Philosophical Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cphp20

Let the Brain Explain the Mind: the Case of Attention
Maria Ruz

Version of record first published: 12 Dec 2006

To cite this article: Maria Ruz (2006): Let the Brain Explain the Mind: the Case of Attention, Philosophical Psychology, 19:4, 495-505
To link to this article: http://dx.doi.org/10.1080/09515080600806583

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Let the Brain Explain the Mind: the Case of Attention

Maria Ruz

Oversimplified conceptions of cognitive neuroscience regard the goal of this discipline as the localization of previously discovered and validated cognitive processes. Research however is showing how brain data goes far beyond this translation role, as it can be used to help in explaining human cognition. Knowing about the brain is useful in building and redefining taxonomies of the mind and also in describing the mechanisms by which cognitive phenomena proceed. The present paper takes the cognitive system of attention as a model research field to exemplify how biological knowledge can be used to advance the psychological theories explaining mental phenomena.

Keywords: Functionalism; Human Cognition; Mind; Brain; Attention

1. Introduction

Characterizing how the human mind works is one of the greatest challenges for current science. In the last decades, the introduction of neuroimaging techniques has fostered the union among several disciplines aimed at studying how the brain builds up the mind (Gazzaniga, Ivry, & Mangun, 1998). This results from the combination of behavioral research in cognitive psychology, brain data from neuroimaging techniques, cognitive deficits in neuropsychological patients, electrophysiological experiments in nonhuman primates, biologically driven computational models, and recent genetic studies. Indeed, the use of convergent evidence coming from different fields and research modalities is the most powerful characteristic of cognitive neuroscience. Despite the youth of this discipline, considerable progress has been made in data and theoretical models that propose specific mappings between cognitive functions and brain structures (Gazzaniga, 2004; Humphreys, Duncan, & Treisman, 1999). Many of these investigations represent examples of
how the different levels of analysis applicable to complex systems such as the human mind-brain can exchange information that generates bidirectional constrains and promotes the advance of knowledge.

The core message of the present paper is that data on the brain will play an essential role in explaining the human mind. Many theorists, however, understand the sole role of research in cognitive neuroscience to be the localization of already described and validated cognitive processes or mental states in the brain. This oversimplified conception of the relevance of brain data for cognition leads to many criticisms regarding the impossibility of such localization (e.g., Uttal, 2001) or the lack of relevance of localization in the brain to the psychological theory (e.g., Fodor, 1999). Things change however when the brain is allowed a role in explaining the mind by helping in the development of taxonomies of cognitive processes or by uncovering the mechanisms by which they exert their function.

The goal of this article is to show that an optimistic position regarding the usefulness of brain data to the psychological theory is not only wishful thinking but relies on theoretical and empirical facts. In an attempt to highlight the relevance of neuroscience data to cognition, I describe how research in the field of attention has benefited enormously by the consideration of the brain as a valuable source of information. First, cognitive neuroscience has supported models affording attention the status of a cognitive system by showing how it is composed by three attentional networks that map onto a set of reliable brain regions linked to specific neurotransmitters, separable genetic bases and developmental courses. In this way, attention has changed in the last few decades from mainly being a variable used to explain several research problems, to being an autonomous system that has the right to be explained on its own. Second, brain data has facilitated a change in the focus of research from attentional effects on behavior to the study of the attentional mechanisms or the processes by which attention biases the flow of information in other processing systems.

I begin by briefly summarizing two of the most important set of critiques made of cognitive neuroscience research. Next I will try to show how these critiques do not hold true for the study of attention, a field in which brain data is proving to be highly valuable to advancing our knowledge.

