

Technical Report 1337

Macroscopic Neural Theories of Cognition

William R. Uttal
Arizona State University

March 2014



**United States Army Research Institute
for the Behavioral and Social Sciences**

Approved for public release; distribution is unlimited.

**U.S. Army Research Institute
for the Behavioral and Social Sciences**

**Department of the Army
Deputy Chief of Staff, G1**

Authorized and approved for distribution:

**MICHELLE SAMS, Ph.D.
Director**

Research accomplished under contract
for the Department of the Army by:

Arizona State University

Technical review by:

David Costanza, The George Washington University
Paul Gade, The George Washington University

NOTICES

DISTRIBUTION: Primary distribution of this Technical Report has been made by ARI. Address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: DAPE-ARI-ZXM, 6000 6th Street (Bldg 1464 / Mail Stop: 5610), Fort Belvoir, Virginia 22060-5610

FINAL DISPOSITION: Destroy this Technical Report when it is no longer needed. Do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE				<i>Form Approved</i> OMB No. 0704-0188	
1. REPORT DATE (DD-MM-YYYY) March 2014		2. REPORT TYPE Final		3. DATES COVERED (From - To) September 2012 – December 2013	
4. TITLE AND SUBTITLE Macroscopic Neural Theories of Cognition				5a. CONTRACT NUMBER W5J9CQ-12-C-0033	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 611102	
6. AUTHOR(S) William R. Uttal				5d. PROJECT NUMBER B74F	
				5e. TASK NUMBER 4901	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Arizona State University 660 S Mill Ave Ste 312 Tempe, AZ 85281-3670				8. PERFORMING ORGANIZATION REPORT	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Institute for the Behavioral & Social Sciences 6000 6 th Street (Building 1464 / Mail Stop 5610) Fort Belvoir, VA 22060-5610				10. SPONSOR/MONITOR'S ACRONYM(S) ARI	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) Technical Report 1337	
12. DISTRIBUTION/AVAILABILITY STATEMENT: Distribution Statement A: Approved for public release; distribution unlimited.					
13. SUPPLEMENTARY NOTES Contracting Officer's Representative POC: Dr. Gerald F. Goodwin & Dr. Lynne M. Rochette					
14. ABSTRACT This article explores the question, "are the available conceptual foundations, the statistical techniques, and the empirical data of cognitive neuroscience sufficiently robust to serve as a foundation for a neuroreductionist explanatory theories of the mind-brain relationship?" The answer proposed here is that there is insufficient reason to think that such a theory is likely to emerge in the near future. In this context, a theory means a comprehensive explanation of how cognitive or mental processes are explained by neural mechanisms. There are two main contenders for this kind of theory building. The first is based on macroneural measurements such as those produced by brain imaging equipment. The alternative is a microneural approach epitomized by the Hebb (1949) Conjecture. The Hebbian approach asserts that it is most likely that cognitive processes of all kinds are mediated by the concurrent collective microneural state (not the summed, pooled, or accumulated macroneural properties) of an assembly of individual neurons. Although there is insufficient evidence at the microneural level to robustly support this conjecture, difficulties with the macroneural approach and the logic of the microneural approach suggest that Hebb was essentially correct. Thus, overarching theories of the mind-brain relation are unlikely in the foreseeable future.					
15. SUBJECT TERMS Neuroscience, Mind-Brain, Brain Imaging, Theory, Cognition					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Unlimited Unclassified	18. NUMBER OF PAGES 47	19a. NAME OF RESPONSIBLE PERSON Dorothy Young
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER 703-545-2316

Technical Report 1337

Macroscopic Neural Theories of Cognition

William R. Uttal
Arizona State University

Foundational Science Research Unit
Gerald F. Goodwin, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
6000 6th Street, Building 1464
Fort Belvoir, VA 22060

March 2014

Approved for public release; distribution is unlimited.

MACROSCOPIC NEURAL THEORIES OF COGNITION

EXECUTIVE SUMMARY

Research Requirement:

This article is the final report for the research project entitled “*Macroscopic Neural Theories of Cognition*” funded under BAA W5J9CQ-11-R-2000 Part B Basic Research. The general goal of this research was to review the current status of theory in cognitive neuroscience, particularly those macroneural theories influenced by image processing devices such as the fMRI. The specific task was to determine what appear to be possible theories of a mind-brain relationship and discuss the advantages and disadvantages of each. The ultimate goal of the project was to draw a comprehensive appraisal of current cognitive neuroscience theory so that a tentative answer could be made to the question, “are the available conceptual foundations, the statistical techniques, and the empirical data of cognitive neuroscience sufficiently robust to serve as a foundation for a neuroreductionist explanatory theories of the mind-brain relationship?”

Procedure:

This project was primarily a literature search and review effort followed by a synthesis of the reviewed material into a publishable estimate of the current state of macroneural theory.

Findings:

The main finding emerging from this research project was the author’s judgment that the links between neurophysiological and psychological are still not a level that would justify their use as an alternative to behavioral research. Despite wide-spread misunderstanding about how far cognitive neuroscience has gone, no overarching theory of how neural processes produce behavioral and mental responses is yet at hand nor is any foreseeable in the future.

Utilization and Dissemination of Findings:

At the present time the deliverable article is under review by a journal for publication. In addition, a manuscript of a book (partially written at the time this project commenced) was completed. It presents a more complete discussion of the topics necessarily abbreviated in the preparation of the article.

MACROSCOPIC NEURAL THEORIES OF COGNITION

CONTENTS

	Page
MACROSCOPIC NEURAL THEORIES OF COGNITION.....	1
CONCEPTUAL BARRIERS TO THEORY BUILDING.....	2
Levels of Analysis.....	2
Complexity.....	3
Loss of Information with Response Pooling.....	5
Underdetermination	6
Poor Definition of Cognitive Constructs	6
Inaccessibility	7
The Search for Objectivity.....	8
False Analogies between Sensory, Motor, and Cognitive Processes.....	8
The Neurologizing of Psychological Language.....	9
STATISTICAL BARRIERS TO THEORY BUILDING.....	9
General Problems.....	10
The Misleading Null Hypothesis Significance Test and $p < .05$	11
Ignoring the Number of Tests	12
Sample Size.....	13
Double Dipping.....	14
Erroneously Comparing Significances	15
Other Statistical Issues.....	15
EMPIRICAL BARRIERS TO THEORY BUILDING.....	16
The Case against Localization	16
On the reliability of fMRI Data	18
Technical Artifacts.....	23
NEW DEVELOPMENTS; CAN THEY PROVIDE A PATHWAY TO NEUROREDUCTIVE THEORY?	23
New Macroneural Developments	24
New Microneural Developments	26
CONCLUSIONS.....	28
REFERENCES	30

MACROSCOPIC NEURAL THEORIES OF COGNITION

In the last three decades, traditional physiological psychology has undergone a revolution, not only in name (it is now generally known as cognitive neuroscience), but also in substance and available technology. With the invention of functional Magnetic Resonance Imaging (fMRI) by Ogawa, Lee, Kay, and Tank (1990), brain images using the Blood Oxygen Level Dependent (BOLD) measures have largely replaced traditional methods involving surgery, trauma, stimulation, and electrophysiological recording.

Currently, the number of papers that deal specifically with comparisons of cognitive processes and fMRI brain images is difficult to determine because not all articles using this method are related to what is currently referred to as the mind-brain problem (some have been directed quite successfully to problems in neuroanatomy or neurophysiology.) 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 they are brain responses, more or less correlated with cognitive processes, there is a compelling face validity presented to researchers by these new machines. The promise is that we are on the verge of a breakthrough in the profound and hitherto intractable mind-brain problem: How does the activity of tangible brain tissue produce intangible cognitive processes? The hope persists that we will be able to meld what turns out to be an enormous database into a coherent, valid, and comprehensive theory of mind-brain relations. This face validity 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 to their interrelationship. At the least, it has been expected that correlations could be found between cognitive and neural responses that might serve as heuristics for plausible future theories of mind-brain relationships.

The hope of an overarching theory of mind-brain relations remains largely that—an unfulfilled hope. Despite the large number of publications and the large number of investigators working with the fMRI tool, no one has yet been able to provide a satisfactory answer to the challenge proposed by this classic conundrum. Perhaps this is because this approach is so new; only a little more than two decades have passed since the fMRI method became available. However, there is another possibility, namely that there exist a number of conceptual, statistical, and empirical barriers that suggest that the mind-brain problem may not be solvable at the macroneural level of analysis opened to us by fMRI systems. Despite the fact that both cognitive processes and fMRI images are products of the brain and there is an extensive corpus of comparative results, theoretical progress from the raw data to even partial explanatory statements remains rare and speculative. There remains considerable doubt, buttressed by a substantial and increasing amount of critical research, that the relation between brain images and cognitive process is not going to be solved by even the most ingenious applications of macroneural technology. Many philosophers have adopted a conservative point of view with regard to this problem. Rakover (2011), a psychologist-philosopher, lists a mainstream group of philosophers including Putnam (1975), McGinn (1989), Palmer (1999), Ludwig (2003), Heil (2003), and Dietrich and Hardcastle (2005) as all having arrived at this same conclusion.

