused regions, the computations that they are performing will likely be highly specialized. That is, the neural systems processing the information from whichever object is the current focus of the animal, be it exogenous (such as counting a series of objects in the world) or endogenous (such as mentally counting the number of papers published by a rival colleague), will be computed upon in a similar way by the neurons within the frontoparietal control networks.

Attributing computational specificity to a particular region is not to suggest that the brain is composed of discrete modules, each performing specialized computations. Rather, subsections of the overall brain network display properties (such as specialized architecture, cell types, or connectivity), which afford computational roles to individual brain regions that emerge through the interaction with the rest of the network. For example, there might exist a subcircuit incorporating the motor cortex, putamen, and motor thalamus that is particularly important for the execution of a specific motor pattern (say, moving the right index finger), but the output of this circuitry would be simply unable to function out of context (i.e. in vitro), as its computational role arises directly from its location in space (i.e., within the brain, which is itself within a larger organism) and time (i.e., the manner in which its been developed over learning). As a field, cognitive neuroscience has been attempting to understand these computations by manipulating behavior through the use of neuropsychological tasks; however, to properly understand the functional role of each subnetwork within the brain, it may be more useful to describe the computational role of each circuit in context. If such an undertaking were performed, we imagine that some regions perform a well-defined computational role regardless of context (e.g.,

retinal cells processing light), whereas others are more context-dependent (like prefrontal cortex), flexibly altering their role based on the activity of the overall brain network.

Refining cognitive ontologies. Many of the issues highlighted in this commentary could be solved through the progressive refinement of a cognitive ontology that more accurately describes the relationships between the many and varied computational capacities instantiated within the brain (Poldrack et al. 2011). Indeed, given our current lack of conceptual clarity as to the appropriate ontological framework in which to interrogate the brain, any attempt to classify behaviors according to our current framework will necessarily be flawed, hence leading to faulty distinctions in the neural architecture of behavior. To proceed, we agree that we should seek to appropriately define the "psychological factors that best capture and account for the differential activity of the brain in various circumstances" (sect. 6, para. 3). However, we differ in our predictions of the likely neurocognitive architectures that will best explain the functional landscape of the brain. Whereas the author would advocate for an architecture in which all "individual neural elements are put to use for multiple cognitive and behavioral ends" (sect. 2, para. 1), we propose that there will likely be a range of fundamental neural computations that, when combined with information regarding the context of the individual neural population within the larger neuronal network (such as the structural connectivity profile or the relative proportion of different neuromodulatory afferents to a region), will effectively explain the functional repertoire of the brain.

That is not to say that these computations will necessarily make a region more "specialized" in a behavioral sense, but rather that the computational abilities of a given region will define its involvement across multiple tasks. For example, regions within the prefrontal cortex that hold information online over time (Curtis 2006; Sakai & Passingham 2003) can be utilized to maintain information during multiple different behavioral tasks, such as those invoking language, social, mathematical and logical reasoning capacities. However, in each case, the computation performed by the region is likely to be similar, whereas the behavioral outcome will differ, depending on the context in which the computation is deployed. Within this framework, regions can also develop over evolutionary or developmental time to become computationally specialized (Bassett et al. 2015) and hence become recruited only under the precise situations that require their involvement (such as Broca's area in response to the processing of language-related information; Price 2010).

Conclusion. Together, our arguments suggest a possible reformulation of the author's main hypothesis, which would place specialization and reuse together along a functional spectrum, in which the computational properties associated with each brain system define its potential involvement in the mechanistic definition of a range of behavioral capacities.

The implications of neural reuse for the future of both cognitive neuroscience and folk psychology

Michael Silberstein

Elizabethtown College and UMD College Park Department of Philosophy, Elizabethtown College, Elizabethtown, PA 17022.
silbermd@etown.edu
http://facultysites.etown.edu/silbermd/

Abstract: If neural reuse is true, then: (1) fully escaping phrenology will eventually require an even less brain-centric and mechanistic cognitive neuroscience that focuses on relations and interactions between brain, body, and environment at many different scales and levels across both

space and time, and (2) although scientific psychology must be heavily revised, the autonomy and irreducibility of folk psychology are assured.

Whenever the discussion turns to neuroscience, I tell people if they want to read the best thing written about the brain in this century, they need to read *After Phrenology* (Anderson 2014). I say this up front to put any differences between Anderson and me in perspective. I view these differences as relatively minor, mostly a matter of emphasis. Herein I will focus on this two-part question: If Anderson is right about: (a) neural reuse, redeployment, and multiuse at multiple spatial and temporal scales in the brain and (b) the implications for extended and embodied cognition and ecological psychology (I believe he is right on both counts), then (1) what ought the future of cognitive neuroscience look like, and (2) what does the future hold for the autonomy of folk psychology? He says he is advocating explicitly for giving the brain its scientific voice and hence the likelihood of a significant revision to the vocabulary of cognition (p. xxii). I will argue that if Anderson is right about the nature of the brain, then he needs to move even further away from a focus on mapping structure to function and even closer to a more radically extended and embodied account of mind and behavior. Second, I will argue that while the vocabulary of scientific psychology – that is, RTM and CTM – must change, Anderson's vision of the brain and embodied cognition is a good argument for the autonomy of folk psychology and hence a good argument against reduction.

I agree completely with Anderson when he says "individual pieces of the brain, from cells to regions to networks, are used and reused in a variety of circumstances, as determined by social, environmental, neurochemical, and genetic contexts" and that "what is reuse at one level of organization can be multiuse at another" (p. 36). He says the brain achieves its functions by assembling the right functional coalitions neural and extra-neural, including body, environment, and external artifacts such as symbolic ones, all in the service of adaptive behavior (pp. 242, 302). Hence, intelligence is not so much about local processing as it is about "cooperative connectivity" (240). This implies not just the failure of massive modularity, but also a complex many-many relationship between structure and function at multiple scales. In Silberstein and Chemero (2013), which focuses on networks and graph theory in systems neuroscience, we note that very different neurochemical mechanisms and wiring diagrams can instantiate the same networks and hence perform the same cognitive functions. Indeed, in these models, it is primarily the topological features of various types of small-world networks that explain essential organizational features of brains, as opposed to lower-level, local purely structural features. As Sporns puts it, "a reentrant system operates less as a hierarchy and more as a heterarchy, where super- and subordinate levels are indistinct, most interactions are circular, and control is decentralized" (2011, p. 193). Hence, topological features such as the properties of small-world networks exhibit a kind of universality with respect to lower-level structural details. Structural and topological processes occur at radically different and hard to relate timescales. The behavior and distribution of various nodes such as local networks are determined by their nonlocal or global connections. That is, such global organizational principles or features of brains are not explicable in principle via localization and decomposition.

Anderson therefore is absolutely right that this picture of the brain tells against componentiality and mechanistic explanation as standardly conceived. As he says, "Global function is not built from componential local function, but rather the reverse" (p. 93) and "local neural assemblies are polymorphic and 'multifunctional'" (p. 104). He says then that we must use current and invent new increasingly sophisticated tools for mapping structure to function, such as measures of effective connectivity between various interacting parts of the brain. He gives the example of Granger causality for exploring effective connectivity between large-scale systems of the brain. He describes the bottom-up approach that creates mappings between network motifs (common

configurations of functional relations between neural elements) and causal/functional effects. To do this properly, he says, we must first learn to identify small-scale network configurations in living animals (p. 307).

I certainly agree with all this as far as it goes, but my worry is that it is still too brain-centric, still too focused on the "where strategy," to use Anderson's language. My worry is based on the following: (1) brains and other complex biological systems exhibit multiple realizability at all scales and levels of description defined both structurally and functionally, and (2) as Anderson stresses, brains are embedded in larger physical and social networks that play a huge role in determining their activity and interactions at all scales. He argues for example that "actual social interaction appears *necessary* for language acquisition" (p. 258). Given 1 and 2, the first main point I want to make is that going beyond the brain, we need a neuroscience that is inherently multi- and inter-disciplinary, one that focuses on various complex relationships at multiple scales and levels of organization such as between gene networks, RNA networks, epigenetics, various levels of scale in the brain from neurons to networks, behavior, cultural and social features, phenomenology, and so forth. These relations need to be studied both synchronically and diachronically and at various timescales including the developmental and evolutionary.

The point is that a neuroscience-in-full is inescapably interlevel and deeply historical – a point Anderson makes as well, to be sure. However, I think the point deserves greater emphasis because the focus on mapping structure to function, however sophisticated, comes from the perspective of intervention and manipulation versus what I will call "seeing the big picture." Of course, there is nothing wrong with intervention and manipulation, especially in the biological sciences, but in the long run, with respect to both control and the big picture, the full potential of systems neuroscience will be realized only if we move further away from such mechanistic thinking. Of course, this violates some conceptions of what biology is and potentially starts to make biology look a lot more like mathematical physics in general or the study of many-bodied systems such as condensed matter theory. However, this does not imply a reduction of biology to physics, just an increasing use of the physics tool kit, which – as illustrated by graph theoretical neuroscience – is of course already happening.

Anderson characterizes *the brain* as a primarily action oriented control system as opposed to computationally driven. *The brain* is an action controller responsible for managing the values of salient organism-environment relationships, that is, affordances (p. 307). So *the brain* controls or manages biased affordance competition. Of course, the brain for Anderson is decidedly not a CPU, but why does he say *the brain* is the controller, *the brain* assembles the right functional coalition, and so forth, and not the extended organism itself. The worry is that given his Gibsonian picture, this way of talking about the brain makes it sound too autonomous from the body, the environment, and the affordances of the organism, too CPU like. He talks about the "multidimensional functional biases of individual brain regions as indices of some set of underlying causal dispositions" (p. 307). He talks about this in terms of neural dispositions, differential propensities, and causal powers (pp. 114, 308), and in terms of "personalities of brain regions." He calls this the *dispositional vector account of brain activity*. My worry is that this sort of language suggests not merely that different networks or regions have a tendency to contribute to certain tasks, but that there is something intrinsic and mechanism-like in virtue of which they so contribute. I am not angling for full-blooded holism, and I do not doubt that there are important differences, but I am suggesting these differences, rather than being intrinsic, are primarily fixed by evolutionary, developmental, functional constraints, and historical contingencies. Furthermore, there may be no optimal brain architecture, and even if we discover such TALoNS (p. 94), we may be learning about said constraints and historical contingencies only with respect to humans or a subset thereof: There is probably multiple

realizability with respect to TALoNS as well. And, as Anderson acknowledges, these constraints often get trumped by cultural and social constraints that also get preserved and conserved across time. So I want to push back a little against the picture of brain regions evolving particular dispositions that get combined in various ways like primary colors to collaborate on different tasks.

Rather than speak of brains as control systems with intrinsic dispositions, I want to say that brains are hubs in the graph theoretic sense, but they are living hubs in multilevel dynamical processes that are extended in space and time. The networks in question are physio-bio-cultural, and brains are truly historical artifacts that bear the marks of their origins and development. In this model, extended organisms or agents spanning brain, body, and environment are the primary locus of control, and the structural elements that instantiate such networks are often secondary to the networks themselves (Silberstein & Chemero 2011).

This brings me directly to my second main point about the autonomy of folk psychology. Rather than reduction or elimination, neural reuse and extended cognition so conceived actually support functionalism in the sense that we individuate processes with regard to their role and effects as opposed to their structural constituents. Again, looked at in this way, we find a great deal of multiple realizability within and across human and non-human brains. Anderson places a lot of emphasis on the importance of selection in understanding brain function (p. 296), and I agree, but I think there are others aspects of evolutionary and developmental biology that are equally important and that do not fully reduce to selection. Evolutionary theory has its own version of functionalism in the form of convergent or parallel evolution such as mimicry and flying (McGhee 2011). Focusing on cognitive convergence, creatures with very different brains and selective histories seemingly converge on similar behavioral and cognitive strategies for dealing with "socioecological" problems. For example, corvids do not even possess a prefrontal cortex, yet they exhibit very intelligent behavioral and cognitive strategies similar to primates. There is growing evidence across the board that creatures with very different brains have in many respects converged on relatively similar minds (McGhee 2011). Well-known examples of this from insects to mammals include: tool use, architectural behavior, agricultural behavior, social or collective behavior, mathematical behavior, and language use (McGhee 2011, Ch. 6). There is also a growing consensus that a wide variety of different species with very different brain structures and nervous systems possess some form of not only sophisticated convergent cognition, but also consciousness, self-awareness, and metacognition. Examples abound, such as the mourning behavior of gorillas and dolphins (McGhee 2011, p. 240). I think the best explanation for cognitive convergence is that the affordances, environmental and social features often trump structural neural constraints whether imposed by physics or selection.

So, although I agree with Anderson that the brain must have its voice, what I think it is telling us is that neural reuse is best viewed as a subset of what developmental and evolutionary biologists call plasticity and robustness (Bateson & Gluckman 2011). Although there are many different kinds of both robustness and plasticity, in general robustness refers to relative stability or invariance across environmental, genetic, or cellular transformations, and plasticity refers to cases wherein features of the organism are held constant such as genotype, and yet because of environmental transformations the organism nonetheless manifests very different or unique adaptive traits or behaviors (Bateson & Gluckman 2011, p. 8). Different forms of plasticity include: phenotypic, molecular, variety of neural types, immunological, and behavioral (Bateson & Gluckman 2011, p. 46). Both of these features of complex biological systems are of course at the heart of the epigenetic revolution in biology. As many people have pointed out, robustness and plasticity are two sides of the same coin: "Plasticity is often regulated by robust mechanisms and robustness is often generated by plastic mechanisms" (Bateson & Gluckman 2011, p. 46). It is important to note that, although many biological mechanisms possess the properties of robustness and plasticity, these features cannot in principle be explained mechanistically in terms of localization and decomposition. These are global/systemic and scale-invariant features of such biological systems. Convergent evolution, robustness, and plasticity all go hand-in-hand, and they all point to the strongly extended nature of phenotype, behavior, and cognition. For one take on what such a Gibsonian cognitive neuroscience might look like, see Silberstein and Chemero (2012).

Modularity in network neuroscience and neural reuse

Matthew L. Stanley[a] and Felipe De Brigard[a,b,c]
[a]*Center for Cognitive Neuroscience, Duke University, Durham, NC 27708;*
[b]*Duke Institute for Brain Sciences, Duke University, Durham, NC 27708;*
[c]*Department of Philosophy, Duke University, Durham, NC 27708.*
mls103@duke.edu   felipe.debrigard@duke.edu
www.felipedebrigard.com

Abstract: Neural reuse allegedly stands in stark contrast against a modular view of the brain. However, the development of unique modularity algorithms in network science has provided the means to identify functionally cooperating, specialized subsystems in a way that remains consistent with the neural reuse view and offers a set of rigorous tools to fully engage in Anderson's (2014) research program.

In *After Phrenology: Neural Reuse and the Interactive Brain*, Anderson (2014) offers a valuable framework for understanding neural (re)organization and its relationship to cognitive functioning. This framework, neural reuse, maintains that most regions of the brain are involved in multiple and diverse cognitive tasks, and that the brain is capable of achieving functional flexibility precisely because it can redeploy the same neural structures for a variety of purposes (p. 5). This neurofunctional architecture allegedly stands in stark contrast against a modular view of the brain. However, in our view, recently developed, data-driven modularity algorithms in network science (modularity$_{net}$), which make use of mathematical formalisms from graph theory, remain consistent with Anderson's view while providing a set of rigorous methods to engage in Anderson's research program.

Modularity$_{net}$ algorithms are computed on networks, which are formally described using graph theory (Newman 2006). All networks are composed of differentiable elements of the system (nodes) and pairwise relationships between those elements (edges). In the context of human functional brain networks, each node represents a discrete parcellation of brain tissue, and edges represent measured functional interactions (connectivity) between pairs of nodes (Bullmore & Sporns 2009; Rubinov & Sporns 2010). Modularity$_{net}$ algorithms can be used to identify nodes that form tightly interconnected subgroups of interacting brain regions functionally cooperating to subserve certain tasks. These modularity$_{net}$ algorithms provide valuable information about (1) the extent to which the entire system can be (partially) decomposed into modules$_{net}$, (2) the extent to which the nodes within each module$_{net}$ are preferentially cooperating with nodes within their own module$_{net}$ versus nodes in other modules$_{net}$ in the system, and (3) the different functional roles of specific nodes within each module$_{net}$ (Stanley et al. 2014; Telesford et al. 2011). As such, modularity$_{net}$ constitutes an alternative to dimensionality reduction approaches discussed by Anderson (e.g., diversity variability, Dice's coefficient), while providing more information about the properties of these functionally cooperating groups of regions subserving different tasks. Although modules$_{net}$ are not isolated, autonomous, encapsulated processing units – as demanded by more traditional formulations of "modularity" – they do still perform particular, specialized functions during
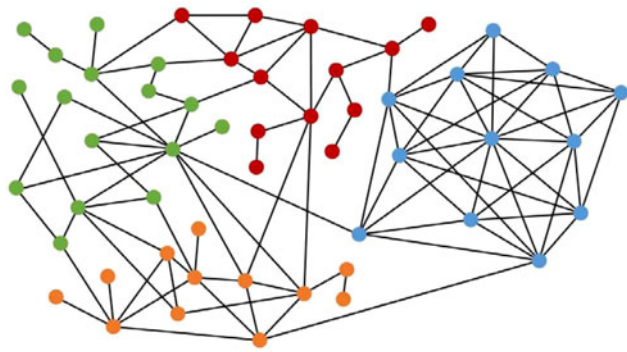
Figure 1 (Stanley & De Brigard). Provides an illustration of possible properties of a network's modular$_{net}$ organization. Suppose each individual node (circle) represents some discrete, predefined portion of the brain, and the links between nodes represent functional interactions between nodes. Node color represents the module$_{net}$ to which that node belongs. Notice that the nodes composing the blue module$_{net}$ are very densely interconnected among themselves, suggesting that those nodes are cooperating to achieve some function, despite remaining interconnected to all other nodes in the system via direct and/or indirect connections. In contrast, the nodes composing the red, orange, and green modules$_{net}$ are not nearly as densely interconnected or clearly defined. The kind of modular structure exhibited by the group of blue nodes allows researchers to maintain the position that modules exist and are responsible for specialized, segregated functions while remaining consistent with Anderson's overarching framework that militates against purely decomposable, strictly domain-specific, encapsulated, and isolated modules.

certain tasks via the interactions between brain regions within each module$_{net}$ itself. Specific modules$_{net}$ are thought to perform specific functions, even though each module$_{net}$ remains interconnected, or integrated, with all other modules$_{net}$ in the system by a complex set of inter-module$_{net}$ connections (Fig. 1). Thus understood, modularity$_{net}$ safeguards a conception of specialized, segregated functioning, which is central to more traditional views of modularity. Indeed, it has been shown that a modular$_{net}$ architecture both exists in functional brain networks and is advantageous because it increases the robustness, flexibility, and stability of the system (Barabasi & Oltvai 2004; Valencia et al. 2009).