2. A Case for the Futility of Brain Data

Among the critiques made of the study of the biological basis of cognition I will consider the two most relevant to the field of attention. The first qualifies research in cognitive neuroscience as a new wave of phrenology, and regards the understanding of higher cognitive functions in the brain as impossible given the lack of modularity of these processes and the nonlinear and dynamic nature of brain functioning (Uttal, 2001). The second line of criticism claims that cognitive processes are independent of its biological implementation, and thus the information about the brain is not relevant to explain how cognition takes place (Fodor, 1999).
2.1. The Impossibility of Mind-Brain Decomposition

In 2001 William Uttal published a book entitled *The New Phrenology* that described many of the problems faced by contemporary research in cognitive neuroscience. Because of these shortcomings, the author claimed that the mapping of cognitive functions in the brain is actually impossible and thus that the discipline is useless. Uttal equates inquiries in modern cognitive neuroscience with those posed by phrenologists lead by Gall in the 18th century, because both accept that the mind is composed by many different functional parts that can be localized in the brain. But according to Uttal, this localization approach is grounded on false assumptions and thus untenable.

First, Uttal (2001) contends that the existence of “a taxonomy of cognitive functions” is questionable given the nonmodularity of high-level processes and thus the search for it is likely to be unsuccessful. Moreover, the dynamic and nonlinear characteristics of brain functioning make it even harder to map cognitive processes on to brain structures, as it is impossible to confidently relate a brain area with a specific function. The scenario posed by Uttal is even worse because the limitations of current neuroimaging techniques make data interpretation and inferences drawn from them severely limited. In light of this, Uttal concludes that the right way to make science progress is to abandon the search for unobservable internal mechanisms and its brain basis and to go back to a “molar” analysis of behavior in terms of its observable inputs and outputs.

2.2. Functionalist Foundations of Cognitive Science

In contrast to Uttal’s position, functionalist doctrines accept that the mapping of cognitive functions on to brain structures possible. However, this localization knowledge is seen as useless for the science of the mind.

The end of the behaviorist era around the 1950s was enabled by the information-processing revolution, based on the so-called *mind-computer analogy*. One of its basic assumptions is that the human mind can be viewed as a complex information-processing machine, and thus it can be decomposed into different functional modules, each specialized in the performance of a set of basic cognitive processes. This set of cognitive systems are further decomposed into more detailed representations and processes, in a recursive manner up to the point of elementary mental operations (see Posner & Rothbart, 1994). David Marr (1982) stressed the notion that there are different epistemic points of view from which complex information-processing systems can be studied. Marr noted that there is no single view of a complex system that explains everything about it. A complete understanding of a system needs a computational theory explaining what the system computes and why it does so, a description of the representations and algorithms that the system uses, and also knowledge about its physical implementation. Although responses at the three levels must be obtained in order to gain a complete understanding of the system, research can be done in each of the levels without knowledge of results in the others.
The same functions and computations can be carried out by very different physical substrates and, for this reason, knowing about the implementation of a given process in the brain is not needed to be able to obtain a complete understanding of the cognitive processes comprising the human mind. An accurate and coherent understanding of the mind will be provided by the psychological theory alone and, once theories are validated and accepted, the set of processes can be translated to their brain basis. As Fodor (1999) jokes, who cares if thinking of teapots happens to be side by side in the brain with taking naps?

3. Attention as a Model of the Usefulness of Knowing About the Brain

It is certainly not an easy task to develop an accurate taxonomy of mental processes. And if such a quest were truly impossible, as some authors claim, incorporating fallible new technologies and brain data would do nothing but add confusion to the story (Uttal, 2001). Many authors, however, do not agree with these conclusions. First, science is a self-corrective enterprise. Theories and taxonomies of the mind proposed today are only in their early stages; they are not definitive but, rather, subject to continuous revisions (Bechtel, 2002; Hubbard, 2003). Second, brain data, far from being useless to the cognitive domain, is one of the relevant sources that can be used to support, modify or falsify proposed mental taxonomies and can even offer insights for devising brand new dissections of mental processes (Churchland, 1986; Posner & DiGirolamo, 2000). Moreover, the limitations of individual research methodologies can be made much less problematic by appealing to the convergence of evidence from multiple sources. This is precisely the most powerful strength of the cognitive neuroscience approach to the study of the human mind and brain (Gazzaniga et al., 1998). This convergent discipline is being successfully applied to many research fields such as perception, memory, learning, executive functions and consciousness (Gazzaniga, 2004). Within these, the study of attention has been one of the fields that has benefited most from this approach (Posner, 2004), due to several reasons outlined below.¹