Conceptual Barriers to Theory Building

Why are theories in cognitive neuroscience, particularly those based on macroneural measurements, so difficult to construct and confirm? To answer such a question requires consideration of some of the factors that inhibit progress in this kind of theory development, including conceptual barriers that confront anyone who has the temerity to attack this most challenging of scientific theoretical enterprises—explaining the relationship between the mind and the brain.

Levels of Analysis

A major barrier to theory building concerns the level at which a problem is being studied. Level of analysis refers to both the scale at which brain activity becomes mental activity and the scale of analyses and measurements. In the present context, a macroneural level of analysis deals with signals that are pools or accumulations of these lower level responses such as fMRI images or electroencephalogram (EEG) recordings. A microneural level of analysis entails investigating the responses of cellular and subcellular units such as neurons and synapses and their respective interactions.

Microneural interactions are the bases of the Hebb conjecture (Hebb, 1949)—a conceptualization that assumes that cognitive processes emerge from the idiosyncratic action of millions (if not billions) of neurons and the thousand fold greater number of synaptic connection. These processes are not measured in terms of their cumulative or summarized values, but instead to the degree they maintain their individual identity in the neuronal network. Despite its conceptual robustness and near universal acceptance, the main problem with the Hebb hypothesis is that it has no empirical basis (and probably never will) given the practical problems preventing simultaneous and independent manipulation of a cognitively significant number of neurons. To establish the validity of the Hebb model would require an unobtainable amount of individual control over an uncountable number of neurons. A microneural experiment would require a brute force empirical approach. Ideally, the brute force technique would require the following preliminary instrumentation steps be accomplished:

1. Identify the neurons and their interconnections that are involved in a cognitive process.
2. Attach a recording device to every neuron in the network.
3. Attach a stimulating device to every neuron in the network.

Obviously, achieving these connections would be a technical tour de force of the first magnitude given that our ability to develop microelectrode arrays suitable for stimulation or recording has so far been limited to a few hundred. The difficulty of determining specifically which neuron is associated with which microelectrode is also daunting. Furthermore, the amount of data obtained by even a few hundred microelectrodes would tax available computational capabilities.

For a host of procedural and instrumentation reasons, neither the instrumental task nor the experimental paradigm can be carried out. For one, it has not proven possible to record

simultaneously and independently from anywhere near the number of neurons that are involved in cognitive processes. Nor are we able to selectively stimulate an identified group of neurons to determine if we can mimic, block, or change a cognitive process.

Because of these technical and procedural limitations, the brute force, microneural strategy of stimulating and recording from a microneural neural network is not implementable and can only be conceived of as a hypothetical experiment. Even if by some magic we were able to produce the necessary technical instruments, constraints on data accumulation and analysis, would overwhelm any conceivable computer system for simple combinatoric reasons. Thus there is a glaring inconsistency between the microneural level at which the Hebb conjecture operates and the macroneural level of brain imaging technology.

In short, microneural theory building in such a many-neuron environment is severely inhibited for two very practical reasons. First, we are overwhelmed by so many unobtainable technical requirements and so much data that it is, for practical reasons, computationally unavailable to us. We have no way of both measuring and regenerating the vast amount of activity in the neuronal network whose states correspond to cognitive processes. Second, at the macroneural situation typified by fMRI images, the necessary microneural data is inaccessible because it no longer exists—it has been averaged out of existence.

Thus, at both the microneural and macroneural levels, the possibility of developing an authentic, valid, overarching theory that truly speaks to the question of how the brain encodes mental activity is remote. For these reasons, both levels of analysis are not promising foundations on which to build a theory of mind-brain relations. However, unlike the microneural level, macroneural experiments can not only be conceived but actually implemented. The question arises, however, of the validity and reliability of this approach to determining how the brain makes the mind. This is the topic of most of the rest of this article.

Complexity

It is a truism that the brain is an extraordinarily complex organ. Indeed, given the idiosyncratic nature of its interconnections, an argument can be made that it is the most complex entity in the universe. Regardless of which level of analysis one is considering, the combinatorial complexity of the brain stresses any possible explanation of its function to its limits for a relatively simple practical reason—there are too many neurons involved in cognitive processes to handle the flood of information they represent. It takes very few interacting neurons to pose an intractable combinatoric problem for any currently conceivable method of analysis. The complexity issue has been known for years. As long ago as the 1960's, Leon Harmon was demonstrating a simple three neuron simulator that produced unpredictable results.

Shortly thereafter, Stockmeyer and Chandra (1979) showed how very simple problems (e.g., minimizing a travelling salesman's itinerary) could require hyper-astronomical processing times. They concluded that the task of minimizing the travelling salesman's route would "require a computer as large as the universe running for at least as long as the age of the universe." (p. 140) Other estimates suggest that if the salesman had as few of 21 stops, an exhaustive solution to the problem would require 77,000 years of even the most powerful computer calculation time.

The work of Karp (1986) on the intractability of many other superficially simple combinatorial problems is especially relevant in this context. Specifically, many, if not most, real world problems are NP-Complete, that is computationally intractable. Although the work of Karp and others is well known in mathematics and computer theory, cognitive neuroscientists seem to be assuming that some of the greatest challenges are in principle solvable when it has already been proven that they are not.

Koch (2012) recently put this problem into perspective when he calculated that the time required to “exhaustively describe a [neural] system” (and this is what we assume we would have to do to “explain how the brain generates the mind) would grow faster than exponentially. As an example, Koch calculated that the 2000 neurons of the visual cortex of a mouse would require a computer system running for 10 million years even if the computer’s powers were simultaneously growing at the modern rate specified by Moore’s law (“computer power doubles roughly every 2 years.”) Koch did go on to note that there were some simplifying conditions comparable to those used to speed up a standard Fourier transform to a “fast” (FFT) version, that could speed up the process by a few orders of magnitude. One way to do this is by grouping neurons into clusters. However, whenever one regularizes a procedure in this manner, there is the ever present possibility that the cognitively relevant organizational properties one is seeking may be tossed out.

Complexity is a result of a number of factors. Obviously sheer numerousness of the components involved in any realistic neuronal network that might be capable of instantiating a cognitive process is one factor. But, as we have already seen, even a modest number of interconnected components can produce combinatorial intractability. Furthermore, nonlinear neural mechanisms pose computational challenges that have not yet been solved by modern mathematics and remain among the most severe challenges confronted by cognitive neuroscience.

Beyond sheer numerousness and nonlinearity, many early theories of neural nets did not scale up well. A simple neural network exhibiting, for example, learning with less than a hundred neurons may suddenly saturate with all simulated neurons responding simultaneously when even a few more neurons are added or a slightly different network configuration introduced. The eventual collapse of the small neuronal network program in the 1980’s (as discussed by Fodor & Pylyshyn, 1988) when attempts were made to scale up the number of synthetic neurons was probably an unavoidable outcome of this intrinsic instability. Although far more simulated neurons are involved in today’s computational neuroscience models, these prototheories are also subject to the curse of numerousness and complexity.

Sheer numerousness, however, pales into significance when one also considers the idiosyncratic nature of the anatomy of the neurons and their interconnections that make up the brain. There are many different types of brain neurons and they are interconnected in a plethora of irregular ways by an abundance of synaptic variations. There are no shortcut methods to overcome this impediment to explanation that do not result in a massive loss of the very information that is presumably the psychoneural equivalent of cognitive processes. It has been argued that any proposed attempt to study the brain at this minute level is beyond both our technology and our mathematics (Lichtman & Denk, 2011).

Loss of Information with Response Pooling

In such a situation, in which an exhaustive microneural analysis is beyond our technical abilities, science tends to turn to available macroneural cumulative measures (e.g., the fMRI or the EEG) and effectively let nature do the accumulative analysis for approximate solutions to complex neuronal network problems. The expectation is that the critical microneural information will be at least partially preserved at the macroneural level. How the cumulative process works neurophysiologically is not completely understood but it may involve summations of either local (e.g., Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001) or spike action potentials (e.g. Mukamel et al., 2005).

Currently, it seems likely that the salient microneural information is not, in fact, preserved in macroneural level measurements. The voltages and ionic currents that are the responses of individual neurons must follow the laws of physics as they are added and subtracted from each other. Once added or accumulated into pooled responses, basic thermodynamic principles argue that they cannot be retrieved any more than an egg can be unscrambled. This constraint applies directly to the pooling of neuronal responses directly or indirectly into signals like the fMRI or the EEG. Thus, there is an innate ambiguity about the origins of any macroneural signal based on 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 macroneural responses—for, example, the subtractive method—produces data that cannot discriminate between microneural brain states. This issue is such an important matter that it has been continuously discussed for almost two decades (Van Orden & Papp, 1997) and is still in need of defense (Roskies, 2010).

