ARE NEUROREDUCTIONIST EXPLANATIONS OF COGNITION POSSIBLE?

William R. Uttal
Arizona State University

ABSTRACT: Are the available conceptual foundations, the statistical techniques, and the empirical data of cognitive neuroscience sufficiently robust to serve as foundations for an overarching neuroreductionist explanation of the mind-brain relationship? There are two main approaches to answering this question that dominate cognitive neuroscience these days. The first is based on accumulated macroneural measurements such as those produced by brain imaging equipment. The alternative is the approach epitomized by the Hebb Conjecture (1949) that asserts that it is most likely that cognitive processes of all kinds are instantiated in the concurrent microneural state of a huge assembly of individual neurons whose coordinated action is the neural equivalent of mental activity. Although there is currently insufficient evidence at the microneural level to robustly support the Hebb conjecture, it is widely accepted that Hebb's approach is probably correct—cognition is a result of the activity of independent, but interacting, neurons that maintain their individual identity. The problem is that both of these approaches are deeply flawed for methodological, conceptual, and empirical reasons. One reason is that simple models composed of a few neurons may simulate behavior but actually be based on completely different neuronal interactions. Therefore, the current best answer to the question asked in the title of this contribution is–probably not.
Key words: behavior, cognition, neuroreductionism, theory

In the last three decades, cognitive neuroscience has undergone a revolution in substance and available technology. With the invention of functional Magnetic Resonance Imaging (fMRI) by Ogawa, Lee, Kay, and Tank (1990), brain images using Blood Oxygen Level Dependent (BOLD) measures have largely replaced traditional methods involving surgery, trauma, stimulation, and electrophysiological recording and a number of methodological breakthroughs. The earliest application of the fMRI technique to cognitive neuroscience is generally considered to have been published by Belliveau et al (1991) during which attempts were made to map out the visual cortex. Currently, the number of papers that deal specifically with comparisons of cognitive processes and brain images is difficult to determine. Nevertheless, it is obvious that experiments comparing fMRI responses and cognitive processes have gone from virtually nothing in the early 1990's to tens of thousands in the 2000's (e.g., Bandettini 2007; Aue, Lavelle & Cacioppo, 2009).

Because these cognitively related fMRI images are brain responses, more or less correlated with cognitive processes, there is a compelling face validity presented to researchers by these results. The promise is that we are on the verge of a breakthrough to the profound and hitherto unsolved mind-brain problem. The face validity of mind-brain equivalence is, however, countervailed by a pervasive logical error that

incorrectly assumes that because the mind is a brain process, any brain activity correlated with cognitive activity must be relevant. At the least, it has been expected that correlations could be found between cognitive and neural responses that might serve as heuristics for future theories of mind-brain relationships. Despite the large number of publications and the large number of investigators working with the fMRI and similar macroneural tools, no clear path using this approach to an overarching theory of the mind-brain relation has yet appeared—only piece-wise correlations between particular locales.

The major alternative approach—the Hebbian, microneural conjecture—based on the concurrent activity of a large number of individual neurons is currently attracting considerable attention and importance in light of the plan by the current United States administration to invest billions of dollars into the task of recording every spike from every neuron in the brain of a behaving organism. The development in the past few years of powerful computers has raised the possibility that sheer numerousness may not be the hindrance thought until only recently. However, the practical and computational aspects of large scale computer simulations of the brain suggest that even the largest computers may not provide an answer to how brain activity accounts for mental activity.

The possibility of an overarching theory of or explanation of how phenomena measured behaviorally are related to neurophysiological measures remains contentious. Many contemporary philosophers have adopted a conservative point of view with regard to the tractability issue on logical grounds including Rakover (2011), McGinn (1989), Ludwig, (2003), and Dietrich and Hardcastle (2005). These and other scholars have arrived at a conclusion of theoretical intractability. (Strong, unequivocal quotes from these scholars and others arguing on both in principal and current in practice bases can be found in Rakover's article). On the other hand, contemporary scholars such as Churchland (1994), Bechtel (2002), and Bickle, (2003) have compellingly argued that although the problem is difficult, it probably will be solved as our knowledge of neurophysiology expands.

There are also many philosophical points of view that seek to find some kind of a relationship between neurophysiology and mental activity that, I believe, do not constitute neuroreductionist explanation in the sense I use here. For example, such concepts as "identity" (Fiegl, 1958) and the classic idea of "emergence" of complex phenomena from simpler phenomena finesse the problem by excluding the specific processes of the transformation from one level to another and, thus, do not offer a possible path to mechanistic explanation. Similarly, a number of other philosophical ideas such as "supervenience" (Kim, 1993) accept the argument that lower level mechanisms determine higher level properties but do not identify the specific mechanisms that must lie in the heart of such a neuroreductionist explanation. Such logical arguments concern only generalities about relationships and, therefore, do not offer specific explanations of the transformation processes that can account for the emergence of the mental from the neural (or, better stated, the transformation from one to the other). Although physicalist in principle, they are nonreductive in practice. Such ideas are actually metaphors that help us to conceptualize but are not

explanations in the sense implied here; they are much better considered to be definitions or ontological premises.

To sharpen the following discussion, it is desirable to define what I mean by "neurally reductive" and "nonreductive approaches," respectively. By a neurally reductive explanation, I refer to a transformation of the properties of a lower level (such as the action of the brain or its constituent parts) to the properties of higher level phenomena (such as cognitive module that may be inferred from behavior).x In a practical sense, this would mean that manipulation of neurophysiological properties and measures so as to reproduce or cause behavioral properties and measures that would allow us to infer underlying mental activity of the kind designated as hypothetical constructs; in other words to solve the mind-brain problem. It is very important to appreciate that the "cognitive processes" are not directly observable and can only be indirectly inferred from publically observable behavior.

Neuroreductionism assumes that the sum of the neural parts and interactions in some way cause or become the mental experience. How they do so is the core of a neuroreductionist explanation. Neuroreductionism does not imply dualism; instead both levels of analysis are assumed to be manifestations of a single level of physical reality—a monism without which cognitive neuroscience cannot survive. A nonreductive approach would solely concentrate on the parts and properties of the higher level (e.g., cognitive processes) and eschew attempts to demonstrate how those of the lower level (e.g. neural properties) produce or become those of the higher level.

An analogy may be drawn between the kind of neuroreductionism I consider here and a much simpler explanation from the physical world. The macroscopic properties of a plastic can be almost entirely explained by the interaction of the properties of the microscopic molecules that are polymerized to produce the plastic. In this analog, what we know about the properties of the constituent molecules (and their interactions) permits a full reductive explanation of the properties of the macroscopic materials they produce. It is this kind of explanation to which cognitive neuroscientists ideally aspire. Unfortunately, the present limitations of our science suggest that because of its complexity no such explanation is on the foreseeable horizon.

The following discussion is intended to alert us to some of the conceptual, procedural, and empirical challenges encountered in the search for a neuroreductionist explanation of cognitive activity. At this point it is not possible to definitively determine whether or not the problem is tractable in any general scientific sense. Nevertheless, it is clear that the identified barriers to progress do at least raise the possibility that current efforts to develop an overarching neuroreductionist theory of mind-brain may not be obtainable. If the goal of an overarching theory can be shown to be currently unrealistic, then psychology may wish to reorient itself, at least for the time being, toward a nonreductionist approach to the study of cognitive processes; one that would emphasize the observation and description of behavior without the neural empirically unsupported pseudoexplanations or metaphors that proliferate in the field today.

Conceptual Barriers To Theory Building

Many of the barriers to the development of neuroreductive explanation in cognitive neuroscience are matters of deep dispute about the fundamental nature of the scientific method. Hopeful expressions of the inevitability of future discoveries, speculative findings, and "converging methods" are often raised as keys to problem solutions already known to be intractable (e.g., Moore, 1956). Past successes in theory development in the physical sciences are often touted as harbingers of success in cognitive neuroscience despite the fact that physics deals with uniform forces (e.g., gravity) that are orders of magnitude simpler than the interactions observed in brain tissue.

In this section I deal with some of the often overlooked conceptual issues that arise when we ask the question of the plausibility of neuroreductionist explanations of cognitive activity.

Level of Analysis

A major property of any proposed explanation concerns the level of analysis at which a problem is conceptualized. By" level," I refer to both the scale at which brain activity becomes cognitive activity and the scale of our analyses and measurements. In the present context, the Hebbian microneural level of analysis entails responses of cellular and subcellular units such as neurons and synapses and their respective idiosyncratic interactions. (A macroneural level of analysis, on the other hand, deals with signals that are pools or accumulations of these lower level responses.)