In what follows, we briefly identify the fundamental principles of neural reuse and show that modularity$_{net}$ is consistent with it. We also suggest that modularity$_{net}$ provides the best set of methods for investigating neural reuse, while preserving attractive features traditionally ascribed to a modular view of the brain.

The *first* claim Anderson makes about neural reuse provides the foundation for his framework: *Brain regions should be used and reused for diverse purposes in various task domains* (p. 9). In using modularity$_{net}$ algorithms to investigate neural reuse, we assume that (1) subgroups of nodes (brain regions) identified as highly interconnected during a task consistently across persons are being used for that task, and (2) nodes can change in module$_{net}$ allegiance between any two conditions. If, in fact, nodes change in their functional properties to be part of different interconnected, functionally cooperating groupings (modules$_{net}$) from task A to task B, then the modularity$_{net}$ algorithm will be able to identify this shift. Because modularity$_{net}$ algorithms are data-driven, the quantity of nodes forming any given module$_{net}$, the spatial locations of nodes within any given module$_{net}$, the consistency of modular$_{net}$ organization across a set of persons, and the extent of interconnectedness of nodes comprising each module$_{net}$ are determined by the very nature of the functional brain network,

and not by the experimenter. Furthermore, modularity$_{net}$ algorithms are designed to admit the possibility that there are no clear, distinct modules$_{net}$ in the network. Consequently, modularity$_{net}$ serves as both a validation of neural reuse between different tasks and as a way to identify those functionally cooperating groups of brain regions subserving any task of interest. Modularity$_{net}$ makes no a priori assumptions about the selectivity of local neural structures for specific task(s). If, in fact, local neural structures are not highly selective and typically contribute to multiple tasks across domain boundaries, then modularity$_{net}$ will show how neural structures (nodes) change modular allegiance by reorganizing their connections for any set of tasks.

Recent work has demonstrated that human functional brain networks exhibit a modular$_{net}$ architecture, but that architecture is neither temporally static nor anatomically fixed. Instead, nodes alter their interactions with other nodes in the network to form new modules$_{net}$ depending upon the demands on the system (Bassett et al. 2011; Cao et al. 2014; Meunier et al. 2014; Moussa et al. 2011; 2014; Stanley et al. 2014). For example, Stanley et al. (2014) demonstrated that the nodes within the module$_{net}$ mostly composed of default-mode brain regions maintained a highly consistent, densely interconnected, functionally cooperating module$_{net}$ during a minimally demanding working memory task. However, when more attentional and working memory resources were required, the previously observed default-mode module$_{net}$ dissolved. During this more difficult working memory task, a different set of brain regions, many of which have traditionally been considered to be involved in working memory processes, became densely interconnected, forming a new, consistent module$_{net}$ across persons. Changes in modular$_{net}$ organization within individual persons have been shown to facilitate behavioral adaptation during simple tasks, further emphasizing the constantly changing, dynamic nature of modules$_{net}$ (Bassett et al. 2011).

The *second* claim Anderson makes about neural reuse concerns the importance of interactions between different neural elements: *Functional differences between task domains are critically reflected in the different patterns of interaction between many of the same elements forming the right neural team for a job* (pp. 9, 46). Modularity$_{net}$ is computed by detecting certain patterns in the interactions between all brain regions comprising the network simultaneously. That is, modularity$_{net}$ takes into account all interactions between each and every brain region and subsequently identifies the subsets of brain regions that tend to be more densely interconnected (i.e., interacting more strongly) among themselves than the rest of the network. The functional groupings of highly interconnected brain regions observed during task A that appear consistently across subjects are thought to actually subserve the cognitive process(es) associated with task A. Although modularity$_{net}$ does still provide relevant information about the spatial locations of nodes comprising each module$_{net}$ in the brain, modularity$_{net}$ provides more important information regarding how certain regions are functionally cooperating during any given task, the extent to which regions are densely interconnected, how that interconnectedness changes across tasks, and the relative importance of specific nodes within modules$_{net}$ for facilitating integrated and segregated neural functioning.

Critically, however, if one grants that achieving a task is really about putting together the right "neural team" (i.e., module$_{net}$), then that module$_{net}$ is engaged in specialized processing for that specific task. Different components of that neural team might be serving different subfunctions, but the components of the team are all working to serve the particular function carried out by the module$_{net}$ itself. So, in investigating the neural basis of cognitive processes, one cannot merely be concerned with integration *between* modules$_{net}$, as Anderson suggests (p. 42). Modularity$_{net}$ algorithms recognize the importance and mutually dependent necessity of both segregation (specialization) within a particular module$_{net}$ and integration between modules$_{net}$. Understood in this way, segregation and integration of function are not wholly

separable. But, nodes in modules$_{net}$ still densely interconnect to form functionally cooperating groups engaging in specialized functioning during specific tasks. By recognizing the importance and mutual necessity of both segregated and integrated functioning, modularity$_{net}$ seems to preserve the one truly valuable tenet of more traditional views of modularity – that is, that there is, in fact, segregated, specialized functioning in the brain that is necessary for cognitive functioning.

The *third* claim Anderson makes about neural reuse implicates evolutionary assumptions: *Later emerging behaviors/abilities should be supported by more disparate neural structures* (p. 9). Modularity$_{net}$ makes no a priori assumptions about which behaviors/abilities are supported by the most scattered set of structures in the brain. Importantly, modules$_{net}$ need not be spatially contiguous in the brain, because edges in functional brain networks are defined as statistical dependencies in neural signal between nodes. Because the size, consistency, and spatial scatter of modules$_{net}$ are determined by the nature of the functional network itself, modularity$_{net}$ actually provides the ideal way to empirically test whether later emerging behaviors/abilities are supported by a highly interconnected, consistent module$_{net}$ comprising a greater proportion of different structures broadly scattered throughout the brain. So, if Anderson's third claim is right, then the data-driven modularity$_{net}$ algorithm should identify the functional groupings of nodes representing the latest developing modules$_{net}$ (during the relevant tasks) as those that exhibit the most noncontiguous and densely interconnected, yet spatially consistent (across persons), spatial scatter throughout the brain.

The *fourth* claim Anderson makes maintains that neural reuse is a guiding functional principle across many different spatial scales: *Neural reuse does not go away, no matter how small the brain region* (pp. 29–30). By using modularity$_{net}$ algorithms, it is possible to investigate neural reuse on many different spatial scales. Even among the existing network analyses of fMRI data alone, researchers have reported results from 70-node to 140,000-node whole brain networks (Stanley et al. 2013). Assuming adequate computational power, modularity$_{net}$ can be computed on networks at any of these spatial scales, and different kinds of information can be extracted from the modular$_{net}$ architecture (or lack thereof) observed at each scale. Recent work has also led to the development of hierarchical modularity$_{net}$ algorithms (Arenas et al. 2008; Meunier et al. 2009; Sales-Pardo et al. 2007), where each module$_{net}$ obtained at the partition of the highest level can further be decomposed into submodules$_{net}$, which in turn can be decomposed into subsubmodules$_{net}$, and so on. This makes it possible to rigorously compare modular$_{net}$ organization at many different spatial scales in order to capture reuse at different levels of organization within the same brain during the same task.

The *fifth* critical claim Anderson makes about neural reuse concerns separate modifiability and decomposability: *Neural reuse does indeed militate against separate modifiability; the brain is not a nearly decomposable system consisting of separately modifiable parts* (pp. 39, 40). To understand the compatibility between Anderson's fifth claim and modularity$_{net}$, it is necessary to distinguish between strong and weak versions of decomposability. Decomposability$_{strong}$ refers to a fully separable system in which any element can be removed or altered without significantly impacting the remaining elements of the system (e.g., a massive modularity view). Decomposability$_{weak}$ refers to a system that has parts that tend to group together to perform specific functions, but in which each part remains connected to the rest of the system through a complex pattern of interactions, such that no functional group can be changed without changing the system as a whole. Although more traditional formulations of modularity operate under the assumption of decomposability$_{strong}$ that further assumes separate modifiability, modularity$_{net}$ only assumes decomposability$_{weak}$, which does not allow for separate modifiability. Functional brain networks are investigated as fully interconnected, interdependent, nonlinear systems. This means

that no module$_{net}$ can be modified in practice without also changing the modular$_{net}$ architecture of the system as a whole. By accepting decomposability$_{weak}$ but not decomposability$_{strong}$, modularity$_{net}$ is able to maintain a particular notion of segregated, specialized functioning while still remaining consistent with neural reuse.

Neural reuse holds tremendous promise as a framework with which to understand and investigate the neural bases of cognitive processes. We have argued that recent developments in network neuroscience offer a particular view of modularity – modularity$_{net}$ – that is consistent with the fundamental tenets of neural reuse. Additionally, we have suggested that these modularity$_{net}$ algorithms constitute an excellent tool for empirically investigating neural reuse, even for researchers reluctant to relinquish the idea of specialized modules altogether.

Neural reuse leads to associative connections between concrete (physical) and abstract (social) concepts and motives

Yimeng Wang and John A. Bargh
Department of Psychology, Yale University, New Haven, CT 06511.
yimeng.wang@yale.edu   **john.bargh@yale.edu**
http://www.yale.edu/acmelab/index.html

Abstract: Consistent with neural reuse theory, empirical tests of the related "scaffolding" principle of abstract concept development show that higher-level concepts "reuse" and are built upon fundamental motives such as survival, safety, and consumption. This produces mutual influence between the two levels, with far-ranging impacts from consumer behavior to political attitudes.

The neural reuse theory has proven remarkably generative and supportive of research activity on embodied cognition, motivation, and behavior. We first present and discuss recent research evidence consistent with a "conceptual scaffolding" principle that is in harmony with the neural reuse perspective. We then discuss how conceptual development in neuroscience and empirical research in social psychology may collaborate in fathoming the neural structure of grounded cognition.

Anderson (2014) notes that conceptual metaphors – structure and logical protocols from one domain guiding and structuring those in another (e.g., life is a journey, love is war) – may be only one "side effect" (p. 26) of the neural reuse process on a larger scale. Evidence from social and consumer psychology on embodied forms of judgment and behavior suggests that higher-order mental conception is not built *from* the more basic functional models as "prototypes" (Anderson 2014, p. 17) as much as it is built *upon* (associatively connected to) the biological groundings of the primal functions, sharing their neural substrates by reuse. Several lines of research provide support for this proposition.

The conceptual scaffolding account (Williams et al. 2009), which was influenced by and is in harmony with the neural reuse principle, argues that more abstract concepts and complex, higher-order mental functions (person perception, self-esteem, value judgment) grow organically from and are hence "built upon" the more fundamental, innate needs, such as for survival, safety, and resource acquisition (consumption). The "built upon" notion leads directly to the assumption that associative connections will be formed between the concrete and the abstract concepts or goals, associations that remain intact throughout the lifespan. This in turn leads to hypotheses regarding the use of physical level concepts in describing more abstract social and psychological phenomena, as documented by extant theory and research on metaphor use (Lakoff & Johnson 1980; Landau et al. 2010). Social relationships (a *close* relationship, a *warm* father, a

sweet thing to do) are commonly described in physical terms, and we easily understand each other when these terms are used. But the "scaffolded" associative connections play an even larger role than in language alone, as many studies have now documented how activation of the physical level concepts (such as by holding something warm, or feeling something rough) spreads along these associative connections to activate and influence analogically related psychological and social level judgments and motivations as well.

For example, research in the sensorimotor domain shows that higher-level social judgments are influenced by concrete physical sensory experiences (Ackerman et al. 2010; Krishna 2012). Ackerman et al. (2010) found that incidentally touching something rough or smooth (rough or smooth backing on the clipboard they were using) influenced participants to rate a social interaction they read about as having gone more roughly or smoothly. Schaefer et al. (2014) replicated this effect across 96 different social interaction scenarios, in the context of an fMRI study in which participants touched something rough or smooth prior to reading each of the scenarios. Schaefer et al. (2014) replicated the rough-smooth priming effect but showed further that the extent of activation of the somatosensory cortex caused by the physical experiences correlated significantly with the extremity of the social judgment of the smoothness of the social interactions.

Moreover, just as cold ambient temperatures and physical objects (e.g., ice cubes) cause a person to feel physically colder, so too does "social coldness" in the form of rejection by others cause a person to actually feel and become physically colder (IJzerman et al. 2012). Physical warmth (as from holding something warm) causes people to become more socially warm (generous, trusting: Inagaki & Eisenberger 2013; Williams & Bargh 2008), and holding something cold causes people to act with less generosity and lower levels of trust in behavioral economic games (Kang et al. 2011). And again, neuroscience studies have shown that the degree of activation of the brain regions that respond to the physical level experience are correlated with the extent of the obtained effect in the more abstract social or psychological domain. In the case of physical warmth and coldness, the same regions of insula become active upon physical and social warmth experiences – holding something warm or texting to one's family and friends (Inagaki & Eisenberger 2013) – and upon physical and social coldness experiences (Kang et al. 2011). Moreover, the same homeostatic mechanisms are activated in the two forms of warmth or coldness: Just as one seeks warmth when cold, Zhong and Leonardelli (2008) demonstrated that after an experience of social exclusion (coldness), participants more often wanted a hot drink or warm food for lunch, compared with an iced drink or cold cuts.

That physical level experiences such as a warm bowl of soup can effectively substitute for missing feelings of social warmth (as after exclusion or rejection by others) is reminiscent of Kurt Lewin's early theory of *goal substitution*, as originally studied by his student Mahler (see Wicklund & Gollwitzer 1982). If a desired goal cannot be met, then substitute activities can often, at least temporarily, satisfy (and turn off) that goal. And so socially rejected participants more strongly prefer physically warm beverages or lunch dishes (Zhong & Leonardelli 2008). Other studies show similar goal substitution effects, even at the abstract level of political values and attitudes.

For example, we all have a need to feel in control of our lives and important (goal-related) outcomes; when this need is threatened (such as when our economic resources are low) people tend to compensate by identifying more strongly with larger, more powerful social entities such as the government or a supernatural, powerful God (Kay et al. 2008). Other, more physical level needs show the same goal substitution effect. Immigrants to one's culture can be viewed as analogous to external germs and viruses entering one's own physical body. Satisfying the goal of disease prevention, such as by getting a flu shot or by washing one's hands with antibacterial disinfectant, would hence be expected to substitute for the more abstract goal of defending one's culture by anti-immigration attitudes. Several studies indeed showed that participants who had been inoculated against the flu virus (and reminded of that fact) had less negative attitudes toward immigration than nonimmunized participants; and in another study, those given a chance to use disinfectant hand wipes also then had less negative immigration attitudes (Huang et al. 2011).

Similarly, in the case of another basic physical-level motivation, consumption of food, more abstract (and modern-day) consumption goals would be expected to be "built upon" or "reuse" that more fundamental motive. In harmony with that prediction, Xu et al. (2015) have shown that not only do hungry participants tend to buy more food at the grocery store, as previous studies had shown, but they also buy more of everything, even non-food items at a department store. The goal of consumption or acquisition of resources in general appears therefore to be built on the more fundamental (and survival relevant) goal of food consumption – such that when that food consumption goal is active (that is, when the person is hungry), the more abstract consumption goals are also more active. Not only is it a good idea, from a pragmatic economic standpoint, to having something to eat before going to the grocery store (Gilbert et al. 2002), but it also appears to be a good idea before doing *any* kind of shopping.

There are cultural and individual differences in the use of particular metaphors, and these would be expected to moderate when the physical-to-abstract influences on judgments and behavior will occur. Some metaphors, such as the warm-cold one, appear from the neuroscience research (as well as social psychological research on impression formation; see Fiske et al. 2007) to be hard-wired and universal, but other metaphors appear in some cultures and not in others. This suggests that more than one mechanism might underlie the metaphorical effects; some might be hard-wired, whereas others are culturally based and semantic, reflecting linguistic customs (Meier et al. 2012b).

At the individual difference level, the recent development of the Metaphor Usage Measure (MUM; Fetterman et al. 2016) shows that a person's tendency to use metaphors in his or her everyday communications to others is related to the probability that their behavior will be influenced by analogous physical experiences. The authors' previous research had shown that eating sweet foods was related to having a "sweet" personality (kind, thoughtful, selfless) and also that people in general were more likely to behave "sweetly" after having consumed sweet foods or drink (Meier et al. 2012a). Research on validation of the MUM scale showed that the more a given individual tended to use metaphors (in general) in common speech, the stronger the correlation in their daily lives between how much sweet foods and drinks they consumed and how many "sweet" behaviors they performed that day. Hence, the extent to which associative connections have formed between the physical and social levels of a concept influences both the use of the metaphorical use of the physical concept in language and its degree of influence (when activated by relevant physical experience) on analogous forms of social behavior.