The implication of the fact that the same macroneural responses can be produced by a variety of different microneural responses is that macroneural fMRI images are neutral with regard to their microneural origins. For example, a total lack of a differential fMRI response may be due to no change in the microneural responses occurring between the experimental and control conditions. However, the accumulated responses of two completely different microneural patterns can also summate to the same cumulative null value. This is the fundamental weakness of the macroneural approach; the critical microneural information essential to understanding the transition from brain state to cognitive activity has been lost due to the pooling process.

The one possible exception to this neutrality of fMRI images is that information may be preserved concerning the respective locations of function-specific nodes, regions, activation sites, or locales on and in the brain. Regardless of the details, finding the brain loci associated with cognitive processes has been the core of the larger portion of the research carried out with fMRI equipment. Should it turn out that the neural representations of specific cognitive activities actually are localized to particular nodes or activation sites on the brain, the location of the respective cognitive responses might be considered to be preserved information. As we see later, however, even this most basic idea of brain organization—localization—may not be empirically supportable.

Underdetermination

Another related 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. Some of this information was lost due to pooling as just discussed—microneural information is not persevered in macroneural signals. However, it is more likely that most of it is due to the fundamental conceptual difficulty that the information necessary to answer a cognitive neuroscience question was never present in available macroneural measurements. Both behavior and formal mathematical models, for example, are underdetermined in that they do not contain enough information to uniquely determine the specific underlying mechanisms that produce the behavior (Moore, 1956). Indeed, behavior and mathematics cannot distinguish among what maybe an innumerable number of alternative possible and plausible underlying mechanisms.

Furthermore, additional experiments cannot always be carried out that will resolve this deficiency. Hilgetag, O’Neil, and Young (1996), for example, 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 hierarchy of activation order because the outcomes are fundamentally underdetermined. Carrying out additional experiments leads, paradoxically, to an increase in the number of plausible mechanisms at a rate that exceeds the number of new experiments that can be designed to determine the response hierarchy. In short, simply collecting more data may not help to resolve issues of underlying mechanism if the data are underdetermined; indeed, it may exacerbate the problem.

Another way to describe underdetermination is more closely linked to potential neural mechanisms. I refer here to the “many to one” constraint. If, for arguments sake, we accept the possibility of explaining behavioral observations in terms of macroneural mechanisms, then there is no reason to assume that a particular mechanism is solely capable of encoding any particular kind of behavior. It is far more likely that there are many different mechanisms that can equally well provide a satisfactory representation of the neural mechanisms associated with a particular behavior, i.e., many neural mechanisms can produce the same behavior. Thus, a behavior is underdetermined with regard to the identification of a particular explanatory neural mechanism. It is also true that the inverse problem posed by the activation of a particular neural mechanism cannot be used as an indicator of a particular cognitive process. This is the problem of “reverse inference” highlighted by Poldrack (2006).

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.

A novel way of emphasizing how this lack of isomorphism between cognitive and neural vocabularies affects our thinking has been provided by Oosterwijk et al. (2012). Emphasizing that the quest to associate specific cognitive constructs with particular brain regions has generally been unfulfilled, they offer an 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 (a tradition that goes back to Aristotle’s faculty psychology) 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)

There are two ways to interpret this statement that are relevant to this discussion. First, psychological processes are not sufficiently well defined to correspond with the detailed responses of the underlying neural machinery. The second, as just noted, is that the brain is not necessarily mapped in the same manner as are cognitive processes. Either or both of these implications suggest that there is a fundamental mismatch between our psychological and neurophysiological languages and concepts. To search for localized psychological constructs in the brain, therefore, may be a search for a chimera.

Inaccessibility

Why should this imprecision of definition and measurement of cognitive processes exist? A major conceptual reason is that cognitive hypothetical constructs are not only inherently difficult to define but also impossible to directly measure. From any point of view, mental experiences are private, intrapersonal, and inaccessible psychological states that can only be experienced by the individual and are not directly sharable with others. Neither introspection nor experimental assays can provide us with direct evidence of the values, properties, and dimensions of mental phenomena. Our reasoning concerning “other minds” is, therefore, limited to analogies based on the Cartesian assumption that “Cogito ergo sum” and the analogous conviction that our fellows must be just like us. This means that we must use behavioral surrogates such as “percent correct” and “reaction time” as indirect measures of cognitive responses, thus further isolating us from the phenomena of primary interest—our cognitive responses.

Another reason that cognitive variables may be inaccessible is that they probably reflect the outcome of extremely complex multi-variate neural responses that cannot be characterized by single (or a few) measures typical of modern cognitive neuroscientific research. The net effect of inadequate definitions as well as inaccessible, underdetermined responses is to place conceptual barriers between the various levels of representation of cognitive activity and, thus, limit the possibilities of developing a theory of mind-brain relationships.

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 great contention. Namely, can any measurement that is made of brain processes be theoretically transparent to cognitive activity? That is—are objective brain measurements such as the fMRI or the EEG capable of correlating sufficiently well with cognitive activity to serve as an “explanation” of that activity?

The main weakness of the a priori expectation that any objective brain response should be able to inform psychology is that the multiple levels of brain coding mean that many objective measurements of brain activity are simply going to be irrelevant to the mind-brain problem. Studies of the chemistry of neurons and synapses, microscopic investigations of structure, and single cell recordings tell us little about the cognitively meaningful activity of what is increasing likely to be widely distributed portions of the whole brain. It is not the molecular chemistry of neurons nor of synapses that opens the door to understanding how the brain works as a cognitive engine; it is more likely to be the aggregate information processing by a host of neurons whose collective (but not cumulative) states that account for mind. This is a formidable barrier to building neuroreductionist theories of cognition.

Nevertheless, the search for objective measures of ill-defined psychological states continues. In doing so it perpetuates the myth that observable, objective, neural brain responses representing cognitive processes are decodable. Even more seriously, the search for objectivity can bias our experiments such that the probability of false positive results is inordinately enhanced.

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 theoretical successes in explaining sensory and motor processes, on the one hand, and the much more complicated matter of understanding cognitive representation. Peripheral sensory and motor signal transmission is characterized by a mainly unidirectional flow of information, relatively simple dimensional correlations between stimuli and neural responses, as well as between stimuli and perceived experiences. 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 have developed powerful techniques for quantifying responses that require but the simplest of “Class A”

discriminative responses (Brindley, 1960). These simplifying conditions do not hold for cognitive process such as decision making or emotion.

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. Indeed, sensory mechanisms often have a repetitive anatomy that is almost crystalline in nature, especially in invertebrates. As a result of the relatively simple anatomy and task of the sensory and motor pathways, it has been much easier to develop theories, indeed, nearly complete explanations, of transmission codes. The heyday of this kind of work was epitomized by the work of Hartline and Ratliff (1957) and Hubel and Wiesel (1965).

The result was general disappointment as the false analogies were drawn between sensory and motor information transmission processes were not easily transferrable to cognitive activities. These false and optimistic analogies persist to this day. In short, determining the neurophysiological basis of cognition is a much more complex process than what apparently seems to be analogous task for sensory and motor processes.

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 functional 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 (e.g., Anderson, 2010; Johnson, 2011) are doing is to develop plausible, but highly speculative, “hypothetical constructs” that could, in principle, describe the behavior but which do not enjoy the support of empirical linkages between the psychological and possible neural mechanisms. The hypothetical mechanisms are then defined in neurophysiological terminology. At best, theories of this type are heuristics that are not likely to rise from prototheoretical hypotheses to robust neural theories. Perhaps we should be reevaluating the important insight about the Conceptual Nervous System offered by Skinner and distinguish between robust neurophysiological entities and “nick names” for hypothetical constructs.

Statistical Barriers to Theory Building

In addition to these conceptual issues, there are a number of statistical difficulties that come to attention recently. The complexity of the statistical procedures required to analyze fMRI brain images is such that there is ample opportunity for statistical artifacts to lead to erroneous

conclusions to be inferred from responses that may be more noise than signal. 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. These include, not controlling for intertrial (Goldfine et al., 2013) or intervoxel (Loring et al., 2002) dependencies, and some are subtle flaws in statistical logic such as nonlinear interactions between spontaneous and evoked responses (He, 2013) buried deep in the analysis procedures. Even others involve the injudicious use of packaged analysis programs (Joppa et al., 2013).

This section reviews some of the more prominent statistical sources of misinformation recently identified by cognitive neuroscientists that may not only lead to false positives but also to the invalidation of a number of current ideas about the representation of cognitive processes by macroneural brain measures.

General Problems

Any discussion of the problems encountered in authenticating research findings in general must include the sharply critical and contentious thoughts of Ioannidis (2005) who, among others, has challenged the entire hypothesis testing approach currently dominating many kinds of science. Pulling no punches, he made the provocative assertion in his title “Why most published research findings are false.” Although Ioannidis does not specifically deal with the cognitive neuroscience, his ideas are clearly relevant to this field as well.

Cognitive neuroscience, if we may extrapolate from Ioannidis’ point of view, is based on marginally sound empirical data. Whether his view is appropriate or not, the current corpus of cognitive neuroscience research must be closely and critically examined for reliability and validity before it can be used as a foundation for theory development. One practical suggestion is that no cognitive neuroscience research should be published without a mandatory replication.