Microneural level interactions are the bases of the Hebb conjecture (Hebb, 1949)—a conceptualization that assumes that cognitive processes emerge from the unpooled (i.e., that their magnitudes are not just added together), but interacting, action of millions if not billions of neurons and the thousand-fold greater number of synaptic connections. Macroneural-level based theories, on the other hand, seek to overcome the sheer numerousness of individual neurons by pooling (either statistically or in terms of their magnitudes) the microneural results into composite measures such as fMRI images or EEG recordings. Each of these levels of analysis is beset with its own difficulties and challenges.

For example, to establish the validity of a Hebb-type microneural theory requires that we control and measure the individual activities of a very large number of neurons. We would have to identify many, if not most, of the neurons that are involved in a cognitive process and then provide some means of stimulating and then recording from them individually in order to provide the essential empirical data. This is necessary since each neuron in the microneural network acts in its own idiosyncratic manner even as it presumably contributes to the ensemble response. In other words, the ensemble of interacting neurons acts as a "super-register" in computer terminology—a collection of units that maintains their individual identity while collectively accounting for the properties of the system. A simulation that results in a behavior is not tantamount to an explanatory theory. A full blown Hebb-

type theory would not be satisfied by a less than exhaustive description—if that is the way the mind emerges from the brain.

Complexity

Obviously, achieving these connections would be a technical tour de force of the first magnitude given that our limited ability to develop microelectrode arrays suitable for stimulation and/or recording is currently limited to a few hundred (e.g., Wark et al, 2013). Furthermore, the amount of data obtained from even a few hundred microelectrodes would tax available computational capabilities as complex patterns of response exponentially proliferate.

Despite the fact that we must consider the study of any microneural network in this direct manner (manipulating and recording individual neuronal responses) to be far beyond current technology, it is at least possible to consider the criteria that would have to be applied to determine either its necessity or sufficiency. Martin, Grimwood, and Morris (2000) have suggested four criteria by means of which such a Gedanken experiment would have to be evaluated to support a particular microneural theory (a "synaptic plasticity and memory" hypothesis in their terms). They carry out this conceptual exercise for learning and memory, but the four formal criteria they proposed hold true for any microneural, neuronal network experiment. I reconstruct them here to more closely fit the present discussion.

Table 1
Four Formal Criteria for Support of Microneuronal Network Theory of Cognition (Based on Martin, Grimwood, and Morris', 2000, Discussion in the Context of Learning and Memory)

(1) Detectability: Changes in the cognitive process must be accompanied by detectable changes in the organization of the supposed neural network.
(2) Mimicry: If the same neuronal network state observed during some cognitive process can be reconstructed, then the same cognitive process should reoccur.
(3) Anterograde Alteration: Anything that prevents relevant neural activity also should prevent the associated cognitive process.
(4) Retrograde Alteration: Any changes in a preexisting neuronal network state should alter the cognitive experience.

Obviously, for a host of procedural and instrumentation reasons, experiments to meet these criteria would be extremely difficult, if not impossible, to carry out. Others have argued that any proposed attempt to study the brain at this minute level is beyond both our technology and our mathematics (Lichtman & Denk, 2011). Their reasons include; (1) the immense diversity of cell types in the brain; (2) the idiosyncratic nature of cellular responses in both space and time; (3) the scale difference between microscopic neurons and their extent over vast macroscopic regions of the brain (brain must be studied "over sizes that span six orders of magnitude," p. 620); and (4) the overwhelming flood of data at the microneural level.

Because of these technical and procedural limitations, the direct, brute force, microneural strategy of stimulating and recording from the necessary individual components of a microneural network is not likely to ever be implemented despite the advent of supercomputers of enormous power. Microneural level theory building in such a many-neuron environment is severely inhibited because of the practical problems associated with processing a representative number of neurons. Even if we could assemble such an experiment, we would be dealing with so much data, that it would be for all practical purposes inaccessible. (On the other hand, at the macroneural level typified by fMRI images, the necessary microneural data is also inaccessible, but for a different reason—it no longer exists; it has been averaged out of existence.) Thus, the possibility of developing an authentic, valid, overarching microneural Hebb-type explanation based on empirical evidence that truly speaks to the question of how the brain encodes mental activity appears to be remote. Regardless of which level of analysis one is considering, the combinatorial complexity of the brain stresses any putative explanation of its operation to its limits.

The complexity issue has been known for years. It takes very few interacting neurons to pose an intractable combinatoric problem for any currently conceivable method of analysis. Stockmeyer and Chandra (1979), among others showed how very simple network problems (e.g., minimizing a travelling salesman's itinerary) could require hyper-astronomical processing times. For example, other estimates suggested that if the salesman had as few as 21 stops, an exhaustive solution to the problem would require 77,000 years of computer time.

Karp (1986), a prominent complexity theorist, reported that the intractability of many other superficially simple combinatorial problems, many of which were analogs of those faced in cognitive neuroscience, had been rigorously demonstrated by mathematicians.

Koch (2012) recently put this problem into a cognitive neuroscience perspective when he calculated that the time required to "exhaustively describe a [neural] system" would grow faster than exponentially. He calculated that the two million neurons of the visual cortex of a mouse would require a computer system running for 10 million years even if computer power was simultaneously growing at the modern rate specified by Gordon Moore's law ("the number of registers or gates that can be packed on a standard size computer chip doubles roughly every 2 years.") Koch did go on to note that there were some simplifying conditions that could speed up the process by a few orders of magnitude but the processing time remained enormous.

Beyond sheer numerousness and nonlinearity, many early theories of neural nets do not scale up well. The eventual collapse of the small neuronal network program in the 1980's when attempts were made to scale up the number of synthetic neurons beyond one hundred was probably an unavoidable outcome of this intrinsic instability.

Loss Of Information With Response Accumulation

In such a situation, in which an exhaustive, direct microneural (neuron by neuron) analysis is currently beyond our technical abilities, cognitive neuroscientists

have turned to available macroneural cumulative measures (e.g., the fMRI) and effectively let nature do the sums for us. How the cumulative process works in which microneuronal responses are cumulated neurophysiologically is not completely understood but it may involve summations of either local potentials (e.g., Logothetis et al, 2001) or spike action potentials (e.g., Mukamel et al, 2005).

Unfortunately, it seems likely that the cognitively salient microneural information is not preserved in macroneural level measurements. Once added or accumulated into pooled responses, the simplest arithmetic ideas (e.g., what two numbers are added together to provide a sum) argue that they cannot be retrieved. Thus, there is an inescapable ambiguity about the origins of any macroneural signal formed from accumulations of microneural activity—a huge variety of different configurations of microneural neuronal activities may produce exactly the same macroneural response. This means that scores based on differences between control and experimental *macroneural* responses cannot discriminate between quite different microneural brain states.

The implication of this insensitivity to differences is that that macroneural fMRI images are neutral with regard to their microneural origins. This is the fundamental weakness of the macroneural approach *if the Hebbian conjecture of microneural coding of cognition is correct*; the critical microneural information essential to understanding the transition from brain activity to cognitive activity has been lost. Of course, if the Hebb conjecture is wrong, loss of microneural information may not be consequential.

Underdetermination

Another classic conceptual issue serving as a barrier to theory building is generically referred to as "underdetermination." Underdetermination implies there is insufficient information in the results available from an experiment or group of experiments to answer the salient questions for which the research was carried out. This is a constraint operating in almost all scientific enterprises; however, it is particularly exacerbated in cognitive neuroscience because of the multitude of plausible and possible alternative neuronal configurations that could underlie an observable behavior. However, it is more likely that most of the effects of underdetermination are due to the fundamental conceptual difficulty that the information necessary to definitively answer a cognitive neuroscience question is not accessible from macroneural or behavioral measurements. This is a theorem of automata theory (Moore, 1956) that is not well known to cognitive scientists but has been extensively discussed by scholars such as Duhem (1914) and Quine (1951). Moore puts it this way:

> Given any [closed] machine S and any multiple experiment performed on S, there exist other machines [internal mechanisms] experimentally indistinguishable from S for which the original experiment would have had the same outcome … This result means that it will never be possible to perform experiments on a completely

unknown machine which will suffice to identify it from among the class of all sequential machines. (p.140)

As an example of how this may apply to cognitive neuroscience, Hilgetag, O'Neil, and Young (1996) have argued on purely mathematical grounds that complex networks such as those proposed by Van Essen, Anderson, and Felleman (1992) for the primate visual system cannot be placed in a unique hierarchy of activation order because the outcomes are fundamentally underdetermined. As a result carrying out additional experiments leads, paradoxically, to an increase in the number of possible but incorrect theories at a rate that exceeds the number of new experiments that can be designed. In short, simply collecting more data may not always help to resolve issues of underlying mechanism if the data are underdetermined; indeed, it may exacerbate the problem.