3.1. Three Networks as a Taxonomy of Attention

‘Attention’ is a word that pervades our daily vocabulary. The Oxford English Dictionary defines it as “the act or state of attending especially through applying the mind to an object of sense or thought” or as “a condition of readiness for such attention involving especially a selective narrowing or focusing of consciousness and receptivity.” In 1890, William James proposed a psychological definition of attention that appealed to the subjective knowledge of the reader and that has become a reference point in the literature:

Every one knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, or consciousness are of its essence. It implies withdrawal from some things in order to deal effectively
with others, and is a condition which has a real opposite in the confused, dazed, scatterbrained state... (p. 404)

A century later, however, Harold Pashler (1998) captured the confusion in the field when he wrote that “No one knows what attention is, and... there may even not be an ‘it’ there to be known about (although of course there might be)” (p. 1).

The tension between these two assertions exemplifies the difficulty of finding a scientific definition of attention to which everyone agrees. The fact that attention is a concept incorporated into science from folk psychology does not help, but the core problem arises from the fact that attention, as a psychological process, is not a unitary phenomenon (James, 1980; Posner & Boies, 1971). Traditional views of attention emphasized its “selective” (Broadbent, 1958) or “energetic” side (Kahneman, 1973). Whereas selection theories aimed at localizing the point in the processing chain (from perception to motor responses) at which irrelevant stimuli were filtered out (i.e., the location of the “bottleneck”), energetic views investigated how resources were divided among tasks or the unitary or multiple nature of this “attentional energy” (Wickens, 1980). However, it was not clear how these phenomena related to each other and how they fit into a general theory of attention.

Trying to define attention as a whole, and attempting localize it as such in the brain is a potentially fruitless approach. However, using a decomposition strategy (Bechtel, 2002) to develop a taxonomy of attention at the appropriate level of explanation (Posner & Rothbart, 1994) may be extremely useful for suggesting provisional definitions of what attention may be. The attentional networks proposed by Posner and collaborators represent an example of a taxonomy of a cognitive process that nowadays receives support from several converging methodologies. Here knowledge about brain areas acts not only as the backbone of the proposal but has also been used to constrain, redefine and suggest new ideas for organizing the attentional networks.

Attention can be conceptualized as a system of different anatomical areas that is composed by three distinct and specialized modules (see Posner & Fan, 2004, for an extended description). The achievement and maintenance of an alert state is mediated by the alerting network; the orientation to sensory objects is carried out by the orienting network; and the executive network is responsible for monitoring and resolving conflicts between computations taking place in different brain areas.

The alerting network, conceptually related to the early “energetic” theories of attention, generates changes in the preparatory state of the organism in expectation of an incoming stimulus. The paradigms employed to study it involve vigilance tasks and the use of warning signals. Anatomically, it comprises areas of the right parietal and frontal hemispheres, and it is mainly related to the neurotransmitter norepinephrine (NE) that arises from the locus coeruleus. Lesions to these brain regions generate deficits in alerting, and drugs modulating NE levels affect its functioning while leaving other attentional dimensions intact (Posner & Fan, 2004).