How cognitive neuroscience data are reported is another general problem area that has recently come to the notice of the scientific community. We have recently been reminded of two serious problems concerning reporting of brain image-based experiments by Carp (2012a; 2012b). In the first of these two papers, he pointed out that the replicability of research in this field depends on adequate information being available concerning the design of an experiment so that the experiments can be reproduced. Yet, despite this universal necessity for replicability, when Carp analyzed 241 recent articles he found that “many did not report critical methodological details” (p. 289) in sufficient detail to permit replication. Not only was there a wide variety of procedures used for similar experiments, but even the same procedure could be reported in quite different ways. Problems were reported concerning the design, the data acquisition, and preprocessing methods, as well as in modeling and display of the final results.

The problem of inadequate reporting is serious; however, another problem of perhaps even greater import to fMRI-based cognitive neuroscience research was also examined by Carp (2012b) in a follow up article aimed specifically in terms of the variability of the analytical methods that were used. He concluded that variations in method could lead to a cryptic increase in false positive results. This hazard is exacerbated by the multiple modes in which data may be analyzed in an fMRI-based cognitive neuroscience experiment.

According to Carp, there are a very large number of alternative parameters of available experimentation and analysis procedures with which to evaluate brain imaging data. Depending on which subset of the parameters were chosen for a particular experiment, alternative conclusions could be drawn from a brain imaging experiment. Carp suggested that ten is a plausible number of the key analytic parameters (including, for example, different kinds of normalization, filtering, and autocorrelation) as well as two or four alternative means of dealing with each of these analytic parameters. On this basis, Carp computed that there were 6,912 unique analytic “pipelines” that could be followed to a conclusion about the presence or absence of a significant effect. As Carp pointed out, however, not all of these pipelines lead to the same conclusion; indeed, there is a possibility of many competing Type I and Type II errors emerging from what were essentially the same measurements depending on the investigators choice of the particular pipeline that was used. In Carp’s words:

... results also varied considerably from one pipeline to another. Estimates of activation strength were highly variable across analytic pipelines; in regions of peak overall activation, significance estimates varied by 8 Z units. (Carp, 2012b, p. 10.)

Although Carp (2012b) does not assert that all experimental results vary across all pipelines—“others varied widely from one pipeline to another” (p. 12). The fact that some pipelines, even the preponderance of them, were to some degree in agreement (a point made most strongly in Carp, 2012a) does not mitigate the fact that there was an increased tendency for false positives to occur as a function of variability in analytical methodology. The conclusion to which we are drawn by this cogent critique and the concept of alternative pipelines is that variation in method might be driving research conclusions to an unknown degree. Should our scientific results vary with method and not converge for different methods, they should not inspire much confidence in their validity as measures of the underlying neural reality.

Similar points have been made by Simmons, Nelson, and Simonsohn (2011) and Bakker and Wicherts (2011) for purely psychological studies. Both groups also noted that there is such enormous flexibility in the way we collect and analyze data that spurious false positives are encouraged. They emphasized several additional factors as contributing to poor decisions that lead to error—(1) ambiguity about the basis for the $p < .05$ criterion, (2) the arbitrariness of when to terminate data collection, as well as (3) simple calculation errors. Bakker and Wicherts, for example, stated that as many as 33% of the 281 articles they reviewed either misreported their findings or made an error in calculation.

The Misleading Null Hypothesis Significance Test and $p < .05$

Lurking in the background of many of the uncertainties regarding the validity of the data obtained in a brain imaging experiment is the role played by the relatively loose criteria for rejection of the null hypothesis. Cognitive neuroscientists have followed their predecessors in psychology in generally accepting relatively low values of p (e.g., 0.05). (In physics, it should be noted that the criterion is often much higher, corresponding to as many as 5 standard deviations for what is arguably a simpler situation.)

There are a number of problems with the relatively low level criterion typically used by cognitive neuroscientists for testing the significance of the null hypothesis (NHST). Since slight changes in the criterion p-value used in an experiment can produce vastly different estimates of the extent of activation regions observed in a brain image, fusing some and separating others, this is an especially serious problem for current approaches to cognitive neuroscience.

Many investigators (for example, Greenwald, Gonzalez, Harris, & Guthrie, 1996; Lieberman & Cunningham, 2009), have raised questions about the (NHST). Greenwald, et al were particularly insightful in summarizing the problems prevalent in the application of this ubiquitous method. Specifically, they suggested that there were three “severe flaws” (p. 176) that reduced the value of NHST methodology in science. The first was that the null hypothesis was almost always false since slight differences in subjects, methods, and experimental conditions were ubiquitous and true control rare; an idea originally proposed by Meehl (1967).

The second reason, according to Greenwald et al. (1996) was that the NHST does not provide the needed information about an experiment. For example, “the degree of truth or belief” (p. 176) was not quantified. This is an idea presented as early as the work of Rozeboom (1960).

Finally, the third flaw was that the NHST is biased against the null hypothesis, an idea originally proposed by Fisher (1935). That is, given any experimental design, there is a stronger tendency to reject the null hypothesis and report a positive result than there is to accept it.

A blistering critique of the role of the ubiquitous p value-based NHST was published by Lambdin (2012). In brief, Lambdin’s argument is that the entire NHST movement in psychology and other sciences is misleading, if not fallacious. It not only badly misrepresents empirical findings but leads to factually incorrect conclusions. Lambdin pointed out that there has been an almost eighty year old movement criticizing the use of this statistical approach dating back to the now classic work of Fisher in the 1930's. The list of investigators who have rejected the use of NHST is long and impressive; unfortunately none of these critics seems to have had much influence on the use of NHST in psychology or neuroscience. (However, some journals, e.g. *The Psychonomic Bulletin and Review*, now refer explicitly to the “vulnerabilities” of NHST in their instructions to authors.)

The relevant point here is that NHST is considered by a persistent group of statistical critics to be a profoundly flawed method of authenticating data. Yet, virtually the entire brain imaging approach to cognition is currently based on this kind of statistical testing. This is fragile ground on which to build a neuroreductionist theory of cognition.

Ignoring the Number of Tests

Bennett, Wolford, and Miller (2009) and Bennett, Baird, Miller, and Wolford (2011) in dramatic demonstrations of a potential flaw in statistical analysis of fMRI images, showed that positive localizations of fMRI brain responses could be recorded from even as unlikely a specimen as a dead salmon. They pointed out that the statistical error that led to this outrageous result was a lack of correction for the number of tests carried out during a typical fMRI brain image experiment—an observation previously made by Genovese, Lazar, and Nichols (2002).

The point is that the more tests that one carries out the more likely that even a low probability event is going to eventually rise to significance. The situation is particularly misleading because many independent tests of the distributed voxels in a brain image are carried out cryptically in analysis procedures. Thus, in order to avoid a substantial number of false positive responses, some kind of a correction for the number of tests has to be made in the criteria for accepting a result as being significant. The most obvious place that this correction can be made is in the p criterion. For example, Loring et al. (2002) pointed out that if one were examining the results for 100,000 voxels, the appropriate p value should be 0.0000005 rather than 0.05 as corrected by the Holm-Bonferroni (Holm, 1979) correction. Other simple corrections are possible. Simply dividing the traditional criterion ($p < 0.5$) by the number of tests produces an acceptable, if approximate modification of p. However, there are negative effects of this correction—making this simple correction reduces the number of true positives as well as the false positives it was designed to control (Genovese et al., 2002).

The error of not accounting for the number of tests is pervasive throughout the cognitive neuroscience literature. Bennett et al. (2011), for example, reviewed published fMRI based literature and found that only 60-75% of the articles made the needed correction. This raises serious questions about not only about the inferences that may be drawn from brain imaging studies but also about the very empirical validity of such hypotheses as the number, localization, or the size and shape of brain activation regions.

Sample Size

A related and perennial problem engendered by the economics of data collection in experiments using fMRI systems is the unfortunately small number of subjects used in a typical experiment. Because of the high costs of running this kind of experiment (about \$1,000 per subject according to some estimates) there has been a tendency to use only a few subjects. This means that the statistical power of experiments is usually compromised. Inadequate sample size and, thus, low power experimental designs have plagued fMRI studies in cognitive neuroscience since their beginning. Wager, Lindquist, Nichols, Kober, and Van Snellenberg (2009), for example, pointed out that in 415 early studies an average of only 11 or 12 subjects was used and in some cases as few as 4 or 5. Given the relatively poor signal-to-noise ratios of fMRI responses, such small numbers must be considered to be inadequate.