False Analogies Between Sensory, Motor, And Cognitive Processes

Another conceptual barrier encountered in the effort to formulate theories relating neural and cognitive processes is the false analogy often drawn between successes in explaining sensory and motor processes, on the one hand, and the much more complicated matter of understanding cognitive representation. However, the two tasks are entirely different. Peripheral sensory and motor signal transmission is characterized by a mainly unidirectional flow of information (although with some centrifugal influences), relatively simple dimensional correlations between stimuli and neural responses, as well as between stimuli and perceived experiences, and tight linkages to physical dimensions and measures. Thus, it is possible, to ask and answer a question such as—what is the neurophysiological code used by the peripheral nervous system to transmit information about the magnitude or quality of a stimulus to more central regions of the brain? The physical dimensions of both the stimulus and the neural response are measurable in units for which the physical sciences have provided well established values such as lumens, decibels, on the one hand and frequency and pattern of firing on the other. Furthermore, psychophysicists such as Brindley (1960) note that techniques for quantifying sensory responses require and can utilize but the simplest of "Class A" discriminative responses—simple "same" or "different" judgments. Furthermore, the microanatomy of the peripheral sensory and motor pathways is much simpler than of those more intricate and interwoven central mechanisms that are presumably involved in cognitive processes and, therefore, compelling results were promptly forthcoming.

These simplifying conditions do not hold for cognitive process such as "decision making" or "emotion." When the neural coding idea was transferred to cognitive processes, the result was general disappointment; none of the simplifying anchors that provided conceptual order to the transmission codes existed in the complex and irregular networks underlying higher level cognitive processes. The result was that successes in understanding these sensory and motor information transmission processes would not assure the tractability of central processing mechanisms. A

similar distinction can be made between old and new phylogenetic regions of the brain; the former being susceptible to functional analyses and the latter being opaque to them.

Poor Definition Of Cognitive Constructs

For science to achieve a high level of theoretical precision there is a profound need that the dimensions and properties of its variables be precisely defined. Yet, psychology is replete with poorly defined terms such as attention, thinking, emotion, and consciousness at a high level of abstraction and words such as learning, recognition, detection, retrieval, and conditioning defined more operationally. Many cognitive terms have multiple meanings that vary with the particular goals of the research. For example, Vimal (2009) tabulated 40 different meanings of the word "consciousness." Uttal (2011), as another example, listed 41 different meanings of the word "learning." The ultimate problem, however, is that none of these fluctuating cognitive terms need necessarily map directly onto neural mechanisms. That is, what is a fundamental elemental process or faculty to a psychologist may mean something very different to the brain. In short, our behavioral taxonomies and those of the brain may not be speaking the same language, and no Rosetta stone translating between them is possible.

An alternative way of emphasizing how this lack of isomorphism between cognitive and neural vocabularies affects our thinking has been provided by investigators such as Oosterwijk et al. (2012). They support the main alternative hypothesis—namely that psychological constructs are encoded by a distributed pattern of general purpose neural mechanisms. One implication of such a conclusion is that, however much we may seek to modularize our psychological taxonomies, there is no reason to assume that the mental components are parsed in the same way by the nervous system. In Oosterwijk et al's words—"the brain does not respect faculty psychology categories." (p. 2110). It should be noted that the kind of distributed neural network described by Oosterwijk et al is very different than the classic PDP model of decades ago.

The Neurologizing Of Psychological Language

A property of many psychological theories is the casual introduction of neurophysiological terminology into their discussions without adequate empirical linkage. Behavioral hypothetical constructs are supplemented by neural concepts and findings based on little more than analogies and borrowed language. Skinner (1950) referred to this use of unsubstantiated neurophysiological concepts as a "Conceptual Nervous Systems" (CNS) noting that "many theorists point out that they are not talking about the nervous system as an actual structure undergoing physiological or bio-chemical changes but only as a system with a certain dynamic output." (p. 194)

What it appears that many investigators are doing is to develop plausible, but highly speculative, "hypothetical constructs" that could, in principle, describe the

behavior. The hypothetical mechanisms are then associated with similar neurophysiological findings or redefined in neurophysiological terminology. At best, theories of this type are heuristics that are not likely to rise from preliminary hypotheses to robust neural theories. It seems appropriate that we should be reevaluating the important insight about the CNS offered by Skinner and distinguish between robust neurophysiological entities and neural "nicknames" for sometimes ingenious hypothetical constructs.

The Search For Objectivity

Despite the difficulty in defining or measuring inaccessible cognitive entities such as thoughts, feelings, experiences and perceptions, there has been a persistent historical effort to seek out objective "hard science" correlates—in particular, brain responses—of cognitive processes. The fundamental ontological premise of all such efforts has been the physicalist assumption that psychological activity is the outcome of brain activity in some so far indefinable way. Few cognitive neuroscientists would disagree with this point. However, there is an epistemological corollary of this assumption that is a matter of greater contention. Namely, can any measurement that is made of brain processes be theoretically transparent to cognitive activity? That is— are objective macroneural measurements such as the fMRI or the EEG capable of representing the neural basis of cognitive activity adequately to serve as an "explanation" of that activity? The answer to this question is probably not in the strong sense of "explanation" I use here. Many "objective" measurements of brain activity are simply going to be irrelevant to the mind-brain problem and are, therefore, invalid as measures of cognitive activity. Studies of the chemistry of neurons and synapses, microscopic investigations of structure, and single cell recordings (e.g., see Bickle, 2006) tell us little about the cognitively meaningful information processing activity of what is increasingly likely to occur in interactions among neurons situated at widely distributed portions of the whole brain. It is not the molecular chemistry of neurons nor of synapses that might open the door to understanding how the brain works as a cognitive engine; it is more likely to be the aggregate information processing by a large number of neurons whose coordinated (but not physically cumulative) states account for the emergence of mental activity.

Statistical And Reporting Barriers To Theory Building

In addition to these conceptual issues, the complexity of the statistical procedures required to extract meaningful conclusions from both macroneural and microneural observations is such that there is ample opportunity for statistical artifacts to lead to erroneous conclusions. These errors come in many guises–some are very simple computational oversights and some are properties of statistics well known to statisticians if not to cognitive neuroscientists. Many of these difficulties with statistical and reporting methodologies are common to all levels of cognitive neuroscience, both macro-and microneural, as well as in many other forms of purely

psychological research. Indeed, statistical problems are common in many other sciences. I concentrate here, however, on a few of those statistical problems that have been identified as being particularly troublesome to macroneural neuroreductionist theory development. Because these sources of bias and artifact have been discussed extensively in the literature, I simply summarize them (and some of the evidence that they remain major problems for macroneural methodologies) in the following list. My general point is that there are major problems with the most basic aspects of cognitive neuroscience—particularly the reproducibility and validity of its empirical observations. Some of these sources of bias will, of course, be overcome by further technical developments; others, however, may represent constraints on theoretical progress that are much more fundamental.

(1) Inadequate reporting of cognitive neuroscience experimental methods that prevent replication (Carp, 2012a; 2012b).

(2) Alternate computation of cognitive neuroscientific analytic trajectories produce different conclusions from the same initial conditions (Carp, 2012a; 2012b).

(3) Ignoring the number of tests (Bennett et al, 2009; 2011 in cognitive neuroscience experiments; Genovese, Lazar, and Nichols (2002); Loring et al (2002). Bennett et al (2011), for example, reviewed a sample of relevant literature and found that only 60-75% of the articles made the needed correction even using an approximate correction (Holm, 1979).

(4) Inadequate sample size (Ihnen et al, 2009; Wager et al (2009). The latter investigators pointed out that in 415 early cognitive neuroscience studies an average of only 11 or 12 subjects was used and in some cases as few as 4 or 5.

(5) Double dipping (Vul et al, 2009; Vul and Kanwisher, 2010; Kriegeskorte et al 2009; Fiedler, 2011). Double dipping comes in many guises, for example in the form of functional localizers that predefine salient areas and, thus, bias cognitive neuroscience analyses (Friston et al, 2010 and Saxe, Bret, and Kanwisher, 2006). Other subtle forms of double dipping include not controlling for intertrial (Goldfine et al, 2013) or intervoxel (Loring et al, 2002) dependencies; ignoring nonlinear interactions between spontaneous and evoked responses (He, 2013); and injudicious use of packaged analysis programs that have built-in biases of this kind (Joppa et al, 2013)

(6) Erroneously comparing significances when one is significant and the other is not (Gelman and Stern, 2006). In a study of this phenomenon, Nieuwenhuis, Forstmann, and Wagenmakers (2011) examined 513 articles from first line neuroscience journals and discovered that approximately 15 percent incorrectly judged a difference between a significant and an insignificant distribution as being significant. This error, according to Nieuwenhuis and his colleagues is "especially common in the neuroimaging literature." (p. 1107)

(7) General inadequacies of the ubiquitous Null Hypothesis Significance Test (NHST) (Greenwald et al, 1996; Lieberman and Cunningham, 2009; and

Lambdin, 2012 among many others) especially when applied to cognitive neuroscience experiments. Eklund et al (2012) reevaluated 1484 data sets that had previously been generated to examine the *rest* activity of the brain. All of these data sets were expected to have little significant cognitive related activity and, thus, approximately 5% false positive responses among the 1484 experiments. Eklund and his colleagues, however, found that significance rates actually ranged from 1% to 70%.