The selection of stimuli from sensory input is the role of the orienting network, which has been mainly studied using “costs and benefits” paradigms (Posner, 1980).
Voluntary changes in the allocation of attention are mediated by areas in the superior parietal lobe, together with the frontal eye fields, superior colliculus, and some nuclei in the thalamus. Reflexive or automatic changes in attention generated by the appearance of unexpected stimuli at novel locations seem to be resolved by activations in the temporo-parietal junction (TPJ). The neurotransmitter acetylcholine (Ach) originating in the basal forebrain is the main chemical regulating the orienting function, as drugs affecting this neurotransmitter modify indices of orienting with no impact on alerting. Many neglect patients, who suffer a deficit in orienting to and detecting events appearing contralateral to their brain lesions, have damage affecting their right TPJ.

Finally, the classic selective nature of attention is reflected in the executive network, which monitors and resolves conflict between different representations in the control of behavior. This usually happens in situations requiring action planning, decision making, error detection, novel or difficult situations, or when overcoming a habitual response is needed. Research using the Stroop task has shown the relevance of lateral prefrontal regions, the anterior cingulate cortex (ACC) and basal ganglia in mediating executive attention. The neurotransmitter most relevant in this case is dopamine (DA) from the ventral tegmental system, and its imbalances are known to affect executive functions. Lesions in these brain areas can lead to deficits in voluntary behavior, loss of planning capabilities, or lack of appropriate social behavior.

In recent years there has been extensive research aimed at supporting this mental taxonomy, with converging evidence coming from different domains. As described above, the three networks have been associated with different psychological functions and brain regions that use distinct neurotransmitters as signals. This is supported by several lines of research including neuroimaging experiments, studies of differential functional loss after focal brain damage, and dissociations using drugs targeted at specific neurotransmitters in alert monkeys and humans (see Posner & Fan, 2004). The ANT (Attention Network Testing) task was specifically devised to obtain behavioral indices of the functioning of the three networks in different populations: children, healthy and neuropsychological adults, psychiatric patients and nonhuman primates (Fan, McCandliss, Sommer, Raz, & Posner, 2002). Using this tool as a model to both operationalize and quantify attention processes, it has been shown that reaction time and accuracy measures of each network are uncorrelated with one another, in the sense that the same individuals can have differential efficiency in the three functions, which strengthens the idea that they are differentiable. Also, the combination of this testing tool with genetic methods suggests that the three networks are linked to distinct genes and have independent heritability, which stands in agreement with the data supporting their different developmental courses in infancy and their separable involvement in some psychiatric pathologies (Raz & Buhle, 2006).

Moreover, results obtained from different neuroimaging methodologies have shown that the classic notion of a single bottleneck that filters out irrelevant information is incorrect. Instead, stimuli can be selected at many different levels
of processing, from early perceptual stages to motor preparation, depending on the
particular task requirements (see, e.g., Luck & Hillyard, 1999). This kind of research
has also shown that on most occasions irrelevant nonselected information is not
wiped out from the system, but is instead processed in different brain regions
(e.g., Ruz et al., 2005). Electrophysiological results have also been highly influential
on energetic theories in the field by, for example, showing that automatic warning
signals do not affect late motor processes (as it was widely assumed) but rather speed
earlier processes dealing with response selection (Hackley & Valle-Inclan, 1999).
Thus, neuroimaging data has changed the theories and consequently the direction
of the research carried out.

The differentiation of attention in three distinct but interrelated networks can be
used as a model taxonomy of a complex mental phenomenon that relies on brain
structure and psychological functioning as an essential backbone. In this case,
knowledge of the localization of cognitive functions in the brain was not attained
only after the mental processes had previously been described. Rather, partial
biological information from neuropsychology and animal neurophysiology was used
to help build the taxonomy from the beginning (Posner & Petersen, 1990). With
time, the acquisition of new behavioral and brain data pushed the modification and
improvement of the categories proposed for attention. This has led to a refining
of their alleged computational functions (mainly of the executive and orienting
networks) and of its constituent parts (Posner & Fan, 2004). This represents a true
example of how interaction between data and constructs of cognitive psychology
and neuroscience can be used to build a taxonomy of a cognitive function that is
supported from several different domains of knowledge.