Another example of how sample size can influence interpretations is found in the work of Ihnen, Church, Petersen, and Schlager (2009)—a study that was highly critical of any putative sex difference it was designed to detect. In a study of sex difference using two groups of 23 male and 23 female subjects, a significant difference was reported between the sex divided groups. However, when the subjects were randomly assorted into two groups of 23 subjects, each with the same number of male and female subjects, other equally significant differences between the two groups were observed. They concluded, therefore:

Perhaps most significantly, our discovery of comparable regions of “group difference” in both task sets undermines the overall reliability of the finding of sex differences, because it demonstrates clearly that the analysis of small samples for group differences

may be confounded by spurious positive findings unrelated to the comparison of interest. (p. 1031)

Just how serious the problem of inadequate sample size may be for cognitive neuroscience is dramatically evidenced by a recent pair of studies (Thyreau et al., 2012; Gonzales-Castillo et al., 2012) in which unusually large sample sizes were analyzed. I discuss these important findings in more detail shortly.

Double Dipping

Statistical difficulties such as those produced by inadequate sample size are fairly obvious and well known. However, there are some comparable errors that are more subtle and prone to be overlooked. Recent attention, for example, has been directed at problem of “double dipping” by Vul, Harris, Winkelman, and Pashler (2009) and Vul and Kanwisher (2010). Vul and his colleagues noticed that tests of significance of a number of fMRI-based experiments were “implausibly high” given the measured reliability of both the fMRI data and behavioral measures of personality and emotion.

On the basis of this observation, Vul et al. (2009) evaluated 54 articles and showed that they were cryptically using the same data to both select and then to measure the responses. In doing so they were correlating two measures that were not independent of each other, thus “double dipping” from their data base. This lack of independence was tracked down to the criteria that were used to threshold the data.

Kriegeskorte, Simmons, Belgowan, and Baker (2009) reported the same kind of spuriously high significance due to a kind of double dipping when correlating brain images and single neuron results. They attributed this artifact to the premature definition of fMRI regions of interest and the inappropriate restriction of the sample of neurons to those with “certain response properties” (p. 535). Fiedler (2011) joined the fray later by arguing that the lack of independence that led to the errors highlighted by Vul, Kriegeskorte, and the others is much more widely spread than just in the neurosciences.

Double dipping comes in many guises. Another manifestation of it is to invoke what are called functional localizers (For a full discussion of the pro and cons of using functional analyzers, see Friston, Rotshtein, Geng, Sterzer, & Henson, 2010 and Saxe, Bret, and Kanwisher, 2006). A functional localizer is the preliminary determination of the results expected in a more detailed experiment so that the main experiment to follow can deal with a smaller number of factors, regions, or dimensions, hopefully resulting in a more precise estimate of activation areas.

The advantage of establishing a preliminary functional localizer is that a much simpler experimental protocol than a full factorial design may be used if the salient data has been restricted to a particular region of the brain. The main disadvantage of premature identifications of a functional localizer is that some important and highly correlated responses or interactions may be missed due to the inappropriate narrowing of the field of view. Therefore, although the process of determining the extent of a functional localizer has major advantages in reducing the amount of data to be processed, it is also susceptible to obscuring phenomena outside of the range of the predetermined responsiveness. In fMRI experiments, where there is a great deal of

uncertainty concerning the extent of an activation region, this approach may seriously misestimate the actual range of a responsive area.

Erroneously Comparing Significances

A number of statisticians have pointed out a widespread error in interpreting different significance values. The problem lies in the subtle fact that a difference between two distributions is not necessarily significant if one is significant and the other is not. Although well known to statisticians (e.g., Gelman & Stern, 2006), this caveat is apparently not well known to cognitive neuroscientists.

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)

Other Statistical Issues

The statistical issues discussed so far are among the most notable of the many potential sources of artifacts. However, there are many other subtle statistical errors that have emerged in recent years that have not yet been removed by the armamentarium of cognitive neuroscience. For example, further corroboration for the production of spurious positive responses under some conditions and with certain types of analytic methods comes from the work of Eklund, Andersson, Josephson, Johansson, and Knutsson (2012) using statistical parameter matching (SPM) methodologies. These investigators reevaluated 1484 data sets that had previously been generated to examine the rest activity of the brain. All of these data sets were assumed to have little significant cognitively related activity since cognitive activity was presumably at rest. For, $p < 0.05$, we, therefore, would have expected 5% positive responses in the large number of data sets. Eklund and his colleagues, however, found that significance rates actually ranged from 1% to 70% when one simply counted the number of significant data sets. Although, they were careful not to generalize this result to other methods, they did show how this particular SPM method was prone to producing surprising high numbers of false positives.

Additionally, Plaut (1995) was among the first to discuss problems with the dissociation technique. Van Orden and Papp (1997) pioneered the rise in our awareness of the problems with the subtraction method. Bluhm (2013) has recently discussed how regions of interest and within group analyses can distort both data and conclusions. Lurking in the background, of course, is the general problem of mistaking “correlation for causation” (Yule, 1899)

The problem is that these biased statistical and methodological errors, despite extensive discussion in the literature, are still prevalent in the cognitive neuroscience literature. Some of them are relative new developments that have come to our attention because of the rise of brain imaging techniques; however, many of them represent persistent problems in dealing with systems in which a small signal is buried in substantial amounts of noise.

Empirical Barriers to Theory Building

The Case against 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 that selectively respond during particular cognitive activities. The currently prevailing metaphor is of the brain as a system of localized, function-specific nodes or regions. In some of the early forms of this idea, the nodes were isolated functionally 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 sparsely distributed, multiple, nonfunction-specific regions of the brain interacting with each other. Theory development has recently become increasingly aimed at studying not only individual regional localization but also how these regions interacted.

The theory loaded question now arises—do localized focal regions or nodes actually exist? The idea of function-specialized local regions, either singly or as an interacting network, is so ingrained in modern cognitive neuroscience that a negative answer to these rhetorical questions would probably require major changes in our ideas concerning the neural bases of cognition.

Major changes or not, 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, cogently expressed by Lindquist and Barrett (2012), among others, is that the brain operates in a much more distributed and holistic fashion, as well as less specialized, than previously suspected. Their constructionist prototheory suggests a promising future change in the direction theory might take; it is conceptually different than the dominant current approach.

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 seem to hold!

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 functional 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)

They also reported the results of an experiment relating the amount of their data being analyzed to the extent of these distributed brain activations. Their results indicated that the more data that were included in the analysis, the greater was the extent of the distributed activity across the brain and less evidence existed for localized function.

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—wide spread distribution of responses across the entire brain for even the most elementary cognitive process. They pointed out that when very large samples were used, even small responses began to achieve significance; this expansion increased until virtually the whole brain was 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.

Oosterwijk et al. (2012) also supported the idea that the brain mechanisms of cognition must be more widely distributed in the brain than is currently held by localization theorists. From their point of view, the assertion that the neural mechanisms of cognitive processes are localized in the form of macroneural function-specific regions scattered across the brain was fundamentally incorrect. Specifically, they noted:

Our results also emphasize the importance of examining distributed patterns of brain activation to understand mental states with different content rather than focusing on single regions ... isolated regions may serve different psychological functions depending on what they are connected to during a given instance. (p. 2125)

The point of these studies is that the traditional search for localized macroneural regions mapping onto behaviorally defined psychological constructs is likely to be fruitless. The brain is not a system of isolated function-specific nodes; it is more likely to be a broadly, even universally, distributed system of neural mechanisms that are neither function-specific nor localizable to any particular region. Such a metaphor is vastly different from the current one guiding much of macroneural thinking about how the brain is organized. Indeed, the general import of this work is that the search for narrowly localized brain regions associable with cognitive processes may be sterile. The conflict between these emerging data and the search for localizable neural mechanisms demands a major change in whatever kind of theoretical model is being pursued.

On the reliability of fMRI Data

This brings us to our next empirical issue—how reliable are fMRI data? In the past two decades, the body of the cognitive neuroscience literature based 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 rare. Indeed, only a few investigators have made a determined effort to study the reliability of cognitively related fMRI findings. Those few who have ventured into this field have typically found that reliability in cognitive neuroscience is elusive and inconsistent and that experiments that seem to be very close in intent and design may produce different estimates of which brain regions are involved in any particular cognitive process.

One of the few reports that have specifically studied the problem of reliability in comprehensive detail was published by Bennett and Miller (2010). In seeking the answer to the question—how reliable are the results from functional magnetic resonance imaging?—they came to the conclusion that:

Our findings speak strongly to the question of why there is no agreed-upon average value for fMRI reliability. There are so many factors spread out across so many levels of influence that it is almost impossible to summarize the reliability of fMRI with a single value. (p. 150)

Although Bennett and Miller remained generally supportive of the fMRI technique for future research, they acknowledged that the problem of reliability is critical and largely overlooked. Yet, reliable empirical data is the foundation of any scientific enterprise. The problem could not be more important to cognitive neuroscience since unacceptable levels of variability are found at all levels of investigation. Observations of inconsistency appear in repeated trials on a single subject, through pools of subjects and interexperiment comparisons, to the meta-analytic studies in which data from many experiments are pooled.