(8) Lurking in the background, of course, is the underappreciated but ever present problem of mistaking "correlation for causation" (Yule, 1899)—a caveat long understood and yet often violated by cognitive psychologists.

Of course, in a certain sense, these are technical matters discussed by statistically oriented scientists in many fields that may someday be overcome by the invention of improved methods. However, the point is that much of the current corpus of findings concerning mind-brain relations is suspect as a result of subtle and not so subtle statistical artifacts that challenge their most basic validity. All of the examples I allude to are framed in the context of cognitive neuroscience. By themselves, these statistical deficiencies do not provide a compelling argument against neuroreductionism; however, they do suggest that the current situation is filled with what are nonreplicable and invalid findings and that the case for a future neuroreductionist explanation of cognitive processes is still problematical.

Empirical Barriers To Theory Building

Conceptual and statistical sources of difficulty notwithstanding, the most compelling influences on theory development are and should be the empirical results. Should the results of equivalent experiments not be consistent, no effort at synthesis or prediction can possibly bear fruit. Although science often acts in an indirect or heuristic manner and the path from data to meaningful understanding is sometime convoluted and indirect, its ultimate great successes rest on a foundation of replicability and validity.

In recent years, a substantial body of new findings has raised doubts about some of the most basic assumptions of cognitive neuroscience. In this section, I discuss some of the changes occurring in the field that should have an impact on theory building.

The Case Against Macroneural Localization

The main goal of cognitively related brain imaging research in the two-plus decades since fMRI methods were introduced has been to determine the spatial localization and extent of activation regions or nodes that selectively respond during particular cognitive activities. In some of the early forms of this idea, the nodes were isolated from each other and the goal was simply to map modular cognitive processes on local brain regions. When subsequent research (starting about the year 2000)

revealed that more than one brain region was almost always associated with every cognitive process, the working metaphor about brain organization began to change. In this new stage, the emerging metaphor became one of interacting, sparsely distributed, multiple, non-specific regions of the brain.

The very important question now arises—do localized focal regions or nodes actually exist? Recent work is beginning to suggest that the metaphor of nodes or localized activation regions is, in fact, not the way the brain is actually organized. The alternative view, expressed by Lindquist et al (2012), among others, is that the brain operates in a much more distributed and holistic fashion, as well as less specialized, than previously suspected.

This is not to suggest any mass action or continuous equipotential field action on the part of the brain. Instead, the idea is that the neuronal mechanisms accounting for cognition are very widely, if not uniformly, distributed throughout the brain. Indeed, it now seems increasingly likely that almost all of the brain may be responding to virtually any kind of cognitive activity once beyond the sensory and motor areas where localization does obviously hold! This is important because it raises questions about the existence of specialized localized regions of the brain, a central postulate of much of today's theory.

Robust evidence supporting widely distributed brain states (as opposed to localized activations) of cognitive activity has now begun to appear in the literature (e.g., Gonzales-Castillo et al 2012; Thyreau et al, 2012.) Both of these groups of investigators presented compelling data that the apparent separation of brain responses into localized regions or operational nodes may be an artifact of inadequate sample sizes, arbitrary p values, statistical biases, and a prevailing presupposition of sparse localization. Gonzales-Castillo and his colleagues, for example, suggested that when adequate sample sizes (up to 500) were used to average brain images, localized activation areas tended to disappear to be replaced by what is nearly a uniform overall activation of the entire brain. They did not exclude some relatively large regional differences but they concluded that:

> … under optimal noise conditions, fMRI activations extend well beyond areas of primary relationship to the task; and blood-oxygen level-dependent signal changes correlated with task timing appear in over 95% of the brain for a simple visual stimulation plus attention control task. (p. 5487)

Thyreau et al (2012) also reported a similar experimental result in which they also used a very large subject sample (n=1326) and found generally the same result—widespread distribution of responses across the entire brain for even the most elementary cognitive process. They pointed out that when very large samples were used, as expected, even small responses began to achieve statistical significance; this expansion increased until virtually all of the regions of the brain were responding significantly. The important implication of their work, like that of Gonzales-Castillo and his colleagues, was that the apparent segregation of the human brain into cognitively specialized regions associated with particular cognitive processes or brain activation nodes was probably an artifact of inadequate sample size.

In sum, the brain is not a system of isolated operation-specific nodes; it is more likely to be a broadly distributed system of neural mechanisms that are neither operation-specific nor localizable whose responses are wide spread and relatively uniform. Although there may be some discrepancies in effect size at each of many points, the idea of specialized locales is generally not supported.

On The Reliability And Replicability Of fMRI Data

The next empirical issue is—how reliable are macroneural data? In the past two decades, the body of the cognitive neuroscience literature based mainly on fMRI measures has ballooned. As is usual in the case of psychological research, there are a very large number of possible parameters that can be studied. Therefore, we would expect that exact replications of experiments to be relatively rare. Indeed, only a few investigators have made a determined effort to study the replicability of cognitively related findings. Those that have typically report that replicability in cognitive neuroscience is elusive and that experiments that seem to be very close in intent and design may produce different estimates of which brain mechanisms are involved in any particular cognitive process. Observations of inconsistency appear in repeated trials on a single subject, in comparisons of pools of subjects and interexperiment comparisons, and among the relatively few meta-analytic studies in which data from a number of experiments are pooled.

The most disconcerting of these comparisons are obtained from single subjects repeating exactly the same protocol. As a measure of replicability this single-subject, repeated-measures design is the most stringent test possible since a maximum number of factors are controlled. Among the few investigators who studied the variability of data obtained from single subjects were Aquirre, Zarahn, and D'Eposito (1998), McGonigle et al (2000), and Miller et al (2002). One of the few recent reports that have specifically studied the problem of reliability in comprehensive detail was published by Bennett and Miller (2010). Examination of the figures from other reports (e.g., Lund et al, 2005), which were not specifically studies of single subject replicability, also shows substantial intrasubject variation for explicitly cognitive tasks. Furthermore, Rau et al (2007), for example, reported that a naming task produced highly variable data concerning the location and extent of Broca's area, one of the most widely accepted associations of brain regions and cognitive processes. This inconsistency was partially ameliorated by pairing it with a noun generating task. Duncan et al (2009) also reported low levels of intrasubject consistency and reliability when brain scans of interest were compared.

The question of reliability or replicability remains contentious, however. Maitra, Roys, and Gullapalli (2002) have argued that variability was modest. Aron, Gluck, and Poldrack (2006) and others reported a high degree of reliability in an experimental protocol—a "probabilistic classification learning" task—with a one–year interval between the test and retest portions of the experiment. However, even this work showed some evidence of changes over the one year period between the original test and the retest.

What may appear to be a conflict in findings may be resolvable by noting that the least variable brain responses were produced by motor tasks (e.g. Maitra, Roys, and Gullapalli) whereas the more variable responses were associated with higher level cognitive tasks (e.g., McGonigle et al, 2000). The problem of single subject reliability remains quite controversial as evidenced in a recent updated discussion of his original work by McGonigle (2012) and reevaluations of these results by Smith et al (2005) who disputed McGonigle et al's (2000) original claim that intersession variability was unacceptably high.

The intrasubject variability of the fMRI responses reported by Aquirre et al, McGonigle, and Miller and their groups, respectively, highlights the inconsistency of macroneural results especially those from repeated measures on a single subject. Although a few investigators (e.g., Mikl et al, 2008) have attempted to develop data manipulation methods to "smooth" the variability, the basic conundrum remains—is variability an artifact of our instrumentation and experimental protocols, or is it a real psychobiological property of our subjects? If the former, future technical developments may resolve the problem: If, on the other hand, the latter is true, then the entire enterprise of using brain imaging in the search for the neural bases of cognition probably should be critically reevaluated.

Whatever the resolution of these debates, there is no question that intrasubject variability is less than intersubject comparisons in an episodic retrieval task. This is shown in Table 2 from Miller et al (2002).

Table 2
Intra- and Intersubject Correlations in an Episodic Memory Task (after Miller et al., 2002.)