Anderson's (2014) innovative model of neural reuse has been of tremendous help in understanding the basis for the numerous and often remarkable findings of physical experiences on higher-level judgments and behavior over the past decade. This body of evidence, especially its increasing focus on mechanism and individual difference moderators, will be important for sorting out the several distinctive theoretical frameworks that have emerged to account for the metaphorical or analogical influence of the body over the mind. These include evolutionary (Kaschak & Maner 2009), linguistic and symbolic (Boroditsky 2001), simulation-based and perceptually grounded (Barsalou 2008), as well as culturally situated (Cohen & Leung 2009; Oyserman 2011). Some of these theories rely more heavily on a semantic or linguistic mechanism, whereas others emphasize a goal-based evolutionary exaption (Anderson 2008b) of functional units in service pursuit of the

same ultimate underlying end-state – the ancient concerns of survival and prosperity, consisting of the same basic set of primal challenges: self-conservation, social affiliation, mating, and power (Kenrick et al. 2002). As embodiment and neural reuse theory and research themselves evolve together in the years to come, a better understanding of the underlying mechanisms and contextual moderators of embodiment and metaphor effects will certainly be gained by staying mindful of new developments in the theory and research on neural reuse.

Author's Response

Reply to reviewers: Reuse, embodied interactivity, and the emerging paradigm shift in the human neurosciences

Michael L. Anderson
Department of Psychology, Franklin & Marshall College, Lancaster, PA 17604-3003.
michael.anderson@fandm.edu
http://www.fandm.edu/michael-anderson

Abstract: In this reply to reviewers, I argue that, although reforming the taxonomy of psychology will lead to great insights in the cognitive sciences, it will not result in 1:1 structure-function mappings in the brain; we should expect to see a great deal of irreducible functional diversity in the brain at multiple spatial scales. I further clarify both the promise and the limitations of the analytic techniques for capturing functional diversity and interrogating the taxonomy of psychology; describe the ways in which neural reuse can help us understand human development; further explore the ways in which my proposals for integrating psychology, neuroscience, and evolutionary biology differ from the approach exemplified by contemporary evolutionary psychology; and lay out some new and hopefully interesting avenues for future research.

Let me begin by offering my sincere thanks to all of the contributors for taking the time to review and evaluate *After Phrenology* (Anderson 2014), and to *Behavioral and Brain Sciences* for hosting the discussion. We are, I think, at an inflection point in the cognitive sciences generally and in neuroscience in particular, and this collection of essays promises to generate much-needed reflection and stimulate even more rapid progress. Given the scientific moment, I am pleased and humbled to see that even the most critical of the reviewers found the book important, useful, comprehensive, and timely. I am especially grateful for the many reviewers who are calling for an even more sweeping rethinking of the sciences of the mind. If the book helps promote a flowering of creative scientific radicalism, that will be a wonderful legacy, indeed. To finally understand the mind we will need many, many minds, working intensely, freely, and creatively. With a bit of luck, discussions like this will speed the way.

R1. A new taxonomy for psychology?

One of the central contentions of *After Phrenology* (Anderson 2014) is that the taxonomy of psychology is in need of

revision, and both **Shine, Eisenberg, & Poldrack** (**Shine et al.**) and **McCaffery & Machery** appear to agree. Where we disagree is what the outcome of such revision will be. In essence, Shine et al. argue that my position in Anderson (2010) – where I hypothesized that there might be a set of fundamental psychological operations ("workings") that could be strictly localized, but would be used in a variety of cognitive processes – is likely closer to the truth, and they think this model of reuse better accounts for the established data. Perhaps. Their arguments are clear and reasonable, and they usefully illuminate a number of outstanding issues regarding how best to respond to the ever-increasing evidence for functional complexity in the brain. And yet, I find the preponderance of evidence to point not back to my 2010 position, but forward to an understanding that can accommodate a range of functional arrangements, from local workings to true, irreducible polyfunctionality. Hence, are McCaffery & Machery led to ask whether I now actually deny the possibility of intrinsic functions, or if am I arguing instead for the less radical position that the best research strategy is to capture brain function in a multidimensional manner, which will sometimes lead to the discovery of fundamental operations ("workings") and sometimes not?

I certainly endorse the latter statement. My position on intrinsic functions is accordingly a bit more nuanced than denial. I think that *some* parts[1] of the brain, but probably not most, may come to have (relatively) fixed, intrinsic functions (as **D'Souza & Karmiloff-Smith** also insist – but note it is a different open question whether a computational vocabulary will be the best way to describe that function; McCaffery & Machery appear to conflate these two issues). I think that *many* parts of the brain will have multiple, repeatable configurations, and hence different functions in each.[2] And I think that *some* parts of the brain will not have intrinsic functions at all, because the role they play in the coalitions of which they are a member will be irreducibly determined by the mutual constraints imposed by the many interacting parts (Anderson 2015), such that the overall function of the larger system may be unanalyzable into subfunctions. Hence, I do endorse the sort of functional gradation for the brain advanced by Shine et al.; I am just expecting the truth to lie a bit further to the left (as it were). That is, local multifunctionality and unanalyzable distributed functions will be much more common than they suspect – and note **Silberstein** argues that neural reality will prove to be further leftward still.

Because I am an advocate for a new and different taxonomy, I find some of **Kaplan & Craver**'s concerns puzzling. For example, they argue that I "implicitly reif[y] the task domains of BrainMap" (para. 12), thereby trapping psychology inside the very taxonomy I am trying to escape. And yet, I *explicitly* reject any such reification and describe in some detail the many means for avoiding it. These include not just applying various machine learning, pattern analysis, dimensional reduction, and component analysis techniques to get at the hidden underlying structure of neuroimaging (and other neurofunctional) data, but also building fingerprints using multiple different initial taxonomies as starting points. And let me be as clear as I can be: Our current taxonomies are just that: *starting points* for the eventual articulation of a new vocabulary of psychological function.

Of course, Kaplan & Craver are right that solving the problem of functional registration is going to be wickedly difficult, for all of the reasons they cite. Any attempt to integrate scientific evidence across experiments risks including and thus mistakenly interpreting noise, and they worry that my results may reflect more noise than signal. I worry, too. That is why I take into account evidence gathered from many, many different methods, including not just fMRI but also TMS, single cell recording, neural attenuation experiments, and more. I believe that the most parsimonious hypothesis that accounts for *all* of these data is that most parts of the brain are functionally diverse. That said, I do *not* think that any of the fingerprints I have built accurately reflect the true functional profiles of any specific brain region. I think they *do* capture something important about the underlying functional dispositions of those regions (and the fact that fingerprints can be used to predict functional cooperation in the brain suggests this is not an unreasonable supposition), but I do not think we yet have the right way of characterizing underlying function.

Hence, here is where we need to separate the quality and accuracy of our specific current results from the value of the method going forward. As I argue in *After Phrenology*, representing brain activity in a multidimensional way and integrating experimental results across disparate domains, paradigms, and populations is the best way to reach an eventual understanding of the brain. If there really is a 1:1 mapping between function and structure to be found, this method will help us find it (and *not* integrating across different experiments will actually delay this discovery). And if there is *not* a 1:1 mapping to be found, then this method is the best we have for capturing the complex underlying truth. If it is right, as **Pessoa**, **Pasqualotto**, and I all argue, that the mapping between structure and function will remain many:many at multiple spatial scales, then however our taxonomy of function evolves, we need to be using tools that respect this fact.

Although I do believe that the analytic tools I advocate will prove essential to our science going forward, **McCaffery & Machery** raise an important concern regarding the use of unsupervised machine learning and dimensional reduction techniques to revise the taxonomy of psychology. As they note, the techniques are underdetermined, require arguably arbitrary decisions (e.g., in setting the number of dimensions or clusters), do not necessarily converge on a unique solution, and are far from guaranteed to produce meaningful results. Hence, I agree that "none of the possible spaces should be interpreted realistically" (para. 8). But this does *not* mean that none of the *dimensions* should be interpreted realistically. I think that these techniques for exploring large collections of data (neuroimaging data, yes, but also anything else we can get our hands on) will prove to be crucial guides to discovery, precisely *because* they will array things along nonintuitive dimensions and group things in surprising ways. And in this context, the *variety* of the techniques and the *range* of solutions is an asset. Yes, we should avoid the reification pitfall – none of these techniques will speak infallibly. But we should also avoid the cynicism pitfall of never trusting what sometimes fails. As **Perlovsky** notes, much is hidden from conscious awareness, and the development of language further problematizes the clear search for models of brain function, because the apparent structure of language itself can be

(and has been) misconstrued to be a reflection of the underlying structure of the mind (a different instance of reification, perhaps?). By potentially defying our intuitions regarding what is neurally and psychologically similar across contexts, these techniques may help us catch a glimpse of what lies beneath. We should take what they show us seriously.

R2. Brains and bodies

Silberstein's opening line that *After Phrenology* is "the best thing written about the brain this century" (para. 1) already offers the core of my response to his concern that I am being too brain-centric: this appearance is simply the result of the fact that this is a book *about the brain* and how to study it in light of what the best neuroscience, ecological psychology, and evolutionary biology has to teach us. Hence, I completely agree with Silberstein that good neuroscience must also be what he calls "big picture" biology, and I suspect that part of what it will take to make substantial progress understanding the brain is a reform of graduate training in psychology and neuroscience to include more evolutionary and developmental biology, mathematical physics, and, yes, even philosophy (some of which is happening already). I do not expect or advocate for every scientist to master every relevant field, in the manner of TV-show physicians, but if we are to get the best advantage out of the greater interdisciplinary collaboration that *everyone* realizes is necessary, then we need to achieve sufficient familiarity with related disciplines to support effective communication.

Pezzulo usefully highlights one area where such opportunities have been insufficiently explored. He writes: "Progress in the field will benefit from a 'new alliance' between proponents of embodied and action-oriented views [of] cognition and computational modelers including roboticists" (para. 4). I completely agree – and in fact called for exactly this at the most recent conference on Advances in Cognitive Systems! I certainly hope that people heed Pezzulo's (and my) call, and I appreciate the many pointers he offers to underinvestigated theories and phenomena. I will certainly be urging *my* students to follow out some of these ideas.

Interestingly, I think the research path that **Pezzulo** advocates can *also* be an important part of addressing **Silberstein**'s concerns. I agree with Silberstein and Pezzulo that the notion of *constraint* is going to provide a crucial organizing frame for the study of organism-environment systems (Anderson 2015), and cognitive robotics looks to be one important technique for exploring the essential questions of how the environmental and other constraints that Silberstein rightly highlights help determine brain function, and how brains are organized to take advantage of said constraints for cognitive and behavioral ends. As I note above, I am open to the notion that some parts of the brain lack intrinsic functions that can be localized, because they will be irreducibly determined by properties of whole systems, and I am also open to the possibility that *some* intrinsic functions can in fact be identified. Similarly, as Pezzulo advocates, I remain open to the discovery that *some* elements of *some* brain mechanisms will be best understood as internal representations, but I do not expect them to be central to our overall account of cognition.

Silberstein, Pezzulo, and I are in complete accord that the neurosciences (and the cognitive sciences more generally) need to continue to move toward more embodied, embedded, dynamic, control-theoretical models of psychological phenomena.

Whereas authors including **Pasqualotto**, **Pessoa**, and **Guida, Campitelli, & Gobet** (**Guida et al.**) focus primarily on the neural evidence for the ubiquity of reuse, **Wang & Bargh** offer instead an excellent discussion of some of the cognitive evidence for and effects of reuse, including perceptual grounding – what they call "conceptual scaffolding" – and goal substitution. For example, they detail the fascinating ways in which physical sensations – of warmth, texture, and the like – both influence and are influenced by interpersonal social judgments of closeness and cohesion. My only concern with their account is that, although the associative connections they suggest are responsible for this influence are surely part of the story, in fact explicit associations may not always be necessary to generate the effects they discuss. As I have argued along with Penner-Wilger (e.g., Penner-Wilger & Anderson 2013), all that is required to explain a given instance of reuse is that the reused neural element have the right functional properties to support each use. In the case of the finger-number relationship that we discuss, there need not have been direct experience relating fingers to numbers (for example, counting on one's fingers) for the relationship to obtain – it is enough that a particular part of the brain has the right functional structure to aid both sensory and numerical tasks.

Similarly, the connections between the literal physical experiences of closeness, warmth, and sweetness, and their corresponding social metaphors, *might* involve psychological associations, but it might also be simply that the relevant homeostatic mechanism has the right structure to help manage both physical and social relationships. If this is the case, then activating the mechanism for either purpose will often cause multiple effects and experiences, regardless of whether an association has been formed. Naturally, once this functional relationship exists, it could well *lead* to the formation of a psychological association, although not necessarily a conscious one. And I agree that, once these relationships are captured in language, then the causal story becomes significantly more complex and is likely to involve not just low-level neural mechanisms, but also conceptual and cultural ones.

R3. Which way forward?

D'Souza & Karmiloff-Smith suggest that an even greater focus on development could greatly enhance and strengthen the neural reuse framework. The questions they raise are clearly important, and I share their desire to see them addressed. As I tried to showcase in the book, I think that enhancing neuroconstructivism with some ideas and concepts arising within the reuse framework leads to a theory with greater explanatory power than either framework on its own. D'Souza & Karmiloff-Smith's own example of the complex variation observed in an individual's ability to recover from brain injury over developmental time offers a case in point: As they admit, neither the hypothesis of increasing neural commitment (decreasing neural plasticity) nor greater early vulnerability account for all of the data. But by combining these ideas

with two from the neural reuse framework, one can perhaps do a bit better. Why do some early brain injuries result in worse outcomes than those that come a few years later? Perhaps because there is a crucial time for establishing the base-set of local cortical biases that will be woven into functional neural coalitions; if the "normal" base-set is missing crucial members or otherwise altered, so too will be the developmental trajectory that the individual will follow. So, from the perspective of the reuse framework, the observed "early vulnerability" is not intrinsic to the early-developing neurons themselves, but rather it is an aspect of the delicate early interactions between Hebbian plasticity and neural reuse. Similarly, reuse can potentially shed light on why it is sometimes possible to reverse the perceptual narrowing that is otherwise a crucial part of the language learning process. For reuse, much of the apparent decrease in plasticity over development comes *not* because Hebbian tuning is intrinsically difficult to reverse, but because any given local network becomes incorporated into multiple functional coalitions, and interactions with the other partnerships tend to reinforce existing configurations – a situation I describe using the evolutionary notion of "burden." Neural commitment – functional tuning – does of course happen. But neurons also become *burdened* by the multiple uses they support, and this is an important contributor to the observed loss of behavioral and perceptual plasticity. In situations where that burden can be lessened or released, then the underlying neuroplasticity can be released and retuned. This kind of developmental thinking is in fact central to *After Phrenology*, although it is true that I did not discuss as much *developmental phenomena and data* as might have been desirable. I would like to thank the authors for highlighting some of these important findings, and I look forward to the time when we can answer the many crucial questions they raise here.

Parkinson & Wheatley and **Stanley & De Brigard** argue that the greater use of pattern analysis and graph theory (respectively) will smooth the path ahead. I agree, discuss both approaches at length in the book, and welcome the renewed emphasis these authors provide here. Parkinson & Wheatley may be right that neuroscience has finally turned decisively away from modular architectures (**Pessoa** and **Badcock, Ploeger, & Allen** [**Badcock et al.**] suggest the same), at least rhetorically, but a stroll through the poster session at any Society for Neuroscience meeting will quickly reveal that *many* bad habits remain – of assigning domain-restricted specialized functions to regions studied only under a single narrow range of conditions, of strictly separating perception, cognition, and action, both psychologically and neurally, and of making unwarranted reverse inferences, just to name a few (although **McCaffery & Machery** are right about the conditions under which reverse inference can be informative, in practice these conditions are rarely met or even explicitly considered). So perhaps the arguments against modular thinking have some work to do yet. In any case, I agree that getting in to the habit of analyzing data for distributed patterns will be an important part of the behavioral therapy that is (still) needed and promises to open up new horizons in our understanding of brain function – a promise that Parkinson & Wheatley beautifully illustrate in their own work on distance perception.

Similarly, **Stanley & De Brigard** are surely right that graph theory will remain a crucial tool for studying the

networks of the brain (Sporns 2011), and I agree that finding "modules"[3] can sometimes lead to useful insights about brain structure and function. My only concern, one that **Pessoa** shares, is that the techniques for identifying network communities assume that each node is a member of exactly one community (which may also belong to a hierarchy). Stanley & De Brigard are right to emphasize that graph theory can help us detect *dynamic and changing* affiliations, but if nodes can *also* be members of more than one community in a given moment, then community detection algorithms will generally miss this. This does not lessen the importance of network analysis, but it does mean the results should be interpreted with due caution.

