Book Symposium: William Uttal’s The New Phrenology

Précis of The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*

WILLIAM R. UTTAL
Department of Industrial Engineering, Arizona State University, Tempe, AZ 85287, USA
(Accepted: 26 April 2002)

It has long been my concern that, in our efforts to exorcise speculative and philosophical considerations from what we all want to be an objective empirical science, cognitive neuroscientists have overlooked an important methodological tool – critical analysis of the foundation concepts, axioms, assumptions and premises that underlie our day-to-day research efforts. The present book, in an effort to correct this oversight, takes aim at a foundation issue in modern cognitive neuroscience – Are cognitive processes associated with particular brain regions? This is the basic question of localization.

There is a long history of active research by neuroscientists of all stripes seeking answers to this question. In my earlier (1981) book, I dealt extensively with the problem from the point of view of the techniques that were available at that time – mainly direct electrical recording and stimulation and extirpative surgery – as well as with the empirical findings that had accumulated up to that time concerning the localization of cognitive processes in the brain.

In recent years, an entirely new dimension has been added to the study of where cognitive processes might be so localized – the new noninvasive tomographic techniques using radioactivity, X-rays, magnetic fields, or even more exotic means that promise to allow us to peer into the human brain while it is actively engaged in cognitive processing. With our ability to extract extraordinary three-dimensional images of brain anatomy and function and, as some have argued, even signatures of mental activity, by newly developed computational procedures, more and more of my colleagues have been drawn to this powerful approach to the localization question. The enormous amount of activity in this field in the past decade characterizes a topic of great potential importance and possible contribution to our understanding of the relation between brain structure and cognitive activity.
Among the many cognitive processes that have been localized in a specific region of the brain using imaging techniques are visual imagery and rivalry, semantic information processing, memory, pattern discrimination, expectations, language, decision making, attention, emotions, and even the anticipation of pain.

As promising as it might seem at first to be, it takes only a modest amount of scrutiny of historical precedents and the current logic of the field to appreciate that many cognitive neuroscientists have in many cases gone off without careful consideration of the potential problems that lay just below the surface of the entire localization enterprise. I argue in this book that the excitement of the marvelous three-dimensional imaging machines and the allure of noninvasive studies of the human brain have been so enticing that often important caveats have been ignored and research has been carried out without adequate consideration of the logic of this research field.

Is a careful scrutiny of localization research important at this time? One has only to examine the table of contents of one of the world’s leading scientific journals – *Science* – to appreciate the current impact and popularity of localization research using imaging techniques. Virtually all of the articles from 1998 to the date of completion of the manuscript of this book (mid-2000) dealing with cognitive processing included an MRI or PET image. Conventional psychophysical studies based on purely behavioral observations have been almost completely excluded from this important journal. Furthermore, in some of our most prestigious universities, the traditional cognitive (previously known as experimental) psychology programs have been replaced by cognitive neuroscience activities that depend almost exclusively on the positive and implicit acceptance of assumptions about the cerebral localization of hypothetical cognitive modules and the pursuit of these assumptions with imaging techniques. Researchers originally trained to exercise their classic psychophysical and cognitive experimental skills are now working with devices that were not even imaginable during their graduate training. The speed which this new approach has taken over the activity of many cognitive psychologists and neuroscientists and the number of people (as well as the enormous resources) committed to this approach suggests that any criticism of this style of research is likely to be subject to severe counter-criticism.

On one matter, however, there should be no dispute – the modern version of the localization research program is not without its conceptual as well as technical difficulties, and these difficulties should, at least, be discussed. One part of this discussion must concern the logical structure of the field. I point out in this book that the very act of searching for the brain locales of specific cognitive processes begs two fundamental questions. The first of these inadequately considered questions is – What are the cognitive processes that are being localized? This question is answered by the usually unexpressed, but deeply held, axiomatic presupposition that cognitive processes are, indeed, defined in a sufficiently precise manner for them to be dealt with as separable and distinct modules.
The second of these “begged” questions is answered by an equally prejudicial and unjustified assumption. The presumed answer is that these putative cognitive modules or components are actually localized (i.e., represented or encoded) by neuronal activity in circumscribed regions of the brain as opposed to being broadly distributed throughout the brain. If one accepts these presumptive answers, there is little impediment to looking for and eventually finding data that supports the original assumptions. The facts of the matter, however, are that both of these issues are still controversial and yet to be rigorously resolved. Obviously, they should be subject to extensive and careful discussion before they lead to broad agreement on an erroneous model of the relation between the brain and the mind.