	S.C.	K.B.	B.B.	H.G.	C.C.	B.K.
S.C.	.63	.12	.11	.19	.08	.11
K.B.		.47	.19	.25	.19	.23
B.B.			.40	.29	.25	.25
H.G.				.50	.27	.30
C.C.					.43	.20
B.K.						.44

In this table, correlations are shown between subjects and within subjects for two repetitions of an experiment in episodic memory. The low correlations between different subjects are not surprising. However, even the higher, but still relatively modest levels of correlation, when the same subjects were retested (as indicated by the scores along the diagonal) suggests a lack of consistency that challenges the very basis of the macroneural approach.

A high level of variability is also observed when comparing experiments. A graphic depiction of interexperiment variability within a restricted domain of cognitive neuroscience—emotions—is shown in Figure 1. This figure is a typical map

of the raw distribution of activation areas reported from a number of what were presumed to be related experiments.

Figure 1. 437 Activation Peaks Generated from 162 Experiments on Emotion
(From Kober et al, 2008, with the Permission of the Publisher.)
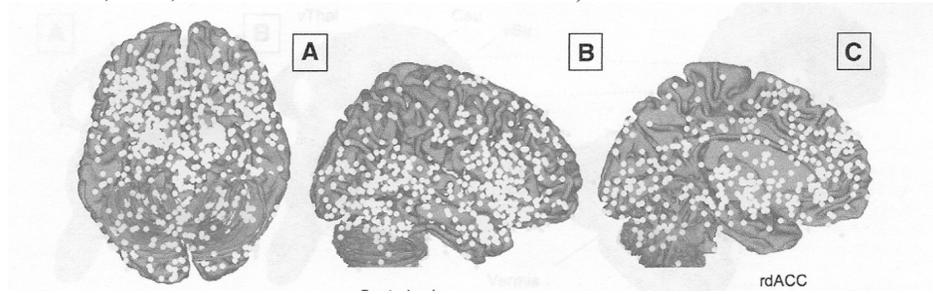
These records show the very wide distribution of responses to what were the same stimuli for a large number of experiments. This is compelling evidence for the large amount of variability observed in fMRI experiments.

The next step in considering variability is to compare meta-analyses. A graphic comparison of seven meta-analyses carried out for working memory is shown in Table 3 to illustrate the diversity of results obtained in this comparison of meta-analyses.

Table 3
A Comparison of Seven Meta-analyses dealing with Working Memory (from Uttal, 2013).

BA*	Meta A	Meta B	Meta C	Meta D	Meta E	Meta F	Meta G
3							X
4				X			
5	X	X	X	X	X	X	X
7	X		X		X	X	X
8	X				X		
9	X		X		X		X
10	X	X	X		X	X	X
11	X						
13	X	X					X
17	X						
18	X						X
19	X	X	X	X			X
21	X			X			
22	X		X	X			
24	X					X	X
31	X						
32	X	X	X		X	X	X
37	X		X	X			
39	X						
40	X	X	X		X	X	X
44	X		X		X		X
45	X	X			X	X	X
46	X		X		X	X	
47		X				X	

*Plus scattered reports of thalamus, cerebellar, etc., activations.
The activation data in this table are taken from the following articles by column:
A, Wager and Smith (2003); B, Glahn et al. (2005); C, Simmonds, Pekar, and Mostof-
sky (2008); D, Turkeltaub, Eden, Jones, and Zeffiro (2002); E, Owen, McMillan, Lair,
and Bullmore (2005); F, Chein, Fissell, Jacobs, and Fiez (2002); G, Krain, Wilson,
Arbuckle, Castellanos, and Milham (2006).

Note. *BA = Brodmann Areas

The results of this comparison are also clear—there was substantial variation among the areas reported by the seven meta-analyses for what were considered to be nearly the same cognitive process; this illustrates, at another level of analysis the relatively high degree of variability observed in this kind of experiment.

In sum, the relatively rare currently available comparisons of intrasubject, intersubject, interexperiment, inter-meta-analysis variability (see Uttal, 2013 for an extended discussion of reliability in fMRI experiments), raise serious questions about the possibility that a neuroreductionist theory of cognition will be forthcoming. Any plausible theory depends on a foundation of reliable data and there is a serious question about replicability in these macroneural measurements. There is, perhaps, no more important task for cognitive science than the resolution of the empirical crisis in replicability and reliability.

New Developments; Can They Provide A Pathway To Neuroreductive Explanation?

Clearly, there must be a continuing effort to overcome the obstacles to understanding how brain activity becomes mental activity. New methods are constantly being suggested, some of which are aimed at improving data collection and analysis methodology, some at proposing alternative methods of dealing with networks, and some of which try to develop alternative concepts or forms of the basic question of mind-brain relationships. It is questionable, however, if any of these new developments will overcome the difficulties involved in this awesome task. False hopes appear at both the macroneural and microneural levels as logic and empirical evidence fail us. Some of this confusion is due to the misunderstanding that not all progress in brain science or psychology is relevant to cognitive neuroscience. Great strides in brain anatomy and neurophysiology, psychological theory, and computer simulation do not always lead us to sustainable ideas about how the brain produces or becomes cognitive activity.

For example, recent efforts at making the brain visually transparent (Chung and Deisseroth, 2013) are magnificent tools for visualizing the anatomy of brain tissue. However, it is not equally clear how such images might contribute to mind-brain relations. After all is said and done, they are post- mortem studies about macroneural anatomy. Similarly, the development of devices such as the Diffusion Tensor MRI (Westin et al, 2002) has provided detailed information about the anatomy of the tracts interconnecting brain regions; but this anatomical magnificence has not yet been shown to directly contribute to a neuroreductionist explanation of cognitive processes.

An increasing number of investigators have directed their attention to the general problems of network organization. This can be approached from two points of view—large scale and neuronal-level networks. Meehan and Bressler (2012) and Bressler and Menon (2010) dealt with networks in a mainly qualitative way exploring the principles of brain organization that might guide future research. Stam and van Straaten (2012) extend the discussion to the particulars of networks—edges, nodes, and connectivity. Barrett and Satpute (2013) link the operation of these large scale networks to a particular cognitive state—emotion.

Many of the new approaches are based on the idea of networks of interacting macroneural nodes. All of these macroneural theories depend on mapping cognitive processes onto what is now appreciated to be at least a group of brain structures and at most near universal participation by all parts of the brain. They, thus, provide an organizational theme for testing computer models of how the brain might be organized to produce a particular cognitive process. However, despite partially successful simulations such as those by Schroll, Vitay, and Hamker (2012) and by Ashby, Ennis, and Spiering (2007) it adds to the argument that this macroneural network approach is not an appropriate level at which to attack the mind-brain problem.

Alternatives to the traditional approach of mapping specific cognitive modules onto what are now better considered to be combinations of brain nodes are now

beginning to appear. The constructionist approach championed by Lindquist and Barrett (2012) is an effort to shift the conceptual mode of attack on the mind-brain problem. Instead of the traditional effort to selectively map cognitive modules onto localized brain regions, they propose a "constructionist" approach in which there are no operationally committed macroneural nodes, only broadly distributed, multioperational regions that are recruited as necessary to represent or encode cognitive processes. Should this reconceptualization be accurate, much of the work done previously on mapping cognitive processes to specific brain structures (nodes) would become irrelevant. Of course, this new conceptualization is also subject to many of the practical and conceptual problems discussed already.

Furthermore, not all new developments in statistical methodology solve preexisting problems; some may actually exacerbate the problems. The poor signal-to-noise ratios characteristic of brain image data has generated a need for evermore powerful statistical procedures to extract the signals. It is not uncommon, however, for novel techniques designed to overcome spurious measurement artifacts (such as head movements and breathing) to generate new difficulties. Band pass filtering and statistical comparisons with known sources of noise are now sequentially being applied to remove noise from fMRI signals (Hallquist, Hwang, and Luna, in Press). However, this dual correction produces results that depend on the order in which the two corrective steps are applied to the data and, therefore, may lead to not only false data and ultimately false conclusions.

Other new developments have been proposed as means to understanding neuronal networks. For example, Optogenetics (Deisseroth, 2011) has been offered as a method for activating "cell-type specific" patterns of brain activity based on genetic peculiarities by making *types* of cells responsive to light. However, this powerful technique does not, as sometimes implied (e.g., the use of ambiguous phrases "specific optogenetic inhibition of excitatory neurons" by Goshen et al, 2011), permit the idiosyncratic activation or inhibition of individual neurons in the manner required by a putative microneural theory. Instead, this method is limited to synchronously activating many neurons sharing a common neurochemistry. In other words, optogenetics is, like the fMRI, a macroneural approach that does not permit the simultaneous control of individual neurons.