Perlovsky echoes **D'Souza & Karmiloff-Smith** in noting the challenges involved in specifying the underlying mechanisms of neural search; indeed, as I note in the book, this is one area ready for greater research attention. But Perlovsky's way of putting it implies that neural reuse requires the directed capacity to target specific subnetworks with which to establish new connections – as if the brain contained representations of what its various parts were capable of. Were neural reuse a process of network *design,* this might be an apt description, and I would agree with Perlovsky that the prospects for understanding reuse would indeed be dim. But this is a misunderstanding of the mechanism. Instead, as I explicitly argue in the book, what we seem to have is a parallel *search* process and a test/ consolidation mechanism that results in new functional configurations. We see such a mechanism at work in brain-machine interfaces when, during the course of learning to control a new device, the existing coherence pattern in the relevant network is disrupted by both an overall increase in activity and an increase in the *variability* of the activity of each neural element. As I describe in *After Phrenology*, the effect of these two simple changes is to implement a systematic walk through coherence space – that is, it generates a *search* of possible functional configurations, with no anticipatory design goal or other telos required. As this search of possibilities is occurring, reward signals generated by successful trials reinforce the configurations responsible for the success, leading to the eventual consolidation of the effective configurations. I will be the first to admit that we do not yet fully understand how this works; but as **Guida et al.**'s comprehensive review of the expertise literature demonstrates, *that* it works is increasingly established, and we may be closer to understanding *how* than Perlovsky fears.

Perlovsky also asks how it is that we can simultaneously treat external symbols as *objects* to be manipulated, and as *meaningful*. It is an important question. I tried to point toward an answer with my notion of a cultural affordance, which Perlovsky flatteringly calls "beautiful" but then dismisses as "inexorably logical." Here, Perlovsky simply makes a mistake: Cultural affordances may or may not turn out to provide part of the answer to this question, but to assimilate them to logic, as if all perceivable structure had to be objective, symbolic, and representational, is to miss the nature of the *alternative* being offered. Objects in the world are inherently significant to animals because of the relationships they have to abilities and needs. These relationships are directly perceivable and guide an organism's behavior. Cultural practices co-opt this basic behavior-guiding mechanism in a way that results in

shared meaning in the context of dynamic social interactions. Here again, we do not yet understand this very well (but see Atmaca et al. 2008; Marsh et al. 2009; Richardson & Dale 2005; Richardson et al. 2007; Sebanz et al. 2006, to name just a few important contributions), and it is therefore no surprise that some scientists – including Perlovsky; **Badcock et al.,** and others – remain skeptical or believe an affordance-centered cognitive science to be a mere variation on or extension of contemporary representation-centered approaches. This is why, in addition to chiding (with good justification, I might add!) ecological psychologists for largely ignoring the brain, experimental psychologists for largely ignoring evolution (*not* evolutionary psychologists, but we will come to that), and neuroscientists for largely ignoring organisms, I labored to provide a unified framework that could help illuminate for all how this work fits together and is mutually informative. Clearly, it will take more than this one effort for that promise to be fulfilled, but I think it is fair to say that *After Phrenology* offers the most comprehensive framework to date, and I do think it can be used to generate greater interdisciplinary understanding.

R4. Is this the revolution we were promised?

Of course, interdisciplinary accord requires that we adopt and share *some* unifying framework, and **Badcock et al.** are reluctant to adopt this one. Although they agree that massive modularity "can no longer be reasonably sustained" (para. 1) (a concession that, while perhaps trivial for these particular scientists, is in fact going to have profound and far-reaching effects across the behavioral and life sciences as its ramifications come gradually to be more fully understood), they are skeptical of the claim that perception is not reconstructive – a skepticism that stands in the way of recognizing that a paradigm shift may be in the offing. I guess I agree that if we do not give up reconstructive perception (and the perception-first, stimulus-response, sense-think-act model of psychology more generally), then the coming revolution will be *less* radical than it might otherwise be. Fully accommodating neural reuse will mean *only* that we must reimagine the functional architecture of the brain (as **Pasqualatto, Parkinson & Wheatley, Pessoa,** and **Silberstein** all appear to recognize), reconsider the relationships between psychological states and processes once thought distinct (some of the implications of which are reviewed by **Wang & Bargh, Shine et al.,** and **McCaffery & Machery**) and rethink the relationship between evolution, plasticity, and development (as **D'Souza & Karmiloff-Smith** urge us to do in a much more comprehensive way). Giving up on reconstructive perception means that in addition we must fundamentally change our idea of what the brain is *for*, and therefore *how* it does what it does and how we do what we do.

Badcock et al. are skeptical not just that we are going to *get* a paradigm shift, but also that we *need* one, because they think we may already know more-or-less what the brain is for, how it does what it does, and why we do what we do. In support of this conservatism, they point to the success of the predictive coding framework, on the one hand, and the fertility of evolutionary psychology, on the other. These authors are right that the book engages with the predictive coding framework only very indirectly, via a few scattered references to Bayesian networks and a

brief discussion of causal learning (to help excuse this lacuna, I plead the necessity to restrain what was already an overly long book). For the record, then, I think the predictive coding framework represents an impressive and highly successful empirical research project. I think it is undeniable that brain function involves a great deal of prediction – in a sense of "prediction" closely allied with the notion of correlation, as when we commonly say that the value of one variable "predicts" another (height predicts weight; education predicts income, etc.; see Anderson & Chemero 2013 for further discussion). But I do *not* think that the work merits the cognitivist, representationalist, reconstuctive gloss that it is commonly given by its main proponents (e.g., Clark 2013b) – and it is worth noting that **Pezzulo**, who associates predictive coding with the cybernetic framework (Seth 2015), apparently sees things similarly. This is a debate that is only in its infancy (Anderson & Chemero 2013; Barrett & Bar 2009; Gallagher & Bower 2014; Gładziejewski 2016; Hohwy 2013), but my personal starting position is that none of the empirical results emerging from the predictive coding literature *requires* cognitivist, reconstructive explanations. They can be read as compatible with reconstructive perception or with an action-oriented, affordance-based framework (and indeed, I doubt the adjudication between these interpretations is going to rest on results from this literature, but rather on the relative fertility of the competing frameworks going forward – but that is yet another debate). The conclusion of my discussion of causal learning in *After Phrenology* seems equally apropos here:

> As we perceive and act in the world, we are learning to see what the word affords and where and how to intervene to generate preferred outcomes, and we are at the same time inducing the neural structures that make such control possible. The causal knowledge we acquire appears to be best understood as a guide to action, written primarily in the vocabulary of sensorimotor contingencies. In this sense the literature on causal learning appears to be solidly within the pragmatist tradition despite the cognitivist (structuralist) gloss applied to it by many of its proponents. (Anderson 2014, pp. 191–92)

Possibly **Badcock et al.** could counter that the representationalist gloss serves as a research-guiding heuristic, an aid to hypothesis generation to be understood instrumentally rather than as a substantive hypothesis in its own right. To see what is worrisome in such proposals, consider their discussion of massive modularity, where they suggest exactly this: "Regardless of the veracity of massive modularity, evolutionary computational theories continue to guide research in a systematic and highly productive way" (para. 7). In support of this contention, Badcock et al. cite the example of inherited perceptual biases toward threatening stimuli – for example, snake detection – and argue that this can be explained *only* by reference to an evolutionary adaptation. As a critique of *After Phrenology*, this misses the mark, for *of course* I fully support integrating psychology and evolutionary biology and encourage thinking about the mind as one of the many products of natural selection. But I also want to better understand the mechanisms of selection and inheritance, in a way that is sensitive to what neuroscience tells us about the architecture of the brain. It is here that evolutionary psychology, as they conceive it, is in a bit

of a bind. For scientists like Badcock et al. are faced with a choice: either posit a snake-detection module, despite overwhelming evidence that the brain isn't built that way, or shrug your shoulders over the "how" question and move on to the next hypothesis. Badcock et al. seem content with the latter, but I am not, and I do not think anyone else should be, either. Letting each segregated subfield pursue its fancy in isolation is *not* a recipe for scientific understanding.

Still, **Badcock et al.** would be right to worry if the framework I am advocating were in fact unable to provide substantive hypotheses about cognition and behavior. Although time will provide the true test of their contention, as we discover whether my proposals resonate enough to spur the scientific imagination and continue to generate fruitful research activity, it is perhaps worth noting that **Wang & Bargh** call the hypothesis "remarkably generative and supportive of research activity on embodied cognition, motivation, and behavior" (para. 1), and they, along with **Pasqualotto** and **D'Souza & Karmiloff-Smith** further suggest that it can help account for phenomena as diverse as perceptual grounding, goal substitution, synesthesia, and cross-modal plasticity. I can also point to some things that are happening in *my* lab as we continue to test, refine, and further substantiate the reuse framework. Because we argue that affordance processing underlies much of higher-order cognition, including language (Glenberg & Kaschak 2002; Kaschak & Glenberg 2000) and decision making (Cisek 2007; Cisek & Kalaska 2010), we believe that indirectly manipulating affordances – by using real objects rather than pictures in psychological experiments, for example (Snow et al. 2011), or changing the context within which an experiment takes place – will allow us to modulate higher-order cognitive outcomes such as similarity judgments and categorization behavior. Moreover, we have specific hypotheses about how affordances – in a nonmodular brain – integrate with emotions, and hence how emotion induction can change the affordance landscape in ways leading to detectable cognitive and behavioral consequences. The first set of experiments is already designed (and by the time this appears in print, the first data will have been gathered). So please stay tuned; there is plenty more to come!

NOTES

1. By "parts," I mean neurons, regions, networks; I did not *intend* to be conservative about brain ontology revision, as McCaffery & Machery suggest.

2. This was the possibility that my notion of Transiently Assembled Local Neural Subsystems (TALoNS) was intended to capture. I do not know how distinct or fuzzy the distinctions between these configurations will be, or how different the resulting functions can be – here is yet another question ripe for attention.

3. Here, I mean "module" in the network theory sense of the term, which differs significantly from the sense of the term as used in cognitive and evolutionary psychology. In network theory, a module is a set of nodes with a specific structural relationship to one another – generally, they are more densely connected to one another than to other nodes in the graph. This "module" is a structural or topological descriptor. In cognitive psychology and evolutionary psychology, a module is a system with very specific functional properties, such as automatic fast operation, domain dedication, separate modifiability, and so forth. This "module" is a functional descriptor. It is unfortunate and confusing that the same term is used in these very different ways.

References

[The letters "a" and "r" before author's initials stand for target article and response references, respectively]

Aboitiz, F. (1996) Does bigger mean better? Evolutionary determinants of brain size and structure. *Brain, Behavior and Evolution* 47:225–45. [aMLA]

Abrahamse, E., van Dijck, J.-P., Majerus, S. & Fias, W. (2014) Finding the answer in space: The mental whiteboard hypothesis on serial order in working memory. *Frontiers in Human Neuroscience* 8:932. [AG]

Ackerman, J. M., Nocera, C. C. & Bargh, J. A. (2010) Incidental haptic sensations influence social judgments and decisions. *Science* 328:1712–15. [YW]

Agnati, L. F., Guidolin, D., Guescini, M., Genedani, S. & Fuxe, K. (2010) Understanding wiring and volume transmission. Brain Research Review 64:137–59. [aMLA]

Alain, C., Woods, D. L. & Knight, R. T. (1998) A distributed cortical network for auditory sensory memory in humans. *Brain Research* 812:23–37. [AP]

Amad, A., Cachia, A., Gorwood, P., Pins, D., Delmaire, C., Rolland, B., Mondino, M., Thomas, P. & Jardri, R. (2014) The multimodal connectivity of the hippocampal complex in auditory and visual hallucinations. *Molecular Psychiatry* 19:184–91. [AP]

Amedi, A., Raz, N., Pianka, P., Malach, R. & Zohary, E. (2003) Early "visual" cortex activation correlates with superior verbal memory performance in the blind. *Nature Neuroscience* 6:758–66. [AP]

Andersen, R. A., Essick, G. K. & Siegel, R. M. (1985) Encoding of spatial location by posterior parietal neurons. *Science* 230:456–58. [AP]

Anderson, M. L. (2008a) Circuit sharing and the implementation of intelligent systems. *Connection Science* 20(4):239–51. [aMLA]

Anderson, M. L. (2008b) Evolution, embodiment and the nature of the mind. In: *Beyond the brain: Embodied, situated and distributed cognition*, ed. B. Hardy-Vallee & N. Payette, pp. 15–28. Cambridge Scholar's Press. [YW]

Anderson, M. L. (2010) Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences* 33(4):245–66. doi: 10.1017/S0140525X10000853. [arMLA, JBM, CP, AP]

Anderson, M. L. (2014) *After phrenology: Neural reuse and the interactive brain.* MIT Press. [arMLA, PBB, DD'S, AG, DMK, JBM, CP, AP, LPer, LPes, GP, JMS, MS, MLS, YW]

Anderson, M. L. (2015) Beyond componential constitution in the brain: Starburst amacrine cells and enabling constraints. In: *Open MIND*, ed. T. K. Metzinger & J. M. Windt. MIND Group. doi: 10.15502/9783958570429. [arMLA]

Anderson, M. L. & Chemero, T. (2013) The problem with brain GUTs: Conflation of different senses of "prediction" threatens metaphysical disaster. *Behavioral and Brain Sciences* 36(3):204–205. [rMLA]

Anderson, M. L. & Finlay, B. (2014) Allocating structure to function: The strong links between neuroplasticity and natural selection. *Frontiers in Human Neuroscience* 7:918. [aMLA]

Anderson, M. L., Kinnison, J. & Pessoa, L. (2013) Describing functional diversity of brain regions and brain networks. *NeuroImage* 73:50–58. [aMLA, JBM, JMS]

Anderson, M. L. & Penner-Wilger, M. (2013) Neural reuse in the evolution and development of the brain: Evidence for developmental homology? *Developmental Psychobiology* 55(1):42–51. [aMLA]

Anderson, V., Jacobs, R., Spencer-Smith, M., Coleman, L., Anderson, P., Williams, J., Greenham, M. & Leventer, R. (2010) Does early age at brain insult predict worse outcome? Neuropsychological implications. *Journal of Pediatric Psychology* 35(7):716–27. [DD'S]

Anderson, V., Spencer-Smith, M. & Wood, A. (2011) Do children really recover better? Neurobehavioural plasticity after early brain insult. *Brain* 134(8):2197–221. [DD'S]

Arenas, A., Fernandez, A. & Gomez, S. (2008) Analysis of the structure of complex networks at different resolution levels. *New Journal of Physics* 10:053039. doi: 10.1088/1367-2630/10/5/053039. [MLS]

Ashby, W. R. (1952) *Design for a brain*. Wiley. [GP]

Atkinson, J. (1984) Human visual development over the first 6 months of life. A review and a hypothesis. *Human Neurobiology* 3(2):61–74. [aMLA]

Atmaca, S., Sebanz, N., Prinz, W. & Knoblich, G. (2008) Action co-representation: The joint SNARC effect. *Social Neuroscience* 3(3–4):410–20. [rMLA]

Attias, H. (2003) Planning by Probabilistic Inference. *Proceedings of the Ninth International Conference on Artificial Intelligence and Statistics, Key West, FL. January 2003*, ed. C. M. Bishop & B. J. Frey. Society for Artificial Intelligence and Statistics. [GP]

Bach-Y-Rita, P., Collins, C. C., Saunders, F. A., White, B. & Scadden, L. (1969) Vision substitution by tactile image projection. *Nature* 221:963–64. [AP]

Badcock, P. B. (2012) Evolutionary systems theory: A unifying meta-theory of psychological science. *Review of General Psychology* 16:10–23. doi: 10.1037/a0026381. [PBB]

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R. & Halgren, E.