3.2. The Mechanisms of Attention

Brain data can be used as a driving force in building taxonomies of the mind. But
this is not the only way neuroscience helps cognitive theories. One of the criticisms
of brain research is that it is of no use in understanding cognitive processes. From
this perspective, these cognitive processes can be explained and understood a priori;
neuroimaging could be used in a second step just to localize in the brain previously
described cognitive phenomena, but this brings no new information about these
processes (e.g., Fodor, 1999). However, cognitive neuroscience can also help in
understanding the processes per se—i.e., rather than just translating to the brain
previously discovered phenomena, brain data has a crucial role in explaining
how these cognitive mechanisms work.

For many years, behavioral research was the principal source of investigation
in cognitive psychology. During all this time, the field of attention was populated
with a vast list of valuable information about the behavioral effects of manipulating
attention in different ways and about attention-related phenomena. In the last
few decades, however, electrophysiological recordings in conscious monkeys, and,
more recently, neuroimaging studies of humans, have boosted our knowledge
of the mechanisms by which attention operates in the brain. Here it is highly useful
to differentiate between the neural *sources* of attention modulation and the *sites* that are affected by this attentional modulation.

One of the oldest ideas about attention concerns its selective nature. We are surrounded by a myriad of stimuli and the system is in need of selecting only a few of them to control our behavior. Why is this so? Although the teleological nature of this is still unknown, in recent years we have gained some knowledge concerning why this happens (Desimone & Duncan, 1995; Duncan, 2004). Electrophysiological recordings in monkeys show that stimuli compete for representational resources in the brain by mutually suppressing their respective neural responses. This basic effect is observed when the response of a neuron, which is maximal after the presentation of an effective stimulus in its receptive field, is significantly reduced by the presentation of a second stimulus within the same field. That is, the addition of the second stimulus suppresses the effective response of the cell to the first stimulus. In contrast, the discharge of the neuron is not affected if the competing stimulus is presented outside its receptive field, showing that the suppression is driven by mutual interactions within the region of space in which the cell is interested. This effect has been shown in several different places in the brain, suggesting that it is not an isolated but a general phenomenon. However, when attention is directed to the attribute that the neuron represents, this suppressive influence from other stimuli is eliminated and the cell gives its maximal response even in the presence of competing stimuli. That is, attention biases the competition between stimuli and helps the relevant attribute (either a kind of stimuli, a specific spatial location or an attribute such as motion) gain representational resources and dominate the neural response. These attentional modulations have been shown for monkeys and humans in many locations of the brain, including both the dorsal (occipito-parietal) pathway representing spatial locations and the ventral (occipito-temporal) route affecting representations of objects (Kastner, 2004).

So far I have described the mechanisms operating at the *sites* of attentional modulation; those mechanisms by which attention affects processing by biasing competitions throughout the brain. At the same time, there have been models devised to explain how the putative attention regions, the so-called *sources* of attention, operate to be able to bias processing in other brain systems. The theory of Cohen and collaborators (Cohen, Aston-Jones, & Gilzenrat, 2004; Cohen, Dunbar, & McClelland, 1990) on cognitive control (i.e., executive attention) is a good example of such a model. The “guided activation theory of cognitive control” proposes that this control is achieved by means of the activation of the neurons in prefrontal cortex (PFC) representing the appropriate goal for the situation. These goal representations, or task-demand units, generate a bias that guides the flow of activity in the system along the pathway of their associated units and attenuates processing in other paths not relevant to perform the goal. The theory of Cohen and collaborators, apart from localizing where in the brain goals are primarily represented, specifies the *functional requirements* that such goal representations must fulfill to be effective and how these are achieved in the brain.
First, the goal representations have to be actively maintained in PFC in absence of environmental support while the relevant action is performed, avoiding distraction by resisting perturbation from stimuli irrelevant for the task. This function is achieved by means of self-excitatory recurrent connections that give rise to attractor dynamics. These connections allow the appropriate goals to be activated and to bias processing in the relevant systems until the intended action is finished. Secondly, however, the task-demand units must also be able to avoid perseveration—i.e., the representations have to be adaptively updated once the outcome has been achieved and the action is no longer appropriate. This relies on an adaptative gating mechanism mediated by DA from the ventral tegmental area, as well as structures in the basal ganglia. Here, a transient gating signal renders the task-units temporally sensitive to inputs from other neurons that activate the representation of new goals. Due to reinforcement learning, the system knows when to produce a gating signal, by using the DA release as a learning cue that reinforces those associations that predict a better reward. Finally, the system needs a means to know in which situations control is needed and how much of it is required. The theory proposes that the ACC is the structure responsible for detecting the degree of conflict between different goals.2 The higher the conflict, the more control is needed and PFC representations must be activated more strongly. Hence, whereas the ACC is responsible for conflict detection, the PFC and basal ganglia are the centers in charge of allocating control resources to the brain pathways able to realize the desired outcomes.