One problem is that these presumptions (which eventuate in the axiom that separable modules of mind are associated with specific modules of brain) are much too easy to comprehend. Their simplicity, both as an organizing concept and as a criterion for experimental design, is seductive. Thus, they become amenable to the facile design of research that is, at once, self-fulfilling and self-supporting. It is, to make this argument crystal clear, far easier to deal conceptually, theoretically, and empirically with a system of quasi-independent modules than with a highly interactive distributed system. Arguably, it is possible to develop an entire scientific research program on the basis of such presumptions that results in seriously misguided conclusions concerning the true nature of brain and mind. It is all-too-easy to create and then reify such cognitive “hypothetical constructs” than to cope with the true complexity of the underlying psychobiology. Of course, first steps must be taken, but we must guard, as much as possible, against stepping off in a completely incorrect direction.

To understand the localization research program – both its strengths and its weaknesses – it is important to look at it through many windows. It is necessary to examine historical precedents to see if there is any pattern of logic or illogic that could have confounded the development of brain-mind science in the past. It is also necessary to critically examine the technology of the new imaging methods themselves. The complexity and the sophistication of the tomographic devices that are currently being used to look at the central nervous system are so great that often what seems obvious may, on closer examination, turn out to be not quite so true. Many users of these devices are not completely familiar with the technical details and, thus, are susceptible to subtle misinterpretations or obscure artifacts. At a deeper level, the conceptual train of thinking from observation to conclusion must also be considered.

Certainly, a review of the history of localization research is not encouraging. From the time of Gall and Spurzheim, there have been recurrent efforts to pursue research based on the “localization” hypothesis. Both invasive and noninvasive techniques have historically been applied to the problem. Although, many of my readers will not agree with me on this point, it seems that past efforts to apply what seemed to be a promising new technique eventually faded from the scene leaving precious little scientific residue. Extirpative brain surgery, chemical and electrical
stimulation, EEGs, and EVBPs, among others, have come into momentary prominence, generated a lot of excitement, and then all but disappeared. Of course, these research programs and techniques told us something, but all too often it is information restricted to their own microuniverse of research. Rarely has the initial promise and excitement been generally fulfilled.

Given the relatively modest contributions of past attempts to localize cognitive processes in the brain, my goal in this book is to examine the lexicographic, psychological, methodological, technical, and logical foundations of the extensive current effort to carry out the modern version of this task. I argue that the complete system, composed of the brain, the cognitive processes, the research paradigms, and the hardware, is so complex that the opportunities for misunderstanding and misinterpretation are enormous.

It is essential to present a discussion of the technical strengths and weaknesses of these wonderful devices as well as to consider the physiology and psychology underlying the localization enterprise. Researchers today are often far removed from the details of their laboratory and the new equipment; thus they often tend to overstate the implications of their findings, not so much in the often-arcane terminology of the scientific literature, but rather in the false inferences that they leave behind. An entire chapter of this book, therefore, is dedicated to the explanation of both classic and modern research tools. In addition to considering how the older techniques (e.g., brain recordings and conventional X-rays) were used, I present an introduction to the technology of the modern imaging devices including CAT, PET, MRI, and fMRI scanning systems.

However, the hardware and psychobiological constraints are not be the main problem. Indeed, as I studied the localization problem, it seemed to me that a major difficulty encountered in this field pertained to the meaning attributed to and even the reality of the cognitive entities that were being “localized”. My chapter on the history of cognitive modules and faculties is intended to make it clear that, not only do we not yet agree on the definition of these usually vague constructs, but also there has been relatively little convergence on a widely accepted classification system for mental, psychological, or cognitive (your choice of these three near synonyms) components. Arbitrary experimental operations have often become the source of the meaning of cognitive processes. Indeed, I argue that it is not possible to define the cognitive entities that are to be localized without circularity and imprecision. The question repeatedly arose: Is it possible to localize such phantoms – such elusive “hypothetical constructs” – in the brain?