(2006) Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences USA* 103:449–54. [LPer]

Barabasi, A.-L. & Oltvai, Z. (2004) Network biology: Understanding the cell's functional organization. *Nature Reviews Genetics*, 5:101–13. doi: 10.1038/nrg1272. [MLS]

Barbas, H. and Zikopolous, B. (2007) The prefrontal cortex and flexible behavior. *The Neuroscientist* 13(5):532–45. [JMS]

Bargh, J. A. & Shalev, I. (2012) The substitutability of physical and social warmth in daily life. *Emotion* 12(1):154–62. [aMLA]

Bargmann, C. I. (2012) Beyond the connectome: How neuromodulators shape neural circuits. *Bioessays* 34(6):458–65. [aMLA]

Barkow, J., Cosmides, L. & Tooby J., eds. (1992) *The adapted mind*. Oxford University Press. [DD'S]

Barnes, J., ed. (1984) *The complete works of Aristotle: The revised Oxford translation*. Princeton University Press. [LPer]

Barrett, H. C. (2008) Evolutionary psychology is a union of fields, not a subset. *Psychological Inquiry* 19:19–40. doi: 10.1080/10478400701774022. [PBB]

Barrett, H. C. & Kurzban, R. (2006) Modularity in cognition: Framing the debate. *Psychological Review* 113(3):628. [aMLA]

Barrett, L. F. & Bar, M. (2009) See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences* 364(1521):1325–34. [rMLA]

Barrett, L. F. & Satpute, A. (2013) Large-scale brain networks in affective and social neuroscience: Towards an integrative architecture of the human brain. *Current Opinion in Neurobiology* 23:1–12. [aMLA]

Barsalou, L. W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences* 22 (4):577–660. [LPer, GP]

Barsalou, L. W. (2008) Grounded cognition. *Annual Review of Psychology* 59:617–45. [GP, YW]

Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M. & Grafton, S. T. (2011) Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences of the United States of America* 108:7641–46. doi: 10.1073/pnas.1018985108. [MLS]

Bassett, D. S., Yang, M., Wymbs, N. F. & Grafton, S. T. (2015) Learning-induced autonomy of sensorimotor systems. *Nature Neuroscience* 18(5):744–51. [JMS]

Bates, E., Reilly, J., Wilfeck, B., Donkers, N., Opie, M., Fenson, J., Kriz, S., Jeffries, R, Miller, L. & Herbst, K. (2001) Differential effects of unilateral lesions on language production in children and adults. *Brain and Language* 79:223–65. [DD'S]

Bateson, P. & Gluckman, P. (2011) *Plasticity, robustness, development and evolution*. Cambridge University Press. [MS]

Bauer, M., Heinz, A. & Whybrow, P. C. (2001) Thyroid hormones, serotonin and mood: Of synergy and significance in the adult brain. *Molecular Psychiatry* 7 (2):140–56. [aMLA]

Bavelier, D., Levi, D. M., Li, R. W., Yang, D. & Hensch, T. K. (2010) Removing brakes on adult brain plasticity: From molecular to behavioral interventions. *The Journal of Neuroscience* 30(45):14964–71. [DD'S]

Beauchamp, M. S. (2005) See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology* 15:145–53. [AP]

Beauchamp, M. S., Yasar, N. E., Frye, R. E. & Ro, T. (2008) Touch, sound and vision in human superior temporal sulcus. *NeuroImage* 41:1011–20. [AP]

Beer, A. L., Plank, T. & Greenlee, M. W. (2011) Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. *Experimental Brain Research* 213:299–308. [AP]

Beer, A. L., Plank, T., Meyer, G. & Greenlee, M. W. (2013) Combined diffusion-weighted and functional magnetic resonance imaging reveals a temporal-occipital network involved in auditory-visual object processing. *Frontiers in Integrative Neuroscience* 7:5. [AP]

Birdsong, D. (2006) Age and second language acquisition and processing: A selective overview. *Language Learning* 56:9. [DD'S]

Boroditsky, L. (2001) Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology* 43(1):1–22. [YW]

Borra, E. & Rockland, K. S. (2011) Projections to early visual areas V1 and V2 in the calcarine fissure from parietal association areas in the macaque. *Frontiers in Neuroanatomy* 5:35. [AP]

Bressler, S. L. & Menon, V. (2010) Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences* 14:277–90. [LPes]

Brett, M., Johnsrude, I. S. & Owen, A. M. (2002) The problem of functional localization in the human brain. *Nature Reviews Neuroscience* 3(3):243–49. [DMK]

Bruce, C., Desimone, R. & Gross, C. G. (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology* 46:369–84. [AP]

Buckner, R. L., Krienen, F. M. & Yeo, B. T. (2013) Opportunities and limitations of intrinsic functional connectivity MRI. *Nature Neuroscience* 16:832–37. [LPes]

Buller, D. J. (2005) *Adapting minds: Evolutionary psychology and the persistent quest for human nature*. MIT Press.   [PBB]

Bullmore, E. T. & Sporns, O. (2009) Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* 10:186–98. doi: 10.1038/nrn2618.   [MLS]

Burgess, N., Maguire, E. A. & O'Keefe, J. (2002) The human hippocampus and spatial and episodic memory. *Neuron* 35:625–41.   [AG]

Burgess, N., Maguire, E. A., Spiers, H. J. & O'Keefe, J. (2001) A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *NeuroImage* 14:439–53.   [AP]

Buss, D. M., ed. (2005) *The handbook of evolutionary psychology*. John Wiley.   [aMLA]

Buss, D. M. (2012) *Evolutionary psychology: The new science of the mind, fourth edition*. Allyn & Bacon.   [PBB]

Buzsaki, G. (2006) *Rhythms of the brain*. Oxford University Press.   [DD'S]

Buzsáki, G. & Moser, E. I. (2013) Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Natural Neuroscience* 16:130–38. doi: 10.1038/nn.3304.   [GP]

Buzsáki, G., Peyrache, A. & Kubie, J. (2015) Emergence of cognition from action. *Cold Spring Harbor Symposia on Quantitative Biology* 79:41–50. doi: 10.1101/sqb.2014.79.024679.   [GP]

Cacioppo, J. T. & Tassinary, L. G. (1990) Inferring psychological significance from physiological signals. *American Psychologist* 45:16–28.   [LPes]

Campitelli, G., Gobet, F., Head, K., Buckley, M. & Parker, A. (2007) Brain localisation of memory chunks in chessplayers. *International Journal of Neuroscience* 117:1641–59.   [AG]

Cao, H., Plichta, M. M., Schäfer, A., Haddad, L., Grimm, O., Schneider, M. & Tost, H. (2014) Test-retest reliability of fMRI-based graph theoretical properties during working memory, emotion processing and resting state. *NeuroImage* 84:888–900. doi: 10.1016/j.neuroimage.2013.09.013.   [MLS]

Caporael, L. R. (1997) The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review* 1(4):276–98. doi: 10.1207/s15327957pspr0104_1.   [PBB]

Caporael, L. R. (2001) Evolutionary psychology: Toward a unifying theory and a hybrid science. *Annual Review of Psychology* 52:607–28. doi: 10.1146/annurev.psych.52.1.607.   [PBB]

Carandini, M. & Heeger, D. J. (2012) Normalization as a canonical neural computation. *Nature Reviews Neuroscience* 13(1):51–62.   [DMK]

Carruthers, P. (2006) *The architecture of the mind: Massive modularity and the flexibility of thought*. Oxford University Press.   [aMLA]

Casey, B. J., Duhoux, S. & Cohen, M. M. (2010) Adolescence: What do transmission, transition, and translation have to do with it? *Neuron* 67:749–60.   [DD'S]

Chang, S. W. C., Brent, L. J. N., Adams, G. K., Klein, J. T., Pearson, J. M., Watson, K. K. & Platt, M. L. (2013) Neuroethology of primate social behavior. *Proceedings of the National Academy of Sciences of the United States of America* 110(Suppl.):10387–94. doi: 10.1073/pnas.1301213110.   [CP]

Chase, W. G. & Simon, H. A. (1973) Perception in chess. *Cognitive Psychology* 4:55–81.   [AG]

Chemero, A. (2009) *Radical embodied cognitive science*. MIT Press.   [GP]

Chen, Z. & Cowan, N. (2005) Chunk limits and length limits in immediate recall: A reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31:1235–49.   [AG]

Choi, J. M., Padmala, S. & Pessoa, L. (2012) Impact of state anxiety on the interaction between threat monitoring and cognition. *NeuroImage* 59:1912–23.   [LPes]

Chomsky, N. (1957) *Syntactic structures*. Mouton.   [aMLA]

Cisek, P. (1999) Beyond the computer metaphor: Behaviour as interaction. *Journal of Consciousness Studies* 6(11–12):125–42.   [aMLA, GP]

Cisek, P. (2006) Integrated neural processes for defining potential actions and deciding between them: A computational model. *The Journal of Neuroscience* 26:9761–70.   [GP]

Cisek, P. (2007) Cortical mechanisms of action selection: The affordance competition hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1485):1585–99.   [arMLA]

Cisek, P. & Kalaska, J. F. (2005) Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron* 45(5):801–14.   [aMLA]

Cisek, P. & Kalaska, J. F. (2010) Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience* 33:269–98.   [arMLA, GP]

Clark, A. (1997) *Being there: Putting brain, body, and world together again*. MIT Press.   [aMLA]

Clark, A. (1998) *Being there. Putting brain, body, and world together*. MIT Press.   [GP]

Clark, A. (2013a) Are we predictive engines? Perils, prospects, and the puzzle of the porous perceiver. *Behavioral and Brain Sciences* 36(3):233–53. doi: 10.1017/S0140525X12002440.   [PBB]

Clark, A. (2013b) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* 36(3):181–204. doi: 10.1017/S0140525X12000477.   [PBB, GP, rMLA]

Clark, A. & Chalmers, D. J. (1998) The extended mind. *Analysis* 58:10–23.   [GP]

Clark, A. & Grush, R. (1999) Towards a cognitive robotics. *Adaptive Behavior* 7:5–16.   [GP]

Clark, H. H. (1996) *Using language*. Cambridge University Press.   [aMLA]

Cocchi, L., Zalesky, A., Fornito, A. & Mattingley, J. B. (2013) Dynamic cooperation and competition between brain systems during cognitive control. *Trends in Cognitive Sciences* 17:493–501.   [LPes]

Cohen, D. & Leung, A. K. Y. (2009) The hard embodiment of culture. *European Journal of Social Psychology* 39(7):1278–89.   [YW]

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S. & Petersen, S. E. (2014) Intrinsic and task-evoked network architectures of the human brain. *Neuron* 83:238–51.   [LPes]

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A. & Braver, T. S. (2013) Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience* 16(9):1348–55.   [aMLA, LPes]

Confer, J. C., Easton, J. A., Fleischman, D. S., Goetz, C. D., Lewis, D. M., Perilloux, C. & Buss, D. M. (2010) Evolutionary psychology: Controversies, questions, prospects, and limitations. *American Psychologist* 65(2):110–26.   [aMLA]

Connolly, A. C., Gobbini, M. I. & Haxby, J. V. (2012) Three virtues of similarity-based multivariate pattern analysis: An example from the human object vision pathway. In: *Understanding visual population codes: Toward a common multivariate framework for cell recording and functional imaging*, ed. N. Kriegeskorte & G. Kreiman, pp. 335–55. MIT Press.   [CP]

Costafreda, S. G. (2009) Pooling fMRI data: Meta-analysis, mega-analysis and multi-center studies. *Frontiers in Neuroinformatics* 3:33.   [DMK]

Cowan, N. (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24:87–185.   [AG]

Cowan, N. (2011) The focus of attention as observed in visual working memory tasks: Making sense of competing claims. *Neuropsychologia* 49:1401–406.   [AG]

Cowan, N., Chen, Z. & Rouder, J. N. (2004) Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science* 15:634–40.   [AG]

Craik, K. (1943) *The nature of explanation*. Cambridge University Press.   [GP]

Craver, C. (2007) *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford University Press.   [aMLA]

Crawford, C. B. & Krebs, D. (2008) *Foundations of evolutionary psychology*. Erlbaum.   [PBB]

Curtis, C. E. (2006) Prefrontal and parietal contributions to spatial working memory. *Neuroscience* 139(1):173–80.   [JMS]

Damaraju, E., Huang, Y. M., Barrett, L. F. & Pessoa, L. (2009) Affective learning enhances activity and functional connectivity in early visual cortex. *Neuropsychologia* 47:2480–87.   [LPes]

De Gelder, B. & Bertelson, P. (2003) Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences* 7:460–67.   [AP]

Decharms, R. C. & Zador, A. (2000) Neural representation and the cortical code. *Annual Review of Neuroscience* 23:613–47.   [AP]

Dehaene, S. (2005) Evolution of human cortical circuits for reading and arithmetic: The "neuronal recycling" hypothesis. In: *From monkey brain to human brain: A Fyssen Foundation symposium*, ed. S. Dehaene, J.-R. Duhamel, M. D. Hauser & G. Rizzolatti, pp. 133–58. MIT Press.   [GP]

Dehaene, S. & Cohen, L. (2007) Cultural recycling of cortical maps. *Neuron* 56(2):384–98. doi: 10.1016/j.neuron.2007.10.004.   [CP]

Desgranges, B., Baron, J. C. & Eustache, F. (1998) The functional neuroanatomy of episodic memory: The role of the frontal lobes, the hippocampal formation, and other areas. *NeuroImage* 8:198–213.   [AP]

Desimone, R. & Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annual Review of Neuroscience* 18(1):193–222.   [aMLA]

Dewsbury, D. A. (2009) More on evolution and psychology. *American Psychologist* 64:624–25. doi: 10.1037/a0017380.   [PBB]

Diba, K. & Buzsáki, G. (2007) Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience* 10:1241–42. doi: 10.1038/nn1961.   [GP]

Dragoi, G. & Tonegawa, S. (2011) Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature* 469:397–401.   [GP]

Driver, J. & Noesselt, T. (2008) Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron* 57:11–23.   [AP]

Dubois, J., de Berker, A. O. & Tsao, D. Y. (2015) Single-unit recordings in the macaque face patch system reveal limitations of fMRI MVPA. *The Journal of Neuroscience* 35(6):2791–802.   [JMS]

Duffau, H. (2008) The anatomo-functional connectivity of language revisited: New insights provided by electrostimulation and tractography. *Neuropsychologia* 46:927–34.   [AP]

Dunbar, R. I. M. & Barrett, L., eds. (2007) *Oxford handbook of evolutionary psychology*. Oxford University Press.   [PBB]

Edelman, S. (2008) *Computing the mind: How the mind really works*. Oxford University Press.   [aMLA]

Eimas, P. D. (1975) Auditory and phonetic coding of the cues for speech: Discrimination of the /r-l/ distinction by young infants. *Perception and Psychophysics* 18(5):341–47. [DD'S]

Ellis, B. J. & Bjorklund, D. F., eds. (2005) *Origins of the social mind: Evolutionary psychology and child development.* Guilford. [PBB]

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996) *Rethinking innateness: A connectionist perspective on development.* MIT/Bradford Books. [DD'S]

Engel, A. K., Maye, A., Kurthen, M. & König, P. (2013) Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences* 17:202–209. [GP]

Ericsson, K. A. & Kintsch, W. (1995) Long-term working memory. *Psychological Review* 102:211–45. [AG]

Fetterman, A. K., Bair, J. L., Werth, M., Landkammer, F. & Robinson, M. D. (2016) The scope and consequences of metaphoric thinking: Using individual differences in metaphor usage to understand how metaphor functions. *Journal of Personality and Social Psychology* 110(3):458–76. [YW]

Filippetti, M. L., Lloyd-Fox, S., Longo, M. R., Farroni, T. & Johnson, M. H. (2015) Neural mechanisms of body awareness in infants. *Cerebral Cortex* 25:3779–87. [AP]

Finlay, B. L. & Darlington, R. B. (1995) Linked regularities in the development and evolution of mammalian brains. *Science* 268:1578–84. [aMLA]

Finlay, B. L., Darlington, R. B. & Nicastro, N. (2001) Developmental structure in brain evolution. *Behavioral and Brain Sciences* 24:263–307. [aMLA]

Fiske, S. T., Cuddy, A. & Glick, P. (2007) Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences* 11:77–83. [YW]

Flege, J. E., Yeni-Komshian, G. H. & Liu, S. (1999) Age constraints on second-language acquisition. *Journal of Memory and Language* 41(1):78–104. [DD'S]

Fotopoulou, A. (2014) Time to get rid of the "modular" in neuropsychology: A unified theory of anosognosia as aberrant predictive coding. *Journal of Neuropsychology* 8:1–19. doi: 10.1111/jnp.12010. [PBB]

Fox, P. T., Laird, A. R., Fox, S. P., Fox, P. M., Uecker, A. M., Crank, M., Koenig, S. F. & Lancaster, J. (2005) BrainMap taxonomy of experimental design: Description and evaluation. *Human Brain Mapping* 25(1):185–98. [DMK]

Fox, P. T., Lancaster, J. L., Laird, A. R. & Eickhoff, S. B. (2014) Meta-analysis in human neuroimaging: Computational modeling of large-scale databases. *Annual Review of Neuroscience* 37:409–34. [DMK]

Frankenhuis, W. E., Panchanathan, K. & Barrett, H. C. (2013) Bridging developmental systems theory and evolutionary psychology using dynamic optimization. *Developmental Science* 16:584–98. doi: 10.1111/desc.12053. [PBB]

Friston, K. (2010) The free-energy principle: A unified brain theory? *Nature Reviews: Neuroscience* 11:127–38. doi: 10.1038/nrn2787. [PBB]

Friston, K., Daunizeau, J. & Kiebel, S. (2009) Reinforcement learning or active inference? *PLoS One* 4(7):e6421. doi: 10.1371/journal.pone.0006421. [PBB]

Friston, K., Daunizeau, J., Kilner, J. & Kiebel, S. J. (2010) Action and behavior: A free-energy formulation. *Biological Cybernetics* 102:227–60. [GP]

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., FitzGerald, T. & Pezzulo, G. (2015) Active inference and epistemic value. *Cognitive Neuroscience* 6(4):187–214. [GP]

Friston, K. J. (2005) A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 360:815–36. doi: 10.1098/rstb.2005.1622. [PBB]

Friston, K. J., Daunizeau, J., Kilner, J. & Kiebel, S. J. (2010) Action and behavior: A free-energy formulation. *Biological Cybernetics* 102:227–60. doi: 10.1007/s00422-010-0364-z. [PBB, GP]

Fuster, J. M. (1988) *Prefrontal cortex.* Birkhäuser. [AP]

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.* Marsh, Capen & Lyon. [DMK]

Gallagher, S. (2005) *How the body shapes the mind.* Oxford University Press. [GP]

Gallagher, S. & Bower, M. (2014) Making enactivism even more embodied. *AVANT. Pismo Awangardy Filozoficzno-Naukowej* 2:232–47. [rMLA]

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996) Action recognition in the premotor cortex. *Brain* 119:593–609. [AP]

Ganis, G., Thompson, W. L. & Kosslyn, S. M. (2004) Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research* 20:226–41. [AP]

Gauthier, I., Skudlarski, P., Gore, J. C. & Anderson, A. W. (2000) Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience* 3:191–97. [AP]

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A.-M. & Cruciat, C.-M. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415:141–47. [LPes]

Geary, D. C. & Bjorklund, D. F. (2000) Evolutionary developmental psychology. *Child Development* 71:57–65. doi: 10.1111/1467-8624.00118. [PBB]

Georgopoulos, A., Kettner, R. & Schwartz, A. (1988) Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population. *Journal of Neuroscience* 8(8):2928–37. [CP]

Gerstner, W., Kreiter, A. K., Markram, H. & Herz, A. V. (1997) Neural codes: Firing rates and beyond. *Proceedings of the National Academy of Sciences* 94:12740–41. [AP]

Gibson, J. J. (1966) *The senses considered as perceptual systems.* Houghton-Mifflin. [aMLA]