My description has tried to convey that asking how the brain works by means of different technologies is useful for understanding and theorizing about the mechanisms by which cognitive processes operate. The ideas presented are informed mainly from neurophysiological recordings in monkeys, neuroimaging studies in humans, and biologically driven computational models—all of them aided by tasks devised during many years of behavioral research in cognitive psychology. This convergent brain imaging effort has brought work on mechanisms to the fore of attention research, a move that would have been much more difficult with behavioral methods alone.

4. Conclusions: Mind the Brain

Some decades ago, the main goal of cognitive science was to determine the computations of intelligent systems. Thanks to the acknowledgment of a specific level of analysis for cognition, research in this area considerably advanced our knowledge of how humans represent and process information. The independence assumption stemming from the functionalist doctrine was adopted in cognitive science and thus the role for biological data was left as a posteriori description of already described mental phenomena. More than fifty years later, technological developments allow us to translate questions of cognition to ones about the human brain. Cognitive neuroscience is turning out to be a main source of knowledge regarding the neural mechanisms of cognitive processing. The fast development
of cognitive neuroscience is offering an explanation of human cognitive functioning where Marr’s levels of analysis, or function and realizer, are no longer autonomous. Thinking of the human mental operations and their functions as completely independent of their material substrate fails to reap the benefits of recent technological and conceptual developments. Indeed, in a not-so-distant future, theories explaining human cognition may use concepts in which mind and body are no longer understood as independent phenomena.

I have described how neuroscientific data has been a crucial source of information for developing a taxonomy separating attention in three independent networks; how biological knowledge has boosted research on the mechanisms operating at the sites of attentional influence, and also at brain sources of attention. I am aware that the models and ideas presented here are far from being definitive. Understanding what attention means, how to properly dissect it and which set of brain structures and mechanisms are responsible for the myriad of attentional effects described in the literature will be a long-term pursuit. However, advances obtained so far in the cognitive neuroscience of attention suggest that it is actually possible to devise a taxonomy of a high-level system that is supported from many different sources of investigation, in spite of limitations in current technology and the nonlinearity of brain dynamics. Moreover, in doing so they show that taking into account brain functioning is essential. Time will tell whether this integration of knowledge will eventually take us to a complete and coherent understanding of human cognition in terms of the functioning of the brain. Hopefully the discussion presented in this paper can be taken as an indication that we are on the right path.

Notes

[1] In the last few decades, research on attention has exponentially increased. It is out of the scope of the present paper to review all the neuroscientific data that bears on questions related to attention theory. The discussion is thus limited to a few complementary well-known theories in which the implications of brain data are central to cognitive explanations.

[2] Note that this cognitive conflict monitoring hypothesis was developed after various neuroimaging experiments showed that the ACC was activated in many situations requiring effortful control. This represents an example of how a result coming from neuroimaging methodologies generated the posterior elaboration of a cognitive theory devised to explain it.

References


Oxford English Dictionary.