Other approaches attempt to overcome our inability to study microlevel organization by applying mathematical techniques to pool idiosyncratic neuronal responses. Mean-field theory (Arbib, 2003), for example, creates a single representative value to represent the cumulative or averaged activity of the involved neurons. In other words, it converts an intractable many-body (microneural) problem into a one-body (macroneural) problem that is supposedly capable of solution. Similar techniques are commonly used in physics to overcome the numerousness problem by pooling individual responses; however, the idea works there because the interacting force (gravity) is uniform. This simplifying property is not available in real neuronal networks where each neuron and each synapse is responding in a plethora of different patterns of interaction.

55

Empirical support for the postulate that the essence of the mind-brain issue lies at the microneural level of the states of individual neurons is, as we have seen, difficult to obtain but logically compelling. Attention at the microneural level has, therefore, turned in distinctive ways to neuronal networks. Cognition, according to this approach, arises from the integrated, collective, action of many neurons. Neurons, the postulate asserts, maintain their individual operational integrity despite the fact that many of them have to be linked together to represent a cognitive process. To understand neural networks from this point of view, we must preserve the properties of the individual neurons and their interconnections—an almost impossible task.

If a direct neurophysiological approach to the study of cognitively plausible neuronal networks is exceedingly difficult, the main alternative is computer simulation. This approach aims to use what we know of the microanatomy of the brain to construct plausible computational models of what necessarily has to be a simplified simulated neuronal network. A recent review of this approach (de Garis et al, 2010) points out that simulations are now becoming possible in which the action and interaction of millions of neurons can be evaluated. However, they also point out that, with the exception of a few locales, we do not know enough about the actual organization of the brain to make these simulations neuroscientifically significant. This has become a highly controversial matter with some investigators challenging the biological relevance of very large networks of detailed simulations of neurons. "Neuroscientifically insignificant" does not mean, however, that these simulations are not valuable. An enormous amount has been learned about the interactions of complex networks in general. However, simulation of large scale neuronal networks suffers from many of the practical combinatorial problems precluding direct physiological research on real neuronal networks—how should the parameters of each neuron be set and how can their coordinated activity be measured?

An alternative and essentially macroneural way to approach computational simulation of brain processes that does depend on cognitive processes is what Goertzel et al (2010) refer to as Biologically Inspired Cognitive Architecture (BICA). Rather than operating at the microneural level, BICA models are much more closely linked to psychology using modules of cognition such as learning, attention, and working memory as the elements of their simulations. This type of model, long familiar to psychology in the form of block diagram models, is enabled in this new world of supercomputers by their ability to more powerfully and definitively evaluate the plausibility of a descriptive explanation of cognition than was previously possible. Unfortunately, this kind of computer simulation reinforces what may well be an incorrect emphasis on "chunks" of the brain and "modules" of the brain just as the microneural networks of large scale computers direct our attention to that kind of theory. Both of these approaches, I suggest, are examples of technology forcing theory.

Despite the promise each of these two computational approaches has to their respective fields, there is a vast gulf of ignorance regarding how the two levels of simulation (microneural networks and cognitive module networks) might be connected. As Goertzel and his colleagues note:

One fairly obvious phenomenon worth drawing attention to is the gap between the BICA and brain simulation [microneuronal network] approaches. Both approaches seek to leverage recent advances in hardware and neuroscience to create artificial brains, but so far the two approaches display very different strengths (p. 47-48).

Conclusions

This review of the current status of neuroreductionist explanation in cognitive neuroscience highlights some of the formidable barriers that obstruct our hopes of building a comprehensive, overarching, neuroreductionist, physicalist account of how mental activities emerge from brain activities. It is not yet known which of these barriers, if any, might prove to be surmountable and which may permanently obstruct our efforts. However, they collectively raise serious questions about whether or not a neuroreductionist explanation of mind-brain relations will ever be achieved. At the very least, it appears that the current idea of mapping the brain into a pattern of localized, operation-specific, macroneural, activation regions is less and less supportable. This may be true only in the genetic older or peripheral portions of the nervous system, but it certainly seems inappropriate in the context of those brain operations that underlay cognition. On the other hand, practical problems, complexity, and combinatorics probably preclude the microneural approach that seems logically more plausible.

Speaking about the future of the macroneural approach, Poldrack (2010) asserted that it may be an impossible task in its present form. He noted that "a review of the neuroimaging literature suggests that selective association between mental processes and brain structures is currently impossible to find." (p. 754)

By "selective association" he is referring to unique associations between a cognitive process and one or more a brain areas. Poldrack went on, "… nearly every such claim [of a selective association] uncovers counterexamples that are difficult to reconcile with a selective structure-function mapping." (p. 755)

To move forward, Poldrack suggested that we must develop "ontologies" or "taxonomies" of cognitive processes (i.e., better definitions), explicitly accepting the fact that our previous attempts at classifying cognitive processes were inadequate. Considering that psychology has been unsuccessfully attempting to do this for centuries, "ontological" development is likely to be an elusive goal.

If a comprehensive explanation of how cognitive process emerge from brain activity is unachievable in the near future, can we at least hope for some kind of predictive and descriptive prototheory through the use of macroneural measures? There will always be loose metaphors, hypothetical constructs, and behavior-producing simulations that mimic cognitive processes; but a necessary and explanatory neural theory of mental activity is not on the horizon. The evidence so far and the portents for the future simply do not justify any optimism concerning an answer to the most important of human questions—how does brain activity become mental activity?

The best we can do, in my personal opinion, is to hope for some descriptions, metaphors, or other kinds of proto-theories. We obviously will make progress in many different aspects of cognitive neuroscience and invent interesting metaphors and observe correlations. However, a truly neuroreductive explanation of how the mind emerges from the brain in the sense achieved in cosmological theory is not yet available and is not likely to become available without some kind of a major conceptual, empirical, or methodological breakthrough. The common misunderstanding that correlation alone can serve as the basis for a theory is probably as misleading today as it was in Yule's (1899). A solution to the mind-brain conundrum remains as elusive today as it has been for millennia.

What has this to say about the optimum future course of psychology? Although it is impossible to predict the future, the ensemble of difficulties with neuroreductionism suggests that for the time being our most advantageous approach to the problem of understanding psychology will come not from a fruitless neuroreductionist effort but from a science with the following properties or approaches.

Mathematically descriptive: Use formal methods to describe and predict cognitive activity with the understanding that mathematical models can be sufficient but not necessary.

Neuronally nonreductive: Eschew neural explanations of cognition given the many conceptual and empirical problems highlighted in this article. Most cognitive neuroscientists accept the observation of scholars such as Koch and Greenfield that we have "utterly failed to satisfactorily explain how subjective experience is created" (2007, p. 76) and Sporn that "neuroscience still cannot answer the 'big questions' about mind and intelligence" (2011, p. 179).

Cognitively nonreductive: Acknowledge that hypothetical constructs and cognitive modules are underdetermined (MacCorquodale and Meehl, 1948). The idea of parsing psychological activity into modular components may be an experimental convenience, but an impediment to progress.

Empirical: Concentrate on the overt and measurable behavioral correlates and properties of cognitive activity rather than the inferred parameters.

In conclusion, in this article I have pointed out that there are two main empirical approaches to mind-brain explanation. The first is the very popular macroneural approach based on cumulative data forthcoming from brain images. The second is the microneural approach based on the Hebbian conjecture of synaptic changes among neurons. Both of these approaches are currently challenged by major barriers and difficulties, many of which appear to be extremely difficult, if not impossible, to overcome. This should not be interpreted to mean that we should cease to study the brain at all levels and using all approaches; there is much to be learned about both cognition and brain operation. However, no clear trajectory about how to build a bridge between mind and brain has yet emerged. For the moment at least, the answer to the question asked in the title of this article is certainly "not now."

References

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *NeuroImage, 8*(4), 360-369.

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33*(4), 245-+. doi: Doi 10.1017/S0140525x10000853

Arbib, M. A. (Ed.). (2002). *The handbook of brain theory and neural networks (2nd Ed.)*. Cambridge MA: Bradford Books.

Aron, A. R., Gluck, M. A., & Poldrack, R. A. (2006). Long-term test-retest reliability of functional MRI in a classification learning task. *NeuroImage, 29*(3), 1000-1006. doi: DOI 10.1016/j.neuroimage.2005.08.010

Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychological Review, 114*(3), 632-656. doi: Doi 10.1037/0033-295x.114.3.632

Aue, T., Lavelle, L. A., & Cacioppo, J. T. (2009). Great expectations: What can fMRI research tell us about psychological phenomena? *International Journal of Psychophysiology, 73*(1), 10-16. doi: DOI 10.1016/j.ijpsycho.2008.12.017

Bandettini, P. (2007). Functional MRI today. *International Journal of Psychophysiology, 63*(2), 138-145. doi: DOI 10.1016/j.ijpsycho.2006.03.016

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

Bechtel, W. (2002). Decomposing the mind-brain: A long term pursuit. *Brain and Mind, 3,* 229-242.