Gibson, J. J. (1977) The theory of affordances. In: *Perceiving, acting, and knowing: Toward an ecological psychology,* ed. R. Shaw & J. Bransford, pp. 62–82. Erlbaum. [GP]

Gibson, J. J. (1979) *The ecological approach to visual perception.* Erlbaum. [aMLA]

Giedd, J. N. & Rapoport, J. L. (2010) Structural MRI of pediatric brain development: What have we learned and where are we going? *Neuron* 67:728–34. [DD'S]

Gilbert, D. T., Gill, M. J. & Wilson, T. D. (2002) The future is now: Temporal correction in affective forecasting. *Organizational Behavior and Human Decision Processes* 88(1):430–44. [YW]

Giza, C. C. & Prins, M. L. (2006) Is being plastic fantastic? Mechanisms of altered plasticity after developmental Traumatic Brain Injury. *Developmental Neuroscience* 28:364–79. [DD'S]

Gładziejewski, P. (2016) Predictive coding and representationalism. *Synthese* 193(2):559–82. doi: 10.1007/s11229-015-0762-9. [rMLA]

Glenberg, A. (1997) What memory is for. *Behavioral and Brain Sciences* 20:1–55. [GP]

Glenberg, A. M. & Kaschak, M. P. (2002) Grounding language in action. *Psychonomic Bulletin and Review* 9(3):558–65. [arMLA]

Glenberg, A. M., Sato, M. & Cattaneo, L. (2008) Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology* 18(7):R290–91. [aMLA]

Glymour, C. (2001) *The mind's arrows: Bayes nets and graphical causal models in psychology.* MIT Press. [JBM]

Glymour, C. & Hanson, C. (forthcoming) Reverse inference in neuropsychology. *The British Journal for the Philosophy of Science.* [JBM]

Gobbelé, R., Schürmann, M., Forss, N., Juottonen, K., Buchner, H. & Hari, R. (2003) Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. *NeuroImage* 20:503–11. [AP]

Gobet, F. (2000a) Retrieval structures and schemata: A brief reply to Ericsson and Kintsch. *British Journal of Psychology* 91:591–94. [AG]

Gobet, F. (2000b) Some shortcomings of long-term working memory. *British Journal of Psychology* 91:551–70. [AG]

Gobet, F., Lane, P. C. R., Croker, S. C. H., Cheng, P., Jones, G., Oliver, I. & Pine, J. M. (2001) Chunking mechanisms in human learning. *Trends in Cognitive Science* 5:236–43. [AG]

Gobet, F. & Simon, H. A. (1996) Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology* 31:1–40. [AG]

Goddings, A.-L. & Giedd, J. N. (2014) Structural brain development during childhood and adolescence. In: *The cognitive neurosciences, fifth edition,* ed. M. S. Gazzaniga & G. R. Mangun, pp. 15–22. MIT Press. [DD'S]

Gödel, K. (2001) On formally undecidable propositions of *Principia mathematica* and related systems I. In: *Collected works, volume I, publications 1929–1936,* ed. S. Feferman, J. W. Dawson, Jr. & S. C. Kleene, pp. 145–95. Oxford University Press. [LPer]

Gold, K., Havasi, C., Anderson, M. & Arnold, K. (2011) Comparing matrix decomposition methods for meta-analysis and reconstruction of cognitive neuroscience results. *Proceedings of the Twenty-Fourth Annual Conference of the Florida Artificial Intelligence Research Society* 24:21–26. [aMLA]

Goldstein, E. (2009) *Sensation and perception.* Cengage Learning. [AP]

Gould, S. J. (1996) *The mismeasure of man.* WW Norton. [JBM]

Grafton, S. T., Fadiga, L., Arbib, M. A. & Rizzolatti, G. (1997) Premotor cortex activation during observation and naming of familiar tools. *NeuroImage* 6:231–36. [AP]

Gregory, R. L. (1980) Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 290:181–97. doi: 10.1098/rstb.1980.0090. [PBB]

Grush, R. (2004) The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences* 27:377–96. [GP]

Guerreiro, M. J., Putzar, L. & Röder, B. (2015) The effect of early visual deprivation on the neural bases of multisensory processing. *Brain* 138:1499–504. [AP]

Guida, A., Gobet, F. & Nicolas, S. (2013) Functional cerebral reorganization: A signature of expertise? Reexamining Guida, Gobet, Tardieu, and Nicolas' (2012) two-stage framework. *Frontiers in Human Neuroscience* 7:590. [AG]

Guida, A., Gobet, F., Tardieu, H. & Nicolas, S. (2012) How chunks, long-term working memory and templates offer a cognitive explanation for neuroimaging data on expertise acquisition: A two-stage framework. *Brain and Cognition* 79:221–44. [AG]

Guida, A. & Lavielle-Guida, M. (2014) 2011 space odyssey: Spatialization as a mechanism to code order allows a close encounter between memory expertise and classic immediate memory studies. *Frontiers in Psychology* 5:573. [AG]

Guida, A., Leroux, A., Lavielle-Guida, M. & Noël, Y. (2015) A SPoARC in the dark: Spatialization in verbal immediate memory. *Cognitive Science*. doi: 10.1111/cogs.12316.   [AG]

Gupta, A. S., van der Meer, M. A., Touretzky, D. S. & Redish, A. D. (2010) Hippocampal replay is not a simple function of experience. *Neuron* 65:695–705.   [GP]

Hagen, M. C., Franzén, O., McGlone, F., Essick, G., Dancer, C. & Pardo, J. V. (2002) Tactile motion activates the human middle temporal/V5 (MT/V5) complex. *European Journal of Neuroscience* 16:957–64.   [AP]

Haxby, J. V, Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L. & Pietrini, P. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293(5539):2425–30. doi: 10.1126/science.1063736.   [CP]

Hensch, T. K. (2005) Critical period plasticity in local cortical circuits. *Nature Reviews Neuroscience* 6:877–88.   [DD'S]

Hermans, E. J., van Marle, H. J., Ossewaarde, L., Henckens, M. J., Qin, S., van Kesteren, M. T., Schoots, V. C., Cousijn, H., Rijpkema, M., Oostenveld, R. & Fernández, G. (2011) Stress-related noradrenergic activity prompts large-scale neural network reconfiguration. *Science* 334(6059):1151–53.   [aMLA]

Hernandez, A. E. & Li, P. (2007) Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin* 133(4):638.   [DD'S]

Hilgetag, C. C., O'Neill, M. A. & Young, M. P. (1996) Indeterminate organization of the visual system. *Science* 271:776–77.   [LPes]

Hohwy, J. (2013) *The predictive mind*. Oxford University Press.   [PBB, rMLA]

Holtmaat, A. & Svoboda, K. (2009) Experience dependent structural synaptic plasticity in the mammalian brain. *Nature Reviews Neuroscience* 10:647–58.   [aMLA]

Horwitz, B. & Braun, A. R. (2004) Brain network interactions in auditory, visual and linguistic processing. *Brain and Language* 89:377–84.   [AP]

Huang, J. Y., Sedlovskaya, A., Ackerman, J. M. & Bargh, J. A. (2011) Immunizing against prejudice: Effects of disease protection on attitudes toward out-groups. *Psychological Science* 22:1550–56.   [YW]

Huttenlocher, P. R. (1990) Morphometric study of human cerebral cortex development. *Neuropsychologia* 28(6):517–27.   [DD'S]

Huttenlocher, P. R. (1994) Synaptogenesis, synapse elimination, and neural plasticity in human cerebral cortex. In: *Threats to optimal development: Integrating biological, psychological, and social risk factors. The Minnesota symposium on child development, vol. 27*, ed. C. A. Nelson, pp. 35–54. Erlbaum.   [DD'S]

Iachini, T., Ruggiero, G. & Ruotolo, F. (2014) Does blindness affect egocentric and allocentric frames of reference in small and large scale spaces? *Behavioural Brain Research* 273:73–81.   [AP]

IJzerman, H., Gallucci, M., Pouw, W. T., Weissgerber, S. C., Van Doesum, N. J. & Williams, K. D. (2012) Cold-blooded loneliness: Social exclusion leads to lower skin temperatures. *Acta Psychologica* 140:283–88.   [YW]

Inagaki, T. K. & Eisenberger, N. I. (2013) Shared neural mechanisms underlying social warmth and physical warmth. *Psychological Science* 24:2272–80.   [YW]

James, W. (1890) *The principles of psychology*. Henry Holt.   [aMLA]

Jeannerod, M. (2006) *Motor cognition*. Oxford University Press.   [GP]

Jiang, W., Wallace, M. T., Jiang, H., Vaughan, J. W. & Stein, B. E. (2001) Two cortical areas mediate multisensory integration in superior colliculus neurons. *Journal of Neurophysiology* 85:506–22.   [AP]

Johnson, A. & Redish, A. D. (2007) Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience* 27:12176–89.   [GP]

Johnson, J. S. & Newport, E. L. (1989) Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21(1):60–99.   [DD'S]

Johnson, M. H. (2001) Functional brain development in humans. *Nature Reviews Neuroscience* 2(7):475–83.   [aMLA, DD'S]

Johnson, M. H. (2011) Interactive specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience* 1(1):7–21.   [aMLA, DD'S]

Kang, Y., Williams, L., Clark, M., Gray, J. & Bargh, J. A. (2011) Physical temperature effects on trust behavior: The role of insula. *Social Cognitive and Affective Neuroscience* 6:507–15.   [YW]

Kanwisher, N. (2010) Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences, USA* 107(25):11163–70.   [aMLA]

Karmiloff-Smith, A. (1992) *Beyond modularity: A developmental perspective on cognitive science*. MIT Press.   [PBB]

Karmiloff-Smith, A. (1998) Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences* 2(10):389–98.   [DD'S]

Karmiloff-Smith, A. (2015) An alternative approach to domain-general or domain-specific frameworks for theorizing about evolution and ontogenesis. *AIMS Neuroscience* 2(2):91–104.   [DD'S]

Kaschak, M. P. & Glenberg, A. M. (2000) Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language* 43(3):508–29.   [rMLA]

Kaschak, M. P. & Maner, J. K. (2009) Embodiment, evolution, and social cognition: An integrative framework. *European Journal of Social Psychology* 39(7):1236–44.   [YW]

Kauffman, T., Hamilton, R., Keenan, J. P., Warde, A. & Pascual-Leone, A. (2000) The role of visual cortex in tactile Braille reading: The early blind, the sighted, and the blindfolded. *Annals of Neurology* 48:418–19.   [AP]

Kawachi, Y., Grove, P. M. & Sakurai, K. (2014) A single auditory tone alters the perception of multiple visual events. *Journal of Vision* 14:1–13.   [AP]

Kay, A. C., Gaucher, D., Napier, J. L., Callan, M. J. & Laurin, K. (2008) God and the government: Testing a compensatory control mechanism for the support of external systems. *Journal of Personality and Social Psychology* 95:18–35.   [YW]

Kayser, C., Petkov, C. I. & Logothetis, N. K. (2008) Visual modulation of neurons in auditory cortex. *Cerebral Cortex* 18:1560–74.   [AP]

Kelso, J. A. S. (1995) *Dynamic patterns: The self-organization of brain and behavior*. MIT Press.   [PBB, DD'S]

Kenrick, D. T. (2001) Evolutionary psychology, cognitive science, and dynamical systems: Building an integrative paradigm. *Current Directions in Psychological Science* 10:13–17. doi: 10.1111/1467-8721.00104.   [PBB]

Kenrick, D. T., Maner, J. K., Butner, J., Li, N. P., Becker, D. V. & Schaller, M. (2002) Dynamical evolutionary psychology: Mapping the domains of the new interactionist paradigm. *Personality and Social Psychology Review* 6:347–56. doi: 10.1207/S15327957PSPR0604_09.   [PBB, YW]

Kim, M., Ducros, M., Carlson, T., Ronen, I., He, S., Ugurbil, K. & Kim, D.-S. (2006) Anatomical correlates of the functional organization in the human occipito-temporal cortex. *Magnetic Resonance Imaging* 24:583–90.   [AP]

Kinnison, J., Padmala, S., Choi, J. M. & Pessoa, L. (2012) Network analysis reveals increased integration during emotional and motivational processing. *Journal of Neuroscience* 32:8361–72.   [LPes]

Kirsh, D. (2010) *Thinking with external representations*. AI Soc.   [GP]

Kitano, H. (2004) Biological robustness. *Nature Reviews Genetics* 5:826–37.   [DD'S]

Klasios, J. (2014) Our computational nature: Comment on Barrett et al. *Frontiers in Psychology* 5:1348. doi: 10.3389/fpsyg.2014.01348.   [PBB]

Klein, C. (2012) Cognitive ontology and region-versus network-oriented analyses. *Philosophy of Science* 79:952–60.   [JBM]

Knops, A., Thirion, B., Hubbard, E. M., Michel, V. & Dehaene, S. (2009) Recruitment of an area involved in eye movements during mental arithmetic. *Science* 324(5934):1583–85. doi: 10.1126/science.1171599.   [CP]

Kriegeskorte, N., Mur, M. & Bandettini, P. (2008a) Representational similarity analysis: Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience* 2:4. doi: 10.3389/neuro.06.004.2008.   [CP]

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K. & Bandettini, P. A. (2008b) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60(6):1126–41. doi: 10.1016/j.neuron.2008.10.043.   [CP]

Krishna, A. (2012) An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*, 22(3):332–51.   [YW]

Kuhl, P. K. (2004) Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* 5:831–43.   [DD'S]

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. & Nelson, T. (2008) Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences* 363(1493):979–1000.   [DD'S]

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S. & Iverson, P. (2006) Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental science* 9(2):F13–F21.   [DD'S]

Kuhl, P. K., Tsao, F.-M. & Liu, H.-M. (2003) Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the USA* 100 (15):9096–191.   [DD'S]

Laird, A. R., Lancaster, J. L. & Fox, P. T. (2005) BrainMap: The social evolution of a human brain mapping database. *Neuroinformatics* 3(1):65–78.   [DMK, JMS]

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A. & Schroeder, C. E. (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53:279–92.   [AP]

Lakoff, G. & Johnson, M. (1980) *Metaphors we live by*. University of Chicago Press.   [YW]

Landau, M. J., Meier, B. P. & Keefer, L. A. (2010) A metaphor-enriched social cognition. *Psychological Bulletin* 136:1045–67.   [YW]

Landy, D. & Goldstone, R. L. (2009) How much of symbolic manipulation is just symbol pushing? In: *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society, Amsterdam, Netherlands, July 29–August 1, 2009*, ed. N. A. Taatgen & H. van Rijn, pp. 1072–77. Cognitive Science Society.   [aMLA]

Lawson, R., Boylan, A. & Edwards, L. (2014) Where you look can influence haptic object recognition. *Attention, Perception, and Psychophysics* 76:559–74.   [AP]

Lenartowicz, A., Kalar, D. J., Congdon, E. & Poldrack, R. A. (2010) Towards an ontology of cognitive control. *Topics in Cognitive Science* 2:678–92. [JBM]

Lewkowicz, D. J. & Ghazanfar, A. A. (2009) The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences* 13(11):470–78. [DD'S]

Liang, M., Mouraux, A., Hu, L. & Iannetti, G. D. (2013) Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. *Nature Communications* 4:1979. [AP]

Lickliter, R. & Honeycutt, H. (2003) Developmental dynamics: Toward a biologically plausible evolutionary psychology. *Psychological Bulletin* 129:819–35. doi: 10.1037/0033-2909.129.6.819. [PBB]

Lin, L., Osan, R. & Tsien, J. Z. (2006) Organizing principles of real-time memory encoding: Neural clique assemblies and universal neural codes. *Trends in Neurosciences* 29(1):48–57. doi: 10.1016/j.tins.2005.11.004. [CP]

Lindquist, K. A. & Barrett, L. F. (2012) A functional architecture of the human brain: Emerging insights from the science of emotion. *Trends in Cognitive Sciences* 16:533–40. [aMLA]

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E. & Barrett, L. F. (2012) The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences* 35(3):121–43. [aMLA]

Lisman, J. (2015) The challenge of understanding the brain: Where we stand in 2015. *Neuron* 86:864–82. doi: 10.1016/j.neuron.2015.03.032. [GP]

Lisman, J. E., Coyle, J. T., Green, R. W., Javitt, D. C., Benes, F. M., Heckers, S. & Grace, A. A. (2008) Circuit-based framework for understanding neurotransmitter and risk gene interactions in schizophrenia. *Trends in Neuroscience* 31 (5):234–42. [DD'S]

Liu, J., Wang, Z., Feng, L., Li, J., Tian, J. & Lee, K. (2015) Neural trade-offs between recognizing and categorizing own- and other-race faces. *Cerebral Cortex* 25 (8):2191–203. [DD'S]

LoBue, V. & Rakison, D. H. (2013) What we fear most: A developmental advantage for threat-relevant stimuli. *Developmental Review* 33:285–303. doi: 10.1016/j. dr.2013.07.005. [PBB]

Longo, M. R., Azañón, E. & Haggard, P. (2010) More than skin deep: Body representation beyond primary somatosensory cortex. *Neuropsychologia* 48:655–68. [AP]

Lupyan, G. (2015) Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology* 6(4):547–69. doi: 10.1007/s13164-015-0253-4. [PBB]

Machery, E. (2014) In defense of reverse inference. *British Journal for Philosophy of Science* 65:251–67. [JBM]

Maguire, E. A., Valentine, E. R., Wilding, J. M. & Kapur, N. (2003) Routes to remembering: The brains behind superior memory. *Nature Neuroscience* 6:90–95. [AG]

Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503(7474):78–84. [JMS]