Once past the formidable obstacle to cognitive localization of ill-defined cognitive modules, my book goes on to deal with the logic of experimental protocols, of neurophysiology and neuroanatomy, and, much more fundamentally, of the concepts, both implicit and explicit, that underlay the entire enterprise. I explore some of the persisting problems that confronted classic studies using ablation, stimulation, and recording and show how they have once again become salient in the cognitive imaging studies. The well known, but largely overlooked, points of poor
regional demarcation, broad distribution of activity, and complex interconnection are all considered.

What has been perpetuated and repeatedly confirmed through the centuries of concern with the mind-brain problem has been its basic tripartite nature. Few would argue today that we have neural components located in different parts of the nervous system specialized for input and output, on the one hand, and more complex and larger regions that act as integrative or associative mechanisms linking the inputs and the outputs, on the other. However, this tripartite division of the nervous system has been known since antiquity. Beyond that model, the identified divisions, faculties, or modules of the mind have always been idiosyncratic and arbitrary.

The conceptual issues that are deeply embedded in the identifiable assumptions of localization research must also be examined. For example, it is only infrequently expressed that lesion or imaging experiments can only confirm necessity (i.e., participation) but cannot produce unequivocal arguments of the sufficiency of a particular region to encode a cognitive activity. The sufficiency of a brain region to encode a cognitive process is an inevitable and potentially incorrect logical derivative of localization-type thinking. Whenever a region is designated as the “site” of a particular kind of cognitive activity, there is a strong suggestion that it alone is responsible (i.e., sufficient) for the encoding and representation of that activity. The prototypical statement is – “Our study places the cognitive process of X in the Y brain region”. I argue that this statement in most cases is much too strong a conclusion given the fragility of the logical chain and the empirical findings that led to it.

I also discuss the idiosyncratic nature of human neuropsychological research findings. From one point of view, the cognitive deficits associated with brain injury seem to be the most direct evidence for localization. However, the cognitive deficits produced by trauma or surgery are as ill defined as is the extent of the lesions. The uncontrolled variability of both sides of the equation makes the meaning of these unique cases elusive, at best, and totally incorrect, at worst. Few neuropsychological researchers today have heeded the warnings of their predecessors concerning either the gradient of effects produced by a lesion or the uncertainty of its extent.

The design of experiments specifically intended to relate cognitive functions to brain regions is also examined. For example, arbitrary thresholds, the frailties of dissociation methods, and the misuse of complex statistical analyses are considered. All contribute to the complexity of the problem and to the emerging doubts about the validity of the entire localization enterprise.

I also argue that there is a clear distinction to be made in the underlying conceptual structure of the localization theorists and those who support broadly distributed brain mechanisms as the true mechanism for any cognitive process. However, this distinction is often lost in the complex biochemical instrumentation, statistical processing, and logical chain of assumptions underlying the localization research program. In the final chapter, I trace the details of this intricate logical
chain from the idea that glucose consumption is associated with neural activity to the final conclusion that a particular area of the brain is the locale of such an intangible entity as, say, “executive decision making”. I also trace out the logic of the alternative view and point out that only a few key differences in the logical chain are required to produce antithetical interpretations of the nature of the relationship between brain and mind.

Another major problem typically ignored by localization researchers is the complexity of the nervous system. Not only is the brain made up of many parts that are heavily interconnected, but it is also a nonlinear system in which the same molar behavior can be produced by a wide variety of different system organizations. I discuss the work of several investigators who argue on this basis that the observations of overall system activity (i.e., behavior) cannot, in principle, be used to define the internal system organization. Nevertheless, such arguments are ubiquitous throughout the contemporary localization literature.

Practical data problems are also pervasive in the localization research program. Replication is rare, vaguely defined cognitive modules have been attributed to virtually every area of the brain, and the frontal area of the brain responds to almost any cognitive activity (it is likely that these results are not incorrect, but reflect the actual distributed nature of the neural mechanisms underlying behavior). I also highlight very recent research (circa 2000) that challenges the basic localization hypothesis. New techniques (such as the event related fMRI) are producing strong results that are leading investigators to reject the localization hypothesis in favor of a much more difficult, but perhaps more valid, organizing concept – that the broad distribution of brain components and their highly complex interactions actually provide the basis for cognition.