Belliveau, J. W. et al (1991). Functional mapping of the human visual cortex by magnetic resonance imaging, *Science, 254,* 716-719.

Bennett, C. M., & Miller, M. B. (2010). How reliable are the results from functional magnetic resonance imaging? *Year in Cognitive Neuroscience 2010, 1191*, 133-155. doi: DOI 10.1111/j.1749-6632.2010.05446.x

Bennett, C. M., Baird, A. A., Miller, M. B., & Wolford, G. L. (2011). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for proper multiple comparisons correction. *Journal of Serendipitous and Unexpected Results, 1*, 1-5.

Bennett, C. M., Wolford, G. L., & Miller, M. B. (2009). The principled control of false positives in neuroimaging. *Social Cognitive and Affective Neuroscience, 4*(4), 417-422. doi: Doi 10.1093/Scan/Nsp053

Bickle, J. (2003), Philosophy and Neuroscience: *A Ruthlessly Reductive Account*, Dordrecht: Kluwer Academic Press.

Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences, 14*(6), 277-290.

Brindley, G. S. (1960). *Physiology of the retina and the visual pathway*. London: Edward Arnold.

Carp, J. (2012a). The secret lives of experiments: Methods reporting in the fMRI literature. *NeuroImage, 63*(1), 289-300. DOI 10.1016/j.neuroimage.2012.07.004

Carp, J. (2012b). On the plurality of (methodological) worlds: Estimating the analytic flexibility of fMRI experiments. *Frontiers in Neuroscience, 6*, 1-13.

Chein, J. M., Fissell, K., Jacobs, S., & Fiez, J. A. (2002). Functional heterogeneity within Broca's area during verbal working memory. *Physiology & Behavior, 77*(4-5), 635-639.

Chung, K., & Deisseroth, K. (2013). Clarity for mapping the nervous system. *Nature Methods, 10*, 508-513.

Churchland, P. S. (1994). Can neurobiology teach us anything about consciousness? *Proceedings and Addresses of the American Philosophical Association, 67,* 23-40.

de Garis, H., Chen, S., Goertzel, B., & Lian, R. T. (2010). A world survey of artificial brain projects, Part I: Large-scale brain simulations. *Neurocomputing, 74*(1-3), 3-29. DOI 10.1016/j.neucom.2010.08.004

Deisseroth, K. (2011). Optogenetics. *Nature Methods, 8*, 26-29.

Dietrich, E., & Hardcastle, V. G. (2005). *Sisyphus's boulder: consciousness and the limits of the knowable*. Amsterdam: John Benjamins Publishing.

Duhem, P. (1914) *La théorie physique: son objet et sa structure*. Paris: Marcel Riviera & Cie.

Duncan, K. J., Pattamadilok, C., Knierim, I., & Devlin, J. T. (2009). Consistency and variability in functional localisers. *NeuroImage, 46*(4), 1018-1026. DOI 10.1016/j.neuroimage.2009.03.014

Eklund, A., Andersson, M., Josephson, C., Johannesson, M., & Knutsson, H. (2012). Does parametric fMRI analysis with SPM yield valid results? An empirical study of 1484 rest datasets. *NeuroImage, 61,* 565-578.

Feigl, H. (1958). The mental and the physical. In H. Feigl et al (Eds.), *The Minnesota studies in the philosophy of science (vol. II), concepts, theories, and the mind-body problem*. Minneapolis, MN: University of Minnesota Press.

Fiedler, K. (2011). Voodoo correlations are everywhere-not only in neuroscience. *Perspectives on Psychological Science, 6*(2), 163-171. DOI 10.1177/1745691611400237

Friston, K. J., Rotshtein, P., Geng, J. J., Sterzer, P., & Henson, R. N. (2010). A critique of functional localizers. In S. J. Hanson & M. Bunzl (Eds.), *Foundation Issues in Human Brain Mapping* (pp. 3-24). Cambridge MA: MIT Press.

Gelman, A., & Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *American Statistician, 60*(4), 328-331. DOI 10.1198/000313006x152649

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage, 15*(4), 870-878. DOI 10.1006/nimg.2001.1037

Glahn, D. C., Ragland, J. D., Abramoff, A., Barrett, J., Laird, A. R., Bearden, C. E., et al. (2005). Beyond hypofrontality: A quantitative meta-analysis of functional neuroimaging studies of working memory in schizophrenia. *Human Brain Mapping, 25*(1), 60-69.

Goertzel, B., Lian, R. T., Arel, I., de Garis, H., & Chen, S. (2010). A world survey of artificial brain projects, Part II: Biologically inspired cognitive architectures. *Neurocomputing, 74*(1-3), 30-49.

Goldfine, A. M., Bardin, J. C., Noirhomme, Q., Fins, J. J., Schiff, N. D., & Victor, J. D. (2013). Reanalysis of "Bedside detection of awareness in the vegetative state: a cohort study". *Lancet, 381*(9863), 289.

Gonzales-Castillo, J., Saad, Z. S., Handwerker, D. A., Inati, S. J., Brenowitz, N., & Bandettini, P. A. (2012). Whole-brain, time-locked activation with simple tasks revealed with massive averaging and model-free analysis. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 5487-5492.

Goshen, I., Brodsky, M., Prakash, R., Wallace, J., Gradinaru, V., Ramakrishnan, C., & Deisseroth, K. (2011). Dynamics of retrieval strategies for remote memories. *Cell, 147*(5), 1197-1197. DOI 10.1016/j.cell.2011.11.006

Greenwald, A. G., Gonzalez, R., Harris, R. J., & Guthrie, D. (1996). Effect sizes and p values: What should be reported and what should be replicated? *Psychophysiology, 33*(2), 175-183.

Hallquist, M. N., Hwang, K., & Luna, B. (In Press). The nuisance of nuisance regression: Spectral misspecification in a common approach to resting state of fMRI preprocessing reintroduces noise and obscure functional connectivity.

He, B. J. (2013). Spontaneous and task-evoked brain activity negatively interact. *Journal of Neuroscience, 33*(11), 4672-4682.

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory*. New York, NY: Wiley.

Heil, J. (2003). Mental causation. In S. P. Stich & T. A. Warfield (Eds.), *The Blackwell Guide to the philosophy of mind* (pp. 213-234). Malden MA: Blackwell.

Hilgetag, C. C., ONeill, M. A., & Young, M. P. (1996). Indeterminate organization of the visual system. *Science, 271*(5250), 776-777.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70.

Ihnen, S. K., Church, J. A., Petersen, S. E., & Schlaggar, B. L. (2009). Lack of generalizability of sex differences in the fMRI BOLD activity associated with language processing in adults. *NeuroImage, 45*(3), 1020-1032.

Joppa, L. N., McInerny, G., Harper, R., Salido, L.,…, Emmott, S. (2013). Troubling trends in scientific software use. *Science, 340*(6134), 814-815.

Karp, R. M. (1986). Combinatorics, Complexity, and Randomness. *Communications of the ACM, 29*(2), 98-109. DOI 10.1145/5657.5658

Kim, J. (1993). *Supervenience and mind*. New York, NY: Cambridge University Press.

Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T. D. (2008). Functional grouping and cortical-subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *NeuroImage, 42*(2), 998-1031.

Koch, C. (2012). Modular biological complexity. *Science, 337*, 531-532.

Koch, C. & Greenfield, S. (2007). How does consciousness happen? *Scientific American, 297(4),* 76-83. DOI 10.1145/5657.5658

Krain, A. L., Wilson, A. M., Arbuckle, R., Castellanos, F. X., & Milham, M. P. (2006). Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making. *Neuroimage, 32*(1), 477-484.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience, 12*(5), 535-540. DOI 10.1038/Nn.2303

Lambdin, C. (2012). Significance tests as sorcery: Science is empirical-significance tests are not. *Theory & Psychology, 22*(1), 67-90. DOI 10.1177/0959354311429854

Lichtman, J. W., & Denk, W. (2011). The Big and the Small: Challenges of imaging the brain's circuits. *Science, 334*(6056), 618-623. DOI 10.1126/science.1209168

Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: Re-balancing the scale. *Social Cognitive and Affective Neuroscience, 4*(4), 423-428. DOI 10.1093/Scan/Nsp052

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*(11), 533-540. DOI 10.1016/j.tics.2012.09.005

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412*(6843), 150-157.