Marcus, G. F. (2006) Cognitive architecture and descent with modification. *Cognition* 101(2):443–65. doi: 10.1016/j.cognition.2006.04.009. [CP]

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., Thomas, M. S. C. & Westermann, G. (2007) *Neuroconstructivism, vol. I: How the brain constructs cognition.* Oxford University Press. [DD'S]

Marr, D. (1982) *Vision.* W. H. Freeman. [aMLA]

Marsh, K. L., Richardson, M. J. & Schmidt, R. C. (2009) Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science* 1(2):320–39. [rMLA]

Masataka, N., Hayakawa, S. & Kawai, N. (2010) Human young children as well as adults demonstrate "superior" rapid snake detection when typical striking posture is displayed by the snake. *PLoS One* 5:e15122. doi: 10.1371/journal. pone.0015122. [PBB]

Mathy, F. & Feldman, J. (2012) What's magic about magic numbers. Chunking and data compression in short-term memory. *Cognition* 122:346–62. [AG]

Matsuhashi, M., Ikeda, A., Ohara, S., Matsumoto, R., Yamamoto, J., Takayama, M., Satowa, T., Begum, T., Usui, K., Nagamine, T., Mikuni, N., Takahashi, J., Miyamoto, S., Fukuyama, H. & Shibasaki, H. (2004) Multisensory convergence at human temporo-parietal junction – epicortical recording of evoked responses. *Clinical Neurophysiology* 115:1145–60. [AP]

Maye, A. & Engel, A. K. (2011) A discrete computational model of sensorimotor contingencies for object perception and control of behavior. *IEEE International Conference on Robotics and Automation (ICRA 2011), Shanghai, China, May 9–13, 2011,* pp. 3810–15. IEEE. doi: 10.1109/ICRA.2011.5979919. [GP]

McCaffrey, J. (2015) The brain's heterogeneous functional landscape. *Philosophy of Science* 82:1010–22. [JBM]

McGhee, G. (2011) *Convergent evolution: Limited forms most beautiful.* MIT Press. [MS]

McIntosh, A. R. (2000) Towards a network theory of cognition. *Neural Networks* 13:861–70. doi: 10.1016/S0893-6080(00)00059-9. [PBB]

Meier, B. P., Moeller, S. K., Riemer-Peltz, M. & Robinson, M. D. (2012a) Sweet taste preferences and experiences predict prosocial inferences, personalities, and behaviors. *Journal of Personality and Social Psychology* 102:163–74. [YW]

Meier, B. P., Schnall, S., Schwarz, N. & Bargh, J. (2012b) Embodiment in social psychology. *Topics in Cognitive Science* 4:705–16. [YW]

Menon, V. & Uddin, L. Q. (2010) Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function* 214:655–67. [LPes]

Merabet, L. B., Hamilton, R., Schlaug, G., Swisher, J. D., Kiriakapoulos, E. T., Pitskel, N. B., Kauffman, T. & Pascual-Leone, A. (2008) Rapid and reversible recruitment of early visual cortex for touch. *PLoS One* 3(8):e3046: 1–12. [aMLA, AG]

Mesulam, M.-M. (1990) Large-scale neurocognitive networks and distributed processing for attention, language and memory. *Annals of Neurology* 28:597–13. doi: 10.1002/ana.410280502. [PBB]

Mesulam, M. M. (1998) From sensation to cognition. *Brain* 121:1013–52. [LPes]

Meunier, D., Fonlupt, P., Saive, A.-L., Plailly, J., Ravel, N. & Royet, J.-P. (2014) Modular structure of functional networks in olfactory memory. *NeuroImage* 95C:264–75. doi: 10.1016/j.neuroimage.2014.03.041. [MLS]

Meunier, D., Lambiotte, R. & Bullmore, E. T. (2010) Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience* 4:200. doi: 10.3389/fnins.2010.00200. [PBB]

Meunier, D., Lambiotte, R., Fornito, A., Ersche K. D. & Bullmore E. T. (2009) Hierarchical modularity in human functional brain networks. *Frontiers in Neuroinformatics* 3:37. doi: 10.3389/neuro.11.037.2009. [MLS]

Miller, E. K. & Cohen, J. D. (2001) An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* 24:167–202. [aMLA]

Mosch, S. C., Max, J. E. & Tranel, D. (2005) A matched lesion analysis of childhood versus adult-onset brain injury due to unilateral stroke: Another perspective on neural plasticity and recovery of social functioning. *Cognitive Behavioral Neurology* 18(1):5–17. [MLS]

Moussa, M. N., Vechlekar, C. D., Burdette, J. H., Steen, M. R., Hugenschmidt, C. E. & Laurienti, P. J. (2011) Changes in cognitive state alter human functional brain networks. *Frontiers in Human Neuroscience* 5:83. doi: 10.3389/ fnhum.2011.00083. [MLS]

Moussa, M. N., Wesley, M. J., Porrino, L. J., Hayasaka, S., Bechara, A., Burdette, J. H. & Laurienti, P. J. (2014) Age-related differences in advantageous decision making are associated with distinct differences in functional community structure. *Brain Connectivity* 4:193–202. doi: 10.1089/brain.2013.0184. [MLS]

Murray, D. J., Ellis, R. R., Bandomir, C. A. & Ross, H. E. (1999) Charpentier (1891) on the size-weight illusion. *Perception and Psychophysics* 61:1681–85. [aMLA]

Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., Schroeder, C. E. & Foxe, J. J. (2005) Grabbing your ear: Rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebral Cortex* 15:963–74. [AP]

Newman, M. (2010) *Networks: An introduction.* Oxford University Press. [LPes]

Newman, M. E. J. (2006) Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America,* 103:8577–82. doi: 10.1073/pnas.0601602103. [MLS]

Noë, A. (2004) *Action in perception.* MIT Press. [aMLA]

Öngür, D. & Price, J. L. (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex* 10:206–19. [AP]

O'Regan, K. & Noe, A. (2001) A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24(5):939–73. [GP]

Owen, A. M., McMillan, K. M., Laird, A. R. & Bullmore, E. (2005) N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping* 25(1):46–59. [DMK]

Oyserman, D. (2011) Culture as situated cognition: Cultural mindsets, cultural fluency, and meaning making. *European Review of Social Psychology* 22 (1):164–214. [YW]

Padmala, S. & Pessoa, L. (2011) Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience* 23:3419–32. [LPes]

Palla, G., Derenyi, I., Farkas, I. & Vicsek, T. (2005) Uncovering the overlapping community structure of complex networks in nature and society. *Nature* 435:814–18. [LPes]

Parkinson, C., Liu, S. & Wheatley, T. (2014) A common cortical metric for spatial, temporal, and social distance. *Journal of Neuroscience* 34(5):1979–87. doi: 10.1523/JNEUROSCI.2159-13.2014. [CP]

Parkinson, C. & Wheatley, T. (2013) Old cortex, new contexts: Re-purposing spatial perception for social cognition. *Frontiers in Human Neuroscience* 7:645. doi: 10.3389/fnhum.2013.00645. [CP]

Parkinson, C. & Wheatley, T. (2015) The repurposed social brain. *Trends in Cognitive Sciences* 19(3):133–41. doi: 10.1016/j.tics.2015.01.003. [CP]

Pascalis, O., Loevenbruck, H., Quinn, P. C., Kandel, S., Tanaka, J. W. & Lee, K. (2014) On the links among face processing, language processing, and narrowing during development. *Child Development Perspectives* 8(2):65–70. [DD'S]

Pascual-Leone, A. & Hamilton, R. (2001) The metamodal organization of the brain. *Progress in Brain Research* 134:427–45. [AP]

Pasqualotto, A., Finucane, C. M. & Newell, F. N. (2013a) Ambient visual information confers a context-specific, long-term benefit on memory for haptic scenes. *Cognition* 128:363–79. [AP]

Pasqualotto, A., Lam, J. S. Y. & Proulx, M. J. (2013b) Congenital blindness improves semantic and episodic memory. *Behavioural Brain Research* 244:162–65. [AP]

Passingham, R. E., Stephan, K. E. & Kötter, R. (2002) The anatomical basis of functional localization in the cortex. *Nature Reviews Neuroscience* 3:606–16. [JBM]

Peelen, M. V. & Downing, P. E. (2007) Using multi-voxel pattern analysis of fMRI data to interpret overlapping functional activations. *Trends in Cognitive Sciences* 11(1):4–5. doi: 10.1016/j.tics.2006.10.009. [CP]

Penner-Wilger, M. & Anderson, M. L. (2013) The relation between finger gnosis and mathematical ability: Why redeployment of neural circuits best explains the finding. *Frontiers in Psychology* 4. Available at: http://dx.doi.org/10.3389/fpsyg.2013.00877. [rMLA]

Perlovsky, L. I. (1998) Conundrum of combinatorial complexity. *IEEE Transactions PAMI* 20(6):666–70. [LPer]

Perlovsky, L. I. (2001) *Neural networks and intellect: Using model-based concepts*. Oxford University Press (3rd printing). [LPer]

Perlovsky, L. I. (2006) Toward physics of the mind: Concepts, emotions, consciousness, and symbols. *Physics of Life Reviews* 3(1):22–55. [LPer]

Perlovsky, L. I. (2007a) Symbols: Integrated cognition and language. In: *Semiotics and intelligent systems development*, ed. R. Gudwin & J. Queiroz, pp. 121–51. Idea Group. [LPer]

Perlovsky, L. I. (2007b) The mind vs. logic: Aristotle and Zadeh. Society for mathematics of uncertainty. *Critical Review* 1(1):30–33. [LPer]

Perlovsky, L. I. (2009) "Vague-to-Crisp" neural mechanism of perception. *IEEE Transactions Neural Networks* 20(8):1363–67. [LPer]

Perlovsky, L. I. (2010a) Intersections of mathematical, cognitive, and aesthetic theories of mind. *Psychology of Aesthetics, Creativity, and the Arts* 4(1):11–17. doi: 10.1037/a0018147. [LPer]

Perlovsky, L. I. (2010b) Musical emotions: Functions, origin, evolution. *Physics of Life Reviews* 7(1):2–27. doi: 10.1016/j.plrev.2009.11.001. [LPer]

Perlovsky, L. I. (2013a) A challenge to human evolution – cognitive dissonance. *Frontiers in Psychology* 4:179. doi: 10.3389/fpsyg.2013.00179. Available at: http://www.frontiersin.org/cognitive_science/10.3389/fpsyg.2013.00179/full. [LPer]

Perlovsky, L. I. (2013b) Language and cognition – joint acquisition, dual hierarchy, and emotional prosody. *Frontiers in Behavioral Neuroscience* 7:123. doi: 10.3389/fnbeh.2013.00123. Available at: http://www.frontiersin.org/Behavioral_Neuroscience/10.3389/fnbeh.2013.00123/full. [LPer]

Perlovsky, L. I. (2013c) Learning in brain and machine – complexity, Gödel, Aristotle. *Frontiers in Neurorobotics* 7:23. doi: 10.3389/fnbot.2013.00023. Available at: http://www.frontiersin.org/Neurorobotics/10.3389/fnbot.2013.00023/full. [LPer]

Perlovsky, L. I. (2014) Aesthetic emotions, what are their cognitive functions? *Frontiers in Psychology* 5:98. doi: 10.3389/fpsyg.2014.0009. Available at: http://www.frontiersin.org/Journal/10.3389/fpsyg.2014.00098/full. [LPer]

Perlovsky, L. I. (2015) Origin of music and the embodied cognition. *Frontiers in Psychology* 6:538. Available at: http://dx.doi.org/10.3389/fpsyg.2015.00538. [LPer]

Perlovsky, L. I., Deming R. W. & Ilin, R. (2011) *Emotional cognitive neural algorithms with engineering applications. Dynamic Logic: From vague to crisp*. Springer. [LPer]

Pesenti, M., Zago, L., Crivello, F., Mellet, E., Samson, D., Duroux, B., Seron, X., Mazoyer, B. & Tzourio-Mazoyer, N. (2001) Mental calculation in a prodigy is sustained by right prefrontal and medial temporal areas. *Nature Neuroscience* 4:103–107. [AG]

Pessoa, L. (2009) How do emotion and motivation direct executive function? *Trends in Cognitive Sciences* 13:160–66. [LPes]

Pessoa, L. (2012) Beyond brain regions: Network perspective of cognition-emotion interactions. *Behavioral and Brain Sciences* 35:158–59. [AP]

Pessoa, L. (2013) *The cognitive-emotional brain: From interactions to integration*. MIT Press. [LPes]

Pessoa, L. (2014) Understanding brain networks and brain organization. *Physics of Life Reviews* 11:400–35. [LPes]

Pessoa, L., McKenna, M., Gutierrez, E. & Ungerleider, L. G. (2002) Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences USA* 99:11458–63. [LPes]

Petersen, S. E. & Fiez, J. A. (1993) The processing of single words studied with positron emission tomography. *Annual Review of Neuroscience* 16:509–30. [DMK]

Petersen, S. E., van Mier, H., Fiez, J. A. & Raichle, M. E. (1998) The effects of practice on the functional anatomy of task-performance. *Proceedings of the National Academy of Sciences of the United States of America* 95:853–60. [AG]

Petersson, K. M., Elfgren, C. & Ingvar, M. (1997) A dynamic role of the medial temporal lobe during retrieval of declarative memory in man. *NeuroImage* 6:1–11. [AG]

Pezzulo, G. (2011) Grounding procedural and declarative knowledge in sensorimotor anticipation. *Mind and Language* 26:78–114. [GP]

Pezzulo, G. (2012) An active inference view of cognitive control. *Frontiers in Psychology* 3:478. doi: 10.3389/fpsyg.2012.00478. [GP]

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K. & Spivey, M. (2011) The mechanics of embodiment: A dialogue on embodiment and computational modeling. *Frontiers in Cognition* 2:1–21. [GP]

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K. & Spivey, M. J. (2013) Computational grounded cognition: A new alliance between grounded cognition and computational modeling. *Frontiers in Psychology* 3:612. doi: 10.3389/fpsyg.2012.00612. [GP]

Pezzulo, G. & Castelfranchi, C. (2009) Thinking as the control of imagination: A conceptual framework for goal-directed systems. *Psychological Research* 73:559–77. [GP]

Pezzulo, G., Rigoli, F. & Friston, K. (2015) Active inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology* 134:17–35. doi: 10.1016/j.pneurobio.2015.09.001. [GP]

Pezzulo, G., van der Meer, M. A. A., Lansink, C. S. & Pennartz, C. M. A. (2014) Internally generated sequences in learning and executing goal-directed behavior. *Trends in Cognitive Sciences* 18:647–57. doi: 10.1016/j.tics.2014.06.011. [GP]

Pfaff, D. W. (2002) *Hormones, brain and behavior, five-volume set*. Academic Press. [aMLA]

Pfeifer, J. H. & Allen, N. B. (2012) Arrested development? Reconsidering dual-systems models of brain function in adolescence and disorders. *Trends in Cognitive Sciences* 16(6):322–29. Available at: http://doi.org/10.1016/j.tics.2012.04.011. [PBB]

Pfeifer, R. & Scheier, C. (1999) *Understanding intelligence*. MIT Press. [GP]

Pfeiffer, B. E. & Foster, D. J. (2013) Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* 497:74–79. doi: 10.1038/nature12112. [GP]

Pinker, S. (1997) *How the mind works*. Norton. [PBB]

Platt, M. L. (2002) Neural correlates of decisions. *Current Opinion in Neurobiology* 12:141–48. [aMLA]

Ploeger, A., Van Der Maas, H. J. L. & Raijmakers, M. E. J. (2008) Is evolutionary psychology a metatheory for psychology? A discussion of four major issues in psychology from an evolutionary developmental perspective. *Psychological Inquiry* 19:1–18. doi: 10.1080/10478400701774006. [PBB]

Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., Veraart, C. & De Volder, A. G. (2006) Auditory motion perception activates visual motion areas in early blind subjects. *NeuroImage* 31:279–85. [AP]

Poldrack, R. A. (2006) Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences* 10:59–63. [JBM]

Poldrack, R. A. (2010) Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on Psychological Science* 5(6):753–61. [aMLA, JBM]

Poldrack, R. A., Desmond, J. E., Glover, G. H. & Gabrieli, J. D. (1998) The neural basis of visual skill learning: An fMRI study of mirror reading. *Cerebral Cortex* 8:1–10. [AG]

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* 20:1364–72. [aMLA, JBM]

Poldrack, R. A., Kittur, A., Kalar, D. & Miller, E. (2011) The cognitive atlas: Toward a knowledge foundation for cognitive neuroscience. *Frontiers in Neuroinformatics* 5:17. [JMS]

Posner, M. I. & Rothbart, M. K. (2007) Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology* 58:1–23. [AP]

Postle, B. R., Berger, J. S. & D'Esposito, M. (1999) Functional neuroanatomical double dissociation of mnemonic and executive control processes contributing to working memory performance. *Proceedings of the National Academy of Sciences of the United States of America* 96:12959–64. [AG]

Postle, B. R. & D'Esposito, M. (1999) "What"-then-"Where" in visual working memory: An event-related fMRI study. *Journal of Cognitive Neuroscience* 11:585–97. [AG]

Pouget, A., Dayan, P. & Zemel, R. (2000) Information processing with population codes. *Nature Reviews Neuroscience* 1(2):125–32. doi: 10.1038/35039062. [CP]

Powers, W. (1973) *Behavior, the control of perception*. Aldine de Gruyter. [GP]

Price, C. J. (2010) The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences* 1191:62–88. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20392276. [JMS]

Price, C. J., Devlin, J. T., Moore, C. J., Morton, C. & Laird, A. R. (2005) Meta-analyses of object naming: Effect of baseline. *Human Brain Mapping* 25(1):70–82. [DMK]