The most disconcerting of these reliability comparisons are obtained from single subjects repeating exactly the same protocol at later times. As a measure of reliability this single-subject, repeated-measures design is the most stringent test of reliability possible since a maximum number of factors are controlled. Yet, only a few investigators have studied the variability of data obtained from single subjects. Among the earliest were Aquirre, Zarahn, and D'Eposito (1998), McGonigle et al. (2000), and Miller et al. (2002). Examination of the figures from other reports (e.g., Lund, Norgaard, Rosrup, Rowe, & Paulson, 2005) which were not specifically studies of single subject reliability also shows substantial intrasubject variation for explicitly cognitive tasks.

The problem remains contentious, however. Although these single subject experiments generally reported that there was substantial variation in the responses from session to session, other reports from about this same time period (e.g., Maitra, Roys, & Gullapalli, 2002) argued that this observed variability was modest.

Other investigators have also reported relatively low levels of reliability in brain imaging experiments. Rau et al. (2007), for example, reported that a naming task produced highly variable data concerning the location of well known Broca's area. This inconsistency was partially ameliorated by pairing it with a noun generating task (perhaps due to the addition of a task with a larger motor component). Duncan, Pattamadilok, Knierim, and Devlin (2009) also reported low levels of intrasubject reliability when brain scans defining functional regions of interest were compared with results from a more extensive main experiment.

What, therefore, appeared to be a conflict in the data seems to be resolvable by noting that the least variable brain responses were produced by motor tasks such as those used by Maitra et al. (2002) whereas the more variable were associated with higher level cognitive tasks (e.g., McGonigle et al., 2000). Despite this potential accommodation of what at first may seem to be conflicting data, the problem of single subject reliability remains quite controversial as evidenced in a recent updated discussion of his original work by McGonigle (2012).

The intrasubject variability of the fMRI responses reported by Aquirre et al. (1998), McGonigle et al. (2000), and Miller and their groups (2002), respectively, highlight the inconsistency of results, even those from what should be assumed to be under the best possible control—repeated measures on a single subject. For the last decade, there have been a number of other attempts to mitigate the low reliability by post data collection analysis procedures. Although 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 location oriented brain imaging in the search for the neural bases of cognition must be critically reevaluated.

Whatever the resolution of this debate, in fMRI-based experiments, there is no question that intrasubject variability is less for intrasubject than for intersubject comparisons as shown in Table 1 from Miller et al. (2002).

Table 1.
Intra- and Intersubject Correlations Showing High Degree of Variability.
(Miller et al., 2002)

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

In this table, correlations are shown between and within six subjects in two sessions for an episodic retrieval experiment in which individual subjects were tested and then retested under identical conditions. The cross subject comparisons are not surprising; different subjects would be expected to have different fMRI signatures. However, the relatively modest level of correlation for individual subjects suggests a lack of consistency that challenges the macroneural fMRI approach to understanding the neural basis of cognition.

The data concerned with reliability are not, however, without controversy; Aron, Gluck, and Poldrack (2006) and others have found a higher degree of reliability in a slightly different experimental protocol—a “probabilistic classification learning” task with a one-year interval between the test and retest portions of the experiment. They report substantial similarities between the two tests. However, even their work showed some evidence of changes over the one year period between the original test and the retest. For example, all of the responding areas responded more vigorously in the retest than in the test. As they note, this only establishes reliability for one kind of experiment and for the fronto-parietal system in particular. There is obviously much to be done on the issue of reliability. Furthermore, even some of the most obvious inconsistencies have been challenged. For example, Smith et al., 2005, have disputed McGonigle’s claim that intersession variability was unacceptably high.

Indeed, just what is consistent and what is unreliable will always be something of a judgment call. There are few systems for quantifying the degree of consistency of multifactorial data such as these. The most familiar is the Intra-class Correlation Coefficient (ICC) (Shrout & Fleiss, 1979; Caceres, Hall, Zelaya, Williams, & Mehta, 2009). The ICC is valuable for comparing the relationship of two groups of data as opposed to the usual Pearson r which looks as paired values. Thus, it is useful to scale consistency among variable groups of data such as those found in brain images.

Interexperiment variability is also a major source of uncertainty in drawing associations between cognitive and macroneural processes. The literature is so extensive that it is possible to be overwhelmed with the differences in identified brain activation areas from what are supposed to be comparable studies. A graphic depiction of interexperiment variability within a restricted domain of cognitive neuroscience—emotions—is shown in Figure 1. This figure is a map of the raw distribution of activation areas used in a meta-analysis (Kober et al., 2008).

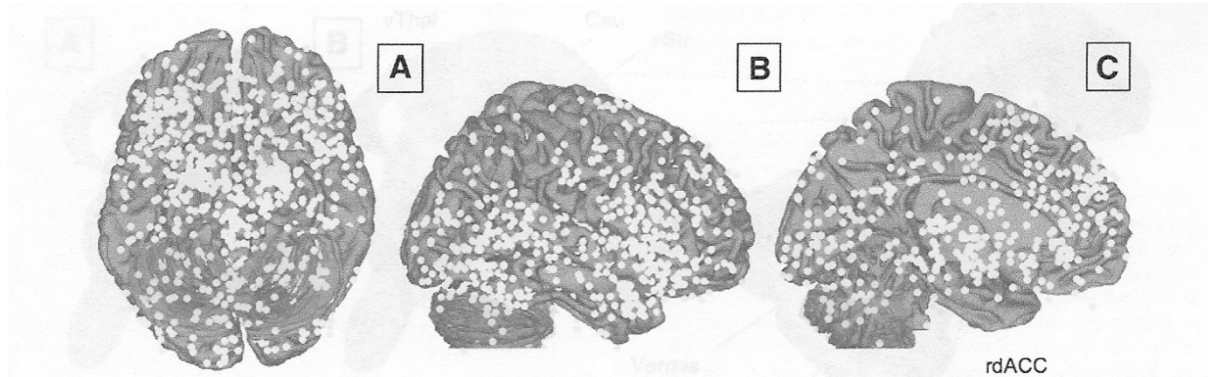


Figure 1.
Raw Activation Peaks from a Meta-analysis of Emotion.
 (Kober, et al., 2008)

This map was generated from 162 experimental reports that resulted in a total of 437 activation peaks. The visual evidence for a low level of reliability in this case is the broad distribution of the activation peaks across virtually the entire brain. To put it most succinctly, multiple experiments purporting to represent the same cognitive process produce widely distributed and variable activations.

The next step in illustrating the variability of fMRI images is to compare the outputs of different meta-analyses for what is purportedly the same cognitive process. As straightforward as the meta-analytic idea (more data produces higher power experiments) is in principle and how important the stability of such data-rich results would be if confirmed, actually carrying out this type of inter-meta-analysis comparison is relatively rare. In one of the few studies of this kind, comparisons of the results from several supposedly comparable meta-analyses were reported Uttal (2013). These results are briefly summarized here.

The results of these meta-analyses were reported in terms of the Brodmann areas (BAs) activated during the respective cognitive processes. A simple ratio measure of the consistency among the identified BAs of the two meta-analyses was used as a measure of reliability. This measure was defined as the number of BAs that were activated by both meta-analyses divided by the number of BAs that were activated by either one alone.

The results of the two-way comparisons were not reassuring; ratio measures for the four pair-wise comparisons were:

- Autobiographical Memory -- 51%
- Emotional Face Processing -- 14%
- Single Word Reading -- 45%
- N-Back Working Memory -- 41%

A graphic comparison of seven meta-analyses carried out for working memory is shown in Table 2. The results of this comparison were also clear—there was substantial variation among the areas reported by the seven meta-analyses.

Table 2.
A Comparison of Seven Meta-analyses dealing with Working Memory.
(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 Mostofsky (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).

The previous cited results were exclusively cortical using the ratio of activated Brodmann areas as the criterion for comparability. Kober et al. (2008) also provided narrative descriptions of both cortical and subcortical brain regions for emotion that were surprisingly inconsistent when compared with the narrative descriptions reported by Murphy, Nimmo-Smith, and Lawrence (2003) for what was purported to be the same cognitive process—emotion. These two meta-analyses were based on moderately large numbers of pooled experiments (162 for the former and 106 for the latter.) With this relatively large amount of data we would expect a high degree of agreement between the two meta-analyses. However, a comparison of the cortical and

subcortical regions identified with emotional processes by these two studies resulted in a ratio measure of only 42%.

In sum, even when sufficiently large amounts of data are pooled to provide a high power statistical analysis, there remains great variability in the findings. Coupled with the magnitude of intrasubject, intersubject, and interexperiment variability, this level of inconsistency raises serious questions about the validity of the data obtained from brain imaging experiments and, thus, about the hope that a neuroreductionist theory of cognition will be forthcoming from this macroneural approach.

Technical Artifacts

Because of the complexity of the brain and cognition as well as the sensitivity of the processing procedures necessary to extract a BOLD signal from background noise, the potential for technical artifacts of one kind or another in this kind of research is substantial. Maitra et al. (2002) reviewed some of the currently known technical artifacts distorting fMRI brain images including, physiological variation, scanner noise, subject motion, and image processing algorithms.