Loring, D. W., Meador, K. J., Allison, J. D., Pillai, J. J., Lavin, T., Lee, G. P., . . . Dave, V. (2002). Now you see it, now you don't: Statistical and methodological considerations in fMRI. *Epilepsy & Behavior, 3*(6), 539-547.

Ludwig, K. (2003). The mind-body problem: An overview. In S. P. Stich & T. A. Warfield (Eds.), *The Blackwell Guide to the philosophy of mind* (pp. 1-46). Malden MA: Blackwell.

Lund, T. E., Norgaard, M. D., Rostrup, E., Rowe, J. B., & Paulson, O. B. (2005). Motion or activity: Their role in intra- and inter-subject variation in fMRI. *NeuroImage, 26*(3), 960-964. DOI 10.1016/j.neuroimage.2005.02.021

MacCorquodale, K. & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review, 55, 95-107*.

Maitra, R., Roys, S. R., & Gullapalli, R. P. (2002). Test-retest reliability estimation of functional MRI data. *Magnetic Resonance in Medicine, 48*(1), 62-70. DOI 10.1002/Mrm.10191

Martin, S. J., Grimwood, P. D., & Morris, R. G. M. (2000). Synaptic plasticity and memory: An evaluation of the hypothesis. *Annual Review of Neuroscience, 23,* 649-711.

McGinn, C. (1989). Can we solve the mind body problem. *Mind, 98*(391), 349-366.

McGonigle, D. J. (2012). Test-retest reliability in fMRI: Or how I learned to stop worrying and love the variability. *NeuroImage, 62*(2), 1116-1120. DOI 10.1016/j.neuroimage.2012.01.023

McGonigle, D. J., Howseman, A. M., Athwal, B. S., Friston, K. J., Frackowiak, R. S. J., & Holmes, A. P. (2000). Variability in fMRI: An examination of intersession differences. *NeuroImage, 11*(6), 708-734. DOI 10.1006/nimg.2000.0562

Meehan, T. P., & Bressler, S. L. (2012). Neurocognitive networks: Findings, models, and theory. *Neuroscience and Biobehavioral Reviews, 36*, 2232-2247.

Meehan, T. P., & Bressler, S. L. (2012). Neurocognitive networks: Findings, models, and theory. *Neuroscience and Biobehavioral Reviews, 36*(10), 2232-2247. DOI 10.1016/j.neubiorev.2012.08.002

Mikl, M., Marecek, R., Hlustik, P., Pavlicova, M., Drastich, A., Chlebus, P., . . . Krupa, P. (2008). Effects of spatial smoothing on fMRI group inferences. *Magnetic Resonance Imaging, 26*(4), 490-503. DOI 10.1016/j.mri.2007.08.006

Miller, M. B., Van Horn, J. D., Wolford, G. L., Handy, T. C., Valsangkar-Smyth, M., Inati, S., . . . Gazzaniga, M. S. (2002). Extensive individual differences in brain activations associated with episodic retrieval are reliable over time. *Journal of Cognitive Neuroscience, 14*(8), 1200-1214.

Moore, E. F. (1956). Gedanken-experiments on sequential machines. In C. E. Shannon & J. McCarthy (Eds.), *Automata Studies* (pp. 129-156). Princeton: Princeton University Press.

Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., & Malach, R. (2005). Coupling between neuronal firing, field potentials, and fMR1 in human auditory cortex. *Science, 309*(5736), 951-954. DOI 10.1126/science.1110913

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E. J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nature Neuroscience, 14*(9), 1105-1107. DOI 10.1038/Nn.2886

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic-resonance-imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America, 87*(24), 9868-9872.

Oosterwijk, S., Lindquist, K. A., Anderson, E., Dautoff, R., Moriguchi, Y., & Barrett, L. F. (2012). State of mind: emotions, body feelings, and thoughts share distributed neural networks. *NeuroImage 62(3)*, 2110-2128.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging. *Human Brain Mapping, 25*(1), 46-59. DOI 10.1002/Hbm.20131

Palmer, S. E. (1999). *Vision science: photons to phenomenology*. Cambridge MA: MIT.

Poldrack, R. A. (2010). Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on Psychological Science, 5*(6), 753-761. DOI 10.1177/1745691610388777

Putnam, H. (1975). *Mind, language, and reality*. New York: Cambridge University Press.

Quine, W. V. O. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *Philosophical Review, 60, 20-43.*

Rakover, S. S. (2011). A plea for methodological dualism and a multiexplanation framework in psychology. *Behavior and Philosophy, 39*, 17-43.

Rau, S., Fesl, G., Bruhns, P., Havel, P., Braun, B., Tonn, J. C., & Ilimberger, J. (2007). Reproducibility of activations in Broca area with two language tasks: A functional MR imaging study. *American Journal of Neuroradiology, 28*(7), 1346-1353. DOI 10.3174/Ajnr.Ao581

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localizers. *NeuroImage, 30*(4), 1088-1096.

Schroll, H., Vitay, J., & Hamker, F. H. (2012). Working memory and response selection: A computational account of interactions among cortico-basalganglio-thalamic loops. *Neural Networks, 26*, 59-74. doi: DOI 10.1016/j.neunet.2011.10.008

Simmonds, D. J., Pekar, J. J., & Mostofsky, S. H. (2008). Meta-analysis of go/no-go tasks, demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia, 46*(1), 224-232. DOI 10.1016/j.neuropsychologia.2007.07.015

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359-1366. DOI 10.1177/0956797611417632

Skinner, B. F. (1950). Are Theories of Learning Necessary. *Psychological Review, 57*(4), 193-216. DOI 10.1037/H0054367

Smith, S. M., Beckmann, C. F., Ramnani, N., Woolrich, M. W., Bannister, P. R., Jenkinson, M., McGonigle, D. J. (2005). Variability in fMRI: A re-examination of inter-session differences. *Human Brain Mapping, 24*(3), 248-257. DOI 10.1002/Hbm.20080

Sporns, O. (2011). Networks of the Brain. *Cambridge, MA, MIT Press*.

Stam, C. J., & van Straaten, E. C. W. (2012). The organization of physiological brain networks. *Clinical Neurophysiology, 123*(6), 1067-1087. DOI 10.1016/j.clinph.2012.01.011

Stockmeyer, L. J., & Chandra, A. K. (1979). Intrinsically difficult problems. *Scientific American, 240*(5), 140-159.

Thyreau, B. et al (2012). Very large fMRI study using the IMAGEN database: Sensitivity-specificity and population effect modeling in relation to the underlying anatomy. *NeuroImage, 61,* 295-303.

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *Neuroimage, 16*(3), 765-780. DOI 10.1006/nimg.2002.1131

Uttal, W. R. (2011). *Mind and Brain: A Critical Appraisal of Cognitive Neuroscience*. Cambridge MA: MIT Press.

Uttal, W. R. (2013). *Reliability in Cognitive Neuroscience: A Meta-meta Analysis*. Cambridge MA: MIT Press.

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information-processing in the primate visual-system - an integrated systems perspective. *Science, 255*(5043), 419-423.

Vimal, R. L. P. (2009). Meanings Attributed to the Term 'Consciousness': An overview. *Journal of Consciousness Studies, 16*(5), 9-27.

Vul, E., & Kanwisher, N. (2010). Begging the question: The nonindependence error in fMRI data analysis. In S. J. Hanson & S. M. Bunzi (Eds.), *Foundational issues for human brain mapping*. (pp. 71-91) Cambridge MA: MIT Press.

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science, 4*(3), 274-290. DOI 10.1111/j.1745-6924.2009.01125.x

Wager, T. D., Lindquist, M. A., Nichols, T. E., Kober, H., & Van Snellenberg, J. X. (2009). Evaluating the consistency and specificity of neuroimaging data using meta-analysis. *NeuroImage, 45*(1), S210-S221. DOI 10.1016/j.neuroimage.2008.10.061

Wager, T. D., & Smith, E. E. (2003 b). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective and Behavioral Neuroscience, 3*(4), 255-274. DOI 10.1016/S1053-8199(03)00078-8

Wark, H. A., Sharma, R., Mathews, K. S., Fernandez, E., Yoo, J., Christensen, B., Tathireddy, P. (2013). A new high-density (25 electrodes/mm) penetrating microelectrode array for recording and stimulating sub-millimeter neuroanatomical structures. *Journal of Neural Engineering, 10*(4), 045003. DOI 10.1088/1741-2560/10/4/045003

Westin, C.-F., Maier, S. E., Mamata, H., Nabavi, A., Jolesz, F. A., & Kikinis, R. (2002). Processing and visualization for diffusion tensor MRI. *Medical Image Analysis, 6*, 93-108.

Yule, G. U. (1899). An investigation in the causes of pauperism in England, chiefly during the last two intercensal decades. *Journal of the Royal Statistical Society, 62*, 249-295.