Price, C. J. & Friston, K. J. (2002) Degeneracy and cognitive anatomy. *Trends in Cognitive Sciences* 6:416–21. doi: 10.1016/S1364-6613(02)01976-9. [PBB]

Price, C. J. & Friston, K. J. (2005) Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology* 22(3–4):262–75. [aMLA, JBM]

Pridmore, B. (2013) *How to be clever*. Lulu Press. [AG]

Proulx, M. J., Ptito, M. & Amedi, A. (2014) Multisensory integration, sensory substitution and visual rehabilitation. *Neuroscience and Biobehavioral Reviews* 41:1–2. [AP]

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A. & Shulman, G. L. (2001) A default mode of brain function. *Proceedings of the National Academy of Sciences, USA* 98:676–82. [aMLA]

Ravassard, P., Kees, A., Willers, B., Ho, D., Aharoni, D., Cushman, J., Aghajan, Z. M. & Mehta, M. R. (2013) Multisensory control of hippocampal spatiotemporal selectivity. *Science* 340:1342–46. [AP]

Ricci, D., Mercuri, E., Barnett, A., Rathbone, R., Cota, F., Haataja, L., Rutherford, M., Dubowitz, L. & Cowan, F. (2008) Cognitive outcome at early school age in term born children with perinatally acquired middle cerebral artery territory infarction. *Stroke* 39:403–10. [DD'S]

Ricciardi, E., Bonino, D., Pellegrini, S. & Pietrini, P. (2014) Mind the blind brain to understand the sighted one! Is there a supramodal cortical functional architecture? *Neuroscience and Biobehavioral Reviews* 41:64–77. [AP]

Richardson, D. C. & Dale, R. (2005) Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science* 29(6):1045–60. [rMLA]

Richardson, M. J., Marsh, K. L. & Baron, R. M. (2007) Judging and actualizing intrapersonal and interpersonal affordances. *Journal of Experimental Psychology: Human Perception and Performance* 33(4):845. [rMLA]

Rigotti, M., Barak, O., Warden, M. R., Wang, X. J., Daw, N. D., Miller, E. K. & Fusi, S. (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497(7451):585–90. [aMLA]

Rissman, J., Gazzaley, A. & D'Esposito, M. (2004) Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage* 23:752–63. [LPes]

Rogers, T. T., Rakison, D. H. & McClelland, J. L. (2004) U-shaped curves in development: A PDP approach. *Journal of Cognition and Development* 5(1):137–45. [DD'S]

Rolls, E. T. & Tovee, M. J. (1995) Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology* 73:713–26. [AP]

Rosenblueth, A., Wiener, N. & Bigelow, J. (1943) Behavior, purpose and teleology. *Philosophy of Science* 10(1):18–24. [GP]

Roux, F. E., Boetto, S., Sacko, O., Chollet, F. & Trémoulet, M. (2003) Writing, calculating, and finger recognition in the region of the angular gyrus: A cortical stimulation study of Gerstmann syndrome. *Journal of Neurosurgery* 99(4):716–27. [aMLA]

Rubinov, M. & Sporns, O. (2010) Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* 52:1059–69. doi: 10.1016/j.neuroimage.2009.10.003. [MLS]

Rusconi, E., Walsh, V. & Butterworth, B. (2005) Dexterity with numbers: rTMS over left angular gyrus disrupts finger gnosis and number processing. *Neuropsychologia* 43(11):1609–24. [aMLA]

Sadato, N., Pascual-Leone, A., Grafman, J., Ibañez, V., Deiber, M.P., Dold, G. & Hallett, M. (1996) Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* 380:526–28. [AP]

Saenz, M., Lewis, L. B., Huth, A. G., Fine, I. & Koch, C. (2008) Visual motion area MT+/V5 responds to auditory motion in human sight-recovery subjects. *Journal of Neuroscience* 28:5141–48. [AP]

Sakai, K. & Passingham, R. E. (2003) Prefrontal interactions reflect future task operations. *Nature Neuroscience* 6(1):75–81. [JMS]

Sales-Pardo, M., Guimera, R., Moreira, A. A. & Amaral, L. A. N. (2007) Extracting the hierarchical organization of complex systems. *Proceedings of the National Academy of Sciences of the United States of America* 104(39):15224–9. doi: 10.1073/pnas.0703740104. [MLS]

Sathian, K. & Zangaladze, A. (2002) Feeling with the mind's eye: Contribution of visual cortex to tactile perception. *Behavioural Brain Research* 135:127–32. [AP]

Schacter, D. L., Addis, D. R. & Buckner, R. L. (2007) Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience* 8:657–61. [GP]

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N. & Szpunar, K. K. (2012) The future of memory: Remembering, imagining, and the brain. *Neuron* 76:677–94. doi: 10.1016/j.neuron.2012.11.001. [GP]

Schaefer, M., Heinze, H.-J. & Rotte, M. (2014) Rough primes and rough conversations: Evidence for a modality-specific basis to mental metaphors. *Social Cognitive and Affective Neuroscience* 9:1653–59. [YW]

Scott, L. S., Pascalis, O. & Nelson, C. A. (2007) A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science* 16(4):197–201. [DD'S]

Scott, S. H. (2012) The computational and neural basis of voluntary motor control and planning. *Trends in Cognitive Sciences* 16:541–49. doi: 10.1016/j.tics.2012.09.008. [GP]

Scott-Phillips, T. C., Dickins, T. E. & West, S. A. (2011) Evolutionary theory and the ultimate-proximate distinction in the human behavioral sciences. *Perspectives on Psychological Science* 6:38–47. doi: 10.1177/1745691610393528. [PBB]

Sebanz, N., Bekkering, H. & Knoblich, G. (2006) Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences* 10(2):70–76. [arMLA]

Sereno, M. I. & Huang, R. S. (2006) A human parietal face area contains aligned head-centered visual and tactile maps. *Nature Neuroscience* 9:1337–43. [AP]

Serino, A., Canzoneri, E. & Avenanti, A. (2011) Fronto-parietal areas necessary for a multisensory representation of peripersonal space in humans: An rTMS study. *Journal of Cognitive Neuroscience* 23:2956–67. [AP]

Seth, A. K. (2015) The cybernetic Bayesian brain – From interoceptive inference to sensorimotor contingencies. In: *Open MIND*: 35(T), ed. T. Metzinger & J. M. Windt. MIND Group. doi: 10.15502/9783958570108. [GP, rMLA]

Shadmehr, R. & Mussa-Ivaldi, F. A. (2012) *Biological learning and control. How the brain builds representations, predicts events and makes decisions*. MIT Press. [GP]

Shannon, C. E. (1948) A mathematical theory of communication. *Bell System Technical Journal* 27(3):379–423. doi:10.1002/j.1538-7305.1948.tb01338.x. [aMLA]

Shibasaki, M. & Kawai, N. (2009) Rapid detection of snakes by Japanese monkeys (Macaca fuscata): An evolutionarily predisposed visual system. *Journal of Comparative Psychology* 123(2):131–35. doi: 10.1037/a0015095. [PBB]

Shinkareva, S. V., Mason, R. A., Malave, V. L., Wang, W., Mitchell, T. M. & Just, M. A. (2008) Using fMRI brain activation to identify cognitive states associated with perception of tools and dwellings. *PLoS One* 3:e1394. [AP]

Shulman, G. L., Ollinger, J. M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Petersen, S. E. & Corbetta, M. (1999) Areas involved in encoding and applying directional expectations to moving objects. *Journal of Neuroscience* 19:9480–96. [AP]

Siegel, M., Buschman, T. J. & Miller, E. K. (2015) Cortical information flow during flexible sensorimotor decisions. *Science* 348(6241):1352–55. Available at: http://doi.org/10.1126/science.aab0551. [JMS]

Silberstein, M. & Chemero, A. (2011) Dynamics, agency and intentional action. In: *Humana Mente: Journal of Philosophical Studies. Special Issue on Agency: From Embodied Cognition to Free Will. Issue 15*, ed. D. Manetti & S. Z. Caiani, pp. 1–20. Edizioni ETS. [MS]

Silberstein, M. & Chemero, A. (2012) Complexity and extended phenomenological-cognitive systems. In: *Topics in Cognitive Science: Special Issue on the Role of Complex Systems in Cognitive Science, vol. 4*, issue 1, ed. G. Van Orden & D. Stephen, pp. 35–50. Wiley. [MS]

Silberstein, M. & Chemero, A. (2013) Constraints on localization and decomposition as explanatory strategies in the biological sciences. *Philosophy of Science* 80 (5):958–70. [MS]

Simon, S. A., de Araujo, I. E., Gutierrez, R. & Nicolelis, M. A. L. (2006) The neural mechanisms of gustation: A distributed processing code. *Nature Reviews Neuroscience* 7(11):890–901. doi: 10.1038/nrn2006. [CP]

Snow, J. C., Pettypiece, C. E., McAdam, T. D., McLean, A. D., Stroman, P. W., Goodale, M. A. & Culham, J. C. (2011) Bringing the real world into the fMRI scanner: Repetition effects for pictures versus real objects. *Scientific Reports* 1:130. doi: 10.1038/srep00130. [rMLA]

Song, S., Miller, K. D. & Abbott, L. F. (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience* 3(9):919–26. [aMLA]

Spivey, M. (2007) *The continuity of mind*. Oxford University Press. [GP]

Sporns, O. (2011) *Networks of the brain*. MIT Press. [MS, rMLA]

Stanley, M. L., Dagenbach, D., Lyday, R. G., Burdette, J. H. & Laurienti, P. J. (2014) Changes in global and regional modularity associated with increasing working memory load. *Frontiers in Human Neuroscience* 8:954. doi: 10.3389/fnhum.2014.00954. [MLS]

Stanley, M. L., Moussa, M. N., Paolini, B. M., Lyday, R. G., Burdette, J. H. & Laurienti, P. J. (2013) Defining nodes in complex brain networks. *Frontiers in Computational Neuroscience* 7:169. doi: 10.3389/fncom.2013.00169. [MLS]

Stein, B. E., Huneycutt, W. S. & Meredith, M. A. (1988) Neurons and behavior: The same rules of multisensory integration apply. *Brain Research* 448: 355–58. [AP]

Stein, B. E. & Stanford, T. R. (2008) Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience* 9:255–66. [AP]

Stephan, H., Baron, G. & Frahm, H. D. (1988) Comparative size of brain and brain components. In: *Comparative primate biology*, ed. H. D. Steklis & J. Erwin, pp. 1–38. Alan R. Liss. [aMLA]

Sternberg, S. (2011) Modular process in mind and brain. *Cognitive Neuropsychology* 28(3–4):156–208. [DD'S]

Stoianov, I., Genovesio, A. & Pezzulo, G. (2016) Prefrontal goal codes emerge as latent states in probabilistic value learning. *Journal of Cognitive Neuroscience* 28(1):140–57. doi: 10.1162/jocn_a_00886. [GP]

Takagi, N. (2002) The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *The Journal of the Acoustical Society of America* 111(6):2887–96. [DD'S]

Takagi, N. & Mann, V. (1995) The limits of extended naturalistic exposure on the perceptual mastery of English *r* and *l* by adult Japanese learners of English. *Applied Psycholinguistics* 16(4):379–405. [DD'S]

Takahashi, T., Kansaku, K., Wada, M., Shibuya, S. & Kitazawa, S. (2013) Neural correlates of tactile temporal-order judgment in humans: An fMRI study. *Cerebral Cortex* 23:1952–64. [AP]

Telesford, Q. K., Simpson, S. L., Burdette, J. H., Hayasaka, S. & Laurienti, P. J. (2011) The brain as a complex system: Using network science as a tool for understanding the brain. *Brain Connectivity* 1:295–308. doi: 10.1089/brain.2011.0055. [MLS]

Thelen, E. & Smith, L. B. (2006) Dynamic systems theories. In: *Handbook of child psychology, volume 1, theoretical models of human development, sixth edition*, ed. W. Damon, vol. ed. R. M. Lerner, pp. 258–312. John Wiley & Sons. [DD'S]

Thomas, M. S. C. (2003) Limits on plasticity. *Journal of Cognition and Development* 4(1):95–121. [DD'S]

Todd, J. J. & Marois, R. (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428:751–54. [AG]

Tomasello, M. (1999) *The cultural origins of human cognition*. Harvard University Press. [aMLA]

Toussaint, M. (2009) Probabilistic inference as a model of planned behavior. *Künstliche Intelligenz*. 3(9):23–29. [GP]

Uddin, L. Q., Kinnison, J., Pessoa, L. & Anderson, M. L. (2014) Beyond the tripartite cognition–emotion–interoception model of the human insular cortex. *Journal of Cognitive Neuroscience* 26:16–27. [JBM]

Uesaki, M. & Ashida, H. (2015) Optic-flow selective cortical sensory regions associated with self-reported states of vection. *Frontiers in Psychology* 6:775. [AP]

Valencia, M., Pastor, M. A., Fernández-Seara, M. A., Artieda, J., Martinerie, J. & Chavez, M. (2009) Complex modular structure of large-scale brain networks. *Chaos* 19:023119. doi: 10.1063/1.3129783. [MLS]

Vallar, G., Lobel, E., Galati, G., Berthoz, A., Pizzamiglio, L. & Le Bihan, D. (1999) A fronto-parietal system for computing the egocentric spatial frame of reference in humans. *Experimental Brain Research* 124:281–86. [AP]

van Dijck, J. P. & Fias, W. (2011) A working memory account for spatial-numerical associations. *Cognition* 119:114–19. [AG]

Van Dijk, K. R. A., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W. & Buckner, R. L. (2010) Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization. *Journal of Neurophysiology* 103:297–321. [AP]

van Gelder, T. (1995) What might cognition be, if not computation? *The Journal of Philosophy* 92(7):345–81. [aMLA]

Varela, F. J., Thompson, E. T. & Rosch, E. (1992) *The embodied mind: Cognitive science and human experience*. MIT Press. [GP]

Verschure, P., Pennartz, C. M. A. & Pezzulo, G. (2014) The why, what, where, when and how of goal-directed choice: Neuronal and computational principles. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369:20130483. [GP]

Vidal, M. & Barrès, V. (2014) Hearing (rivalling) lips and seeing voices: How audiovisual interactions modulate perceptual stabilization in binocular rivalry. *Frontiers in Human Neuroscience* 8:677. [AP]

Vityaev, E. E., Perlovsky, L. I., Kovalerchuk, B. Y. & Speransky, S. O. (2013) Probabilistic dynamic logic of cognition. Invited Article. *Biologically Inspired Cognitive Architectures* 6:159–68. [LPer]

Vogel, E. K. & Machizawa, M. G. (2004) Neural activity predicts individual differences in visual working memory capacity. *Nature* 428:748–51. [AG]

Werker, J. F. & Tees, R. C. (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7(1):49–63. [DD'S]

Werker, J. F. & Tees, R. C. (2005) Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology* 46(3):233–51. [DD'S]

Wicklund, R. & Gollwitzer, P. G. (1982) *Symbolic self-completion*. Routledge. [YW]

Wiener, N. (1948) *Cybernetics: Or control and communication in the animal and the machine*. MIT Press. [GP]

Wikenheiser, A. M. & Redish, A. D. (2015) Hippocampal theta sequences reflect current goals. *Nature Neuroscience* 18:289–94. [GP]

Williams, L. E. & Bargh, J. A. (2008) Experiencing physical warmth influences interpersonal warmth. *Science* 322:606–7. [YW]

Williams, L. E., Huang, J. Y. & Bargh, J. A. (2009) The scaffolded mind: Higher mental processes are grounded in early experience of the physical world. *European Journal of Social Psychology* 39:1257–67. [YW]

Worthen, J. B. & Hunt, R. R. (2011) *Mnemonology: Mnemonics for the 21st century*. Psychology Press. [AG]

Xiao, Y. J. & Van Bavel, J. J. (2012) See your friends close and your enemies closer: Social identity and identity threat shape the representation of physical distance. *Personality and Social Psychology Bulletin* 38(7):959–72. [aMLA]

Xu, A. J., Schwarz, N. & Wyer, R. S. (2015) Hunger promotes acquisition of nonfood objects. *Proceedings of the National Academy of Sciences* 112(9):2688–92. [YW]

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):665–70. [JMS]

Yates, F. A. (1966) *The art of memory*. University of Chicago Press. [AG]

Yeo, B. T., Krienen, F. M., Chee, M. W. & Buckner, R. L. (2014) Estimates of segregation and overlap of functional connectivity networks in the human cerebral cortex. *NeuroImage* 88:212–27. [LPes]

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zollei, L., Polimeni, J. R., Fischl, B., Liu, H. & Buckner, R. L. (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* 106:1125–65. [AP, LPes]

Yopak, K. E., Lisney, T. J., Darlington, R. B., Collin, S. P., Montgomery, J. C. & Finlay, B. L. (2010) A conserved pattern of brain scaling from sharks to primates. *Proceedings of the National Academy of Sciences USA* 107:12946–51. [aMLA]

Zangaladze, A., Epstein, C. M., Grafton, S. T. & Sathian, K. (1999) Involvement of visual cortex in tactile discrimination of orientation. *Nature* 401:587–90. [AP]

Zhong, C. B. & Leonardelli, G. J. (2008) Cold and lonely: Does social exclusion literally feel cold? *Psychological Science* 19(9):838–42. [YW]

Zhu, Q. & Bingham, G. P. (2011) Human readiness to throw: The size-weight illusion is not an illusion when picking the best objects to throw. *Evolution and Human Behavior* 32(4):288–93. doi: 10.1016/j.evolhumbehav.2010.11.005. [aMLA]a