The nearly universally used criterion for localizing a cognitive process (the region displaying the maximum amount of differential activity – the “hot spot”) is also identified as a problem for this kind of research. By shifting one’s standard criterion, activity can be made to appear to be either highly localized or broadly distributed. Furthermore, the basic subtraction technique (the identification of the hot spot by subtracting a control image from an experimental one) used in virtually all imaging laboratories obscures the fact that most of the cerebrum is active during cognitive activity of any kind. Furthermore, there is a tendency to ignore regions of diminished activity – regions that are also critically involved in the cognitive process under study reflecting inhibitory or disinhibitory neural interactions.

In short, my argument is that the enormous effort currently being made to localize vaguely defined cognitive modules as attention or language in particular locations of the exceedingly complex brain is an ill-chosen path for cognitive neuroscience. I believe that it is based on incorrect a priori assumptions that, on close examination, cannot be justified and data that is highly tainted by these incorrect assumptions. Such an approach is leading us astray from thinking (correctly, I argue) about cognitive activity being broadly distributed in the brain. Such a strategy has the serious consequences of creating what we will all ultimately agree
are scientific absurdities. The emerging discussion of “neurotheology” is among the most bizarre outcomes of an extremist localization theory.

My concluding prediction in this book, however, is that we are not likely in the near term to resolve completely any of the issues that I raise. It is just too easy to use the basic localization assumption as a vehicle for misleading, but convenient, research studies. The contradictory idea of a complex, interacting system of brain modules (of unknown number and size) is much more difficult to conceptualize and study. Indeed, in many cases, because of this complexity, long-term solutions are simply not obtainable. It is entirely possible, I argue, that there are well established insurmountable practical and “in principle” barriers to achieving complete understanding that are not yet fully appreciated by many researchers in this field.

There is, in addition, another somewhat sensitive issue that is raised. The entire localization enterprise has taken on a life of its own. Great laboratories have sprung up with multiple expensive imaging devices. Vested economic interests now pollute what should be careful consideration of scientific fundamentals and special scrutiny of the meaning of the empirical results produced by this new version of localization research. Young scholars feel they have to go “where the money is” and, without question, money has been redirected from more conventional kinds of cognitive research. The point I make is that extraneous issues like these are now playing an inappropriate role in this important scientific discussion.

I also appreciate that there are no “killer arguments” that can resolve all of the broader issues that circulate around the localization approach. However, to ignore these issues and simply to claim that we must plunge ahead to make “progress” is irresponsible. If my book does nothing more that raise the entire matter to heightened attention and consideration, I will have achieved my purpose.

Finally, I must also emphasize one other general point. Nothing in any of my books has been intended to discourage explorations and experimentation in either neurophysiology or psychology. Authenticating the achievements of the former and satisfying the need for understanding in the latter should both be high priority goals of science. Furthermore, it takes no great insight to appreciate that the imaging techniques, irrespective of their application in localization research, are, without question, one of the most important medical and scientific developments of the 20th century, certainly comparable to the invention of X-rays in the preceding one. My argument is only with the building of fragile and potentially incorrect conceptual bridges between the two fields. In “simpler” scientific fields, reductionist approaches comparable to the localization hypothesis have worked and will continue to contribute to our understanding of the world of which we are a part. In the study of brain and behavior, however, the multidimensional complexity of the problem introduces new constraints that are not likely to be overcome at any time in the future. To plunge ahead in the hope of solving what a deep analysis suggests are intractable problems flies in the face of new developments in other sciences.
Note


Table of Contents:

Preface
Acknowledgements
1. On Localizing Cognitive Processes: The Problem
2. The Technologies
3. Is A Taxonomy or Even a Lexicon of Cognitive Processes Possible?
4. Technical and Conceptual Problems
5. Summary and Conclusions
Appendix A: The Great Questions of Scientific Psychology
Appendix B: A Comment on the Zeitgeist
References
Index