One technical artifact that is currently getting special attention given our interest in brain activity is the effect of head motion on fMRI brain images. Lund et al. (2005), for example, were among the first to attribute a considerable portion of the observed variability in fMRI images to motion artifacts. Recently, Researchers such as Van Dijk, Sabuncu, and Buckner (2012) and Power, Barnes, Snyder, Schlaggar, and Petersen (2012) all reported that head motions (as small as a .1 millimeter) may have substantial effects on measurements of the brain's interconnecting tracts especially when voxel are small. Under ordinary circumstances this would be merely an arcane result of little general interest to other than specialists in this field. However, macroneural connectivity is now being explored as a biomarker for various cognitive dysfunctions. For example, it has been suggested as a neuroanatomical correlate of autism—an increasingly commonly observed form of behavioral dysfunction (Just, Cherkassy, Keller, Kana, and Minshew, 2007). The suggestion that connectivity estimates are especially sensitive to artifacts of head motion would strongly challenge this association. Other possible motion artifacts include eye motion, swallowing, peristaltic motions, and variations in blood pressure.

Some technical artifacts may be correctable. New fMRI systems, body motion controlling methods, and even mathematical procedures have been proposed to overcome some of these problems. Indeed, all of these technical artifacts may pale in light of the much more influential anatomical, physiological, and behavioral sources of variability that are much less tractable to correction.

New Developments; Can They Provide a Pathway to Neuroreductive Theory?

Clearly in a complex situation like this there is a continued effort to overcome the obstacles to understanding how the brain produces the mind. These “repairs” come in a wide variety, some of which are aimed at improving data 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 ingenious steps forward actually solve some of 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 simulations do not always lead us to theoretical ideas. The flow of bad questions, often brilliantly asked and partially answered, has to be tempered by understanding of what they may or may not mean to the mind-brain problem. The availability of increasingly powerful computers has created a new approach to mind-brain theory generically referred to as computational neuroscience. However, this approach has not yet fulfilled its promise at either the microneural or macroneural level.

New Macroneural Developments

Consider for example, the accomplishments of neuroanatomists such as Sporns (2011). His remarkable book contributed enormously to our knowledge of the structure of the brain. However, as Sporns notes:

And yet, the relationship between the brain and cognition is still only poorly understood. Great progress notwithstanding, neuroscience still cannot answer the “big questions” about mind and intelligence. (p. 179)

What Sporns has highlighted can be said of many other exciting neuroscientific developments. Recent efforts at making the brain visually transparent (Chung & Deisseroth, 2013) are magnificent tools for visualizing the anatomy of brain tissue; however, it is not equally clear what they might contribute little to mind-brain relations. After all is said and done, they are post mortem applications.

Many current 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 function of these large scale networks to a particular cognitive state—emotion.

Most research of this kind is 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. They, thus, provide an organizational theme for testing macroneural 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 is becoming increasingly clear that this macroneural network approach is not the 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 were no functionally committed macroneural nodes, only broadly distributed, multifunctional 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; the whole search methodology of using brain imaging equipment to answer the “where” questions would have to be reevaluated since the distributed networks to which Lindquist and Barrett and others alluded would each be everywhere in the brain.

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 (nuisance regression) with known sources of noise are now sequentially being applied to remove noise from fMRI signals (Hallquist, Hwang, & Luna, 2013). However, this dual correction produces results that depend on the order in which the two corrective steps are applied to the data, another example of the warning raised by Carp (2012b). For example, this double barreled noise reduction process is now known to reintroduce noise signals (that had been removed by the filtering) when the regression methods are applied after the filtering.

Considering the limits that are becoming evident in the macroneural approach, Poldrack (2010), a widely respected contributor to the field, asserted that it may be an impossible task in its present form. He noted:

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 brain areas. Poldrack (2010) 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 from this state of affairs, Poldrack (2010) suggested that we have to develop ontologies or taxonomies of cognitive processes, 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, ontology development is an ambitious goal. Whether or not this can be done for a system such as human cognition with its

adaptability, flexibility, and redundancy remain uncertain. Even if we could improve psychology's terminology and definitions, it is unlikely that this achievement would overcome current challenges at either the macroneural or microneural levels.

Other new developments have been proposed as means to understanding neuronal networks. The study of a network in which the function of individual neurons is preserved is the ultimate goal of the microneural approach. Optogenetics (Deisseroth, 2011) is a recently developed technique that can be used to activate classes of similar neurons by dealing with their specific chemistry or genetic properties. However, this powerful technique does not, as sometimes suggested, permit the idiosyncratic activation or inhibition of individual neurons, each with a different activity pattern. This method is limited to activating many neurons in synchrony. Although phrases such as “specific optogenetic inhibition of excitatory neurons” is commonly used (e.g., Goshen et al., 2011) in actuality they are referring to concurrent control of large numbers of neurons. In other words, optogenetics is, like the fMRI, a macroneural approach that does not permit the simultaneous measurement of individual neurons.

Similarly, Mean-field theory (Arbib, 2003) does not preserve the information recorded from the individual responses. Instead, it 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 problem into a one-body that is supposedly capable of solution. Techniques such as this are commonly used in physics to overcome the numerosness problem by pooling individual responses; however, it works there because the interacting forces are uniform. This is a simplifying property that is not available in real neural networks where each neuron and each synapse is responding idiosyncratically.

New Microneural Developments

It is yet to be determined what the macroneural approach will ultimately contribute to the ultimate solution of the mind-brain problem—there is always the future with surprises that cannot be anticipated. There is, however, an alternative context in which the problem of neuroreductionist theories can currently be attacked—concentration on the microneural aspects of the problem. As we have seen, there are insurmountable barriers to carrying out microneural network experiments directly.

The microneural network approach has bifurcated into two strategies—study of single neurons and studies of networks of these neurons. The first method depends on correlations between cognitive processes and the activity of individual neurons. Much of the single neuron work has been rendered irrelevant and will not be considered further here.

Attention at the microneural level has now turned in distinctive ways to neuronal networks. Cognition, according to this approach, arises from the integrated, collective, but individual, action of many neurons. Neurons, the premise asserts, maintain their individual functional 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.

If a direct empirical approach to the study of cognitively plausible neuronal networks is not possible, the main alternative is computer simulation. This approach aims to use what we know of the microanatomy of the brain to construct a plausible, computational model of what necessarily has to be a simplified simulated neuronal network. A recent review of this approach (de Garis, Chen, Goertzel, & Lian, 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 we do not know enough about the actual organization of the brain, with the exception of a few locales, to make these simulations neuroscientifically significant. This has become a highly controversial matter with some investigators challenging the biological relevance of even very large networks of detailed simulations of neurons. If not neuroscientifically significant, it does not mean that these simulations are not valuable. An enormous amount has been learned about the interactions of complex networks. De Garis and his colleagues, therefore, conclude:

So far, large-scale brain simulations [at the cellular level] have proved useful mainly in terms of refining the equations for modeling neurons and synapses, and helping substantial conceptual models of brain structure and function by connecting these models with detailed electrophysiological data in working simulations. (p. 25)

It also must be noted that selective stimulation of and recording of responses from a network of simulated neurons suffers from many of the practical problems precluding direct physiological research on real neuronal networks. These computer simulations have not yet provided any further foundation for mind-brain theory.

An alternative, and essentially macroneural way to approach computational theory building that does depend more on cognitive processes is what Goertzel, Lian, Arel, de Garis, and Chen (2010) refer to as Biologically Inspired Cognitive Architecture (BICA). Rather than operating at the microscopic level, BICA are much more closely linked to psychology by 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 enhanced in this new world of super computers. The advantage that they offer is the ability to more powerfully and definitively evaluate the plausibility of a descriptive theory of cognition.

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 might be connected. As Goertzel and his colleagues (2010) 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)

Thus, their contribution to theory or explanation in cognitive neuroscience is limited to indirect heuristics and plausibility tests and computer capabilities. Given the current state of the respective sciences, neither approach, however exciting in their own field, is yet able to contribute to bridging the huge gap between brain and mind.

Conclusions

This review of the current status of cognitive neuroscience research highlights some of the formidable barriers that obstruct our hopes of building a comprehensive, overarching theory explaining 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 be permanent. However, collectively they raise serious questions about whether or not a theory of mind-brain relations is in sight. At the very least, it appears that the idea of mapping the brain into a pattern of localized, function-specific, macroneural, activation regions is obsolescent. On the other hand practical problems, complexity, and combinatorics probably preclude a microneural approach. In the former case, there is too much information whereas in the latter there is too little. In the words of Loosemore and Harley (2010):

... we are in a bind: On the one hand, the resolution of these brain imaging studies is not enough to tell us useful things about the functional level, and the future improvements in the technology do not appear to offer the granularity that we need. On the other hand, the level of specificity of the cognitive theories is currently not good enough to make coarse-grained localization theories useful. (p. 240)

If a comprehensive explanation of how cognitive process emerge from brain activity is unachievable based on either of these two approaches in the near future, can we at least hope for some kind of predictive and descriptive prototheory through the use of macroneural measures? The answer to this question is complex. However, given the conceptual, statistical and empirical barriers to theory building in this field, any progress in formulating a neuroreductive, explanatory “theory” of mind-brain relations appears to be highly elusive now and in the foreseeable future. There will always be loose metaphors, hypothetical constructs, and simulations that mimic cognitive processes, but a necessary and explanatory neural theory of the mind is not on the horizon. Although this may seem to be unduly pessimistic, it might be preferable to call it realistic in the same sense as our inability to exceed the speed of light or to build a perpetual motion machine is realistic and not pessimistic. The evidence so far and the portents for the future simply do not justify any optimism concerning a forthcoming answer to the most important of human questions—how does the brain produce the mind?

The best we can do, in my personal opinion, is to hope for some descriptions, metaphors, or other kinds of proto-theories. However, a truly neuroreductive theory of how the mind emerges from the brain may not be available without some kind of a major conceptual breakthrough. A solution to the mind-brain conundrum remains as elusive today as it has been for millennia. If this conclusion is correct, it suggests that psychology should give up the hope of neuroreductionist theory of mind and concentrate on that which it does so well—studying the behavioral and experiential aspects of cognitive activity.

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), 245-313.
- Arbib, M. A. (Ed.). (2003). *The Handbook of Brain Theory and Neural Networks Second Edition*. 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.
- Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychological Review*, 114(3), 632-656.
- 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.
- Bakker, M., & Wicherts, J. M. (2011). The (mis)reporting of statistical results in psychology journals. *Behavior Research Methods*, 43(3), 666-678.
- Bandettini, P. (2007). Functional MRI today. *International Journal of Psychophysiology*, 63(2), 138-145.
- 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(3), 1-12.
- 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), 1-5.
- Bennett, C. M., & Miller, M. B. (2010). How reliable are the results from functional magnetic resonance imaging? *Annals of the New York Academy of Sciences*, 1191, 133-155.
- 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.
- Bluhm, R. (2013). New research, old problems: Methodological and ethical issues in fMRI research examining sex/gender differences in emotion processing. *Neuroethics*, 6(2), 319-330.

- Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: Rmerging 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.
- Caceres, A., Hall, D. L., Zelaya, F. O., Williams, S. C. R., & Mehta, M. A. (2009). Measuring fMRI reliability with the intra-class correlation coefficient. *NeuroImage*, 45(3), 758-768.
- Carp, J. (2012a). The secret lives of experiments: Methods reporting in the fMRI literature. *NeuroImage*, 63(1), 289-300.
- Carp, J. (October 2012). 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 and Behavior*, 77(4-5), 635-639.
- Chung, K., & Deisseroth, K. (2013). Clarity for mapping the nervous system. *Nature Methods*, 10(6), 508-513.
- 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.
- Deisseroth, K. (2011). Optogenetics. *Nature Methods*, 8(1), 26-29.
- Dietrich, E., & Hardcastle, V. G. (2005). *Sisyphus's Boulder: Consciousness and the Limits of the Knowable*. Amsterdam: John Benjamins.
- Duncan, K. J., Pattamadilok, C., Knierim, I., & Devlin, J. T. (2009). Consistency and variability in functional localisers. *NeuroImage*, 46(4), 1018-1026.
- Eklund, A., Andersson, M., Josephson, C., Johansson, M., & Knutsson, H. (2012). Does parametric fMRI analysis with SPM yield valid results? An empirical study of 1484 rest datasets. *NeuroImage*, 61(3), 565-578.
- Fiedler, K. (2011). Voodoo correlations are everywhere-not only in neuroscience. *Perspectives on Psychological Science*, 6(2), 163-171.
- Fisher, R. A. (1935). *The design of experiments*. Edinburgh, London,: Oliver and Boyde.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1-2), 3-71.

- 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.
- 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.
- Glahn, D. C., Ragland, J. D., Abramoff, A., Barrett, J., Laird, A. R., Bearden, C. E., & Velligan, D. I. (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, 289-291.
- 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 memories. *Cell*, 147(3), 678-689.
- 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. (November 15, 2013). The nuisance of nuisance regression: Spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *NeuroImage*, 82, 208-225.
- Hartline, H. K., & Ratliff, F. (1957). Inhibitory interaction of receptor units in the eye of limulus. *Journal of General Physiology*, 40(3), 357-376.
- 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: 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., O'Neill, M. A., & Young, M. P. (1996). Indeterminate organization of the visual system. *Science*, 271, 776-777.
- Holm, S. (1979). A Simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65-70.
- Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28(2), 229-289.
- 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.
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *Plos Medicine*, 2(8), 696-701.
- Johnson, M. H. (2011). Interactive specialization: A domain-general framework for human functional brain development. *Developmental Cognitive Neuroscience*, 1(1), 7-21.
- Joppa, L. N., McInerney, G., Harper, R., Salido, L., Takeda, K., O'Hara, K...Emmott, S. (2013). Troubling trends in scientific software use. *Science*, 340, 814-815.
- Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex*, 17(4), 951-961.
- Karp, R. M. (1986). Combinatorics, Complexity, and Randomness. *Communications of the Acm*, 29(2), 98-111.
- 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.

- 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.
- Lambdin, C. (2012). Significance tests as sorcery: Science is empirical-significance tests are not. *Theory & Psychology*, 22(1), 67-90.
- Lichtman, J. W., & Denk, W. (2011). The Big and the Small: Challenges of Imaging the Brain's Circuits. *Science*, 334, 618-623.
- 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.
- 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.
- 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.
- Loosemore, R., & Harley, T. (2010). Brains and minds: On the usefulness of localization data to cognitive psychology. In S. J. Hanson & M. Bunzl (Eds.), *Foundational Issues in Human Brain Mapping* (pp. 217-240). Cambridge MA: MIT Press.
- 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.
- 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.
- 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.
- 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.
- Meehan, T. P., & Bressler, S. L. (2012). Neurocognitive networks: Findings, models, and theory. *Neuroscience and Biobehavioral Reviews*, 36 (10), 2232-2247.
- Meehl, P. E. (1967). Theory-testing in psychology and physics: Methodological paradox. *Philosophy of Science*, 34(2), 103-115.
- 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.
- 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 fMRI in human auditory cortex. *Science*, 309, 951-954.
- Murphy, F. C., Nimmo-Smith, I., & Lawrence, A. D. (2003). Functional neuroanatomy of emotions: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, 3(3), 207-233.
- 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.
- 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.
- Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge MA: MIT Press.
- Pashler, H. & Harris, C. R. (2012). Is the reliability crisis overblown? Three arguments examined. *Perspectives in Psychological Science*, 7(6), 531-536.
- Plaut, D. C. (1995). Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, 17(2), 291-321.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10(2), 59-63.
- Poldrack, R. A. (2010). Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on Psychological Science*, 5(6), 753-761.
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3), 2142-5154.
- Putnam, H. (1975). *Mind, language, and reality*. New York: Cambridge University Press.
- Rakover, S. S. (2011). A Plea for methodological dualism and a multi-explanation framework in psychology. *Behavior and Philosophy*, 39/40, 17-43.
- Rau, S., Fesl, G., Bruhns, P., Havel, P., Braun, B., Tonn, J. C., & Ilmberger, 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.
- Roseboom, W. W. (1960). The fallacy of the null-hypothesis significance test. *Psychological Bulletin*, 57(5), 416-428.
- Roskies, A. L. (2010). Saving Subtraction: A reply to Van Orden and Paap. *British Journal for the Philosophy of Science*, 61(3), 635-665.
- 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. (February 2012). Working memory and response selection: A computational account of interactions among cortico-basalganglio-thalamic loops. *Neural Networks*, 26, 59-74.

- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420-428.
- 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.
- 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.
- Skinner, B. F. (1950). Are theories of learning necessary. *Psychological Review*, 57(4), 193-216.
- 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.
- 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.
- Stockmeyer, L. J., & Chandra, A. K. (1979). Intrinsically difficult problems. *Scientific American*, 240(5), 140-159.
- Thyreau, B., Schwartz, Y., Thiron, B., Frouin, E., Loth, E., Vollstadt-Klein, S. ... Poline, J-B. (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.
- 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 Dijk, K. R. A., Sabuncu, M. R., & Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage*, 59(1), 431-438.
- Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information-processing in the primate visual-system: An integrated systems perspective. *Science*, 255, 419-423.

- Van Orden, G. C., & Paap, K. R. (1997). Functional neuroimages fail to discover pieces of mind in parts of the brain. *Philosophy of Science*, 64 (4), S85-S94.
- Vimal, R. L. P. (2009). Meanings attributed to the term 'consciousness:' An overview. *Journal of Consciousness Studies*, 16(5), 9-27.
- 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.
- 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*. Cambridge, MA: MIT Press.
- 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.
- Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, 3(4), 255-274.
- Yule, G. U. (1899). An investigation in the causes of changes in pauperism in England, chiefly during the last two intercensal decades. *Journal of the Royal Statistical Society*, 62, 249-295.