Cognitive Neuroscience: Origins and Promise

Michael I. Posner
Sackler Institute

Gregory J. DiGirolamo
University of Illinois at Urbana-Champaign

Both Freud and Wundt had hoped to base psychology on an understanding of the neural basis of mental events. Their efforts were unsuccessful because the structure and function of the human brain was not available for empirical study at the physiological level. Over the last part of this century, there has been amazing growth and vitality in the field of human brain function. In this paper, we trace critical developments in the fields of cognitive psychology, neuropsychology, and brain imaging related to the development of cognitive neuroscience. Cognitive Neuroscience has established that the decomposition of mental events can be unified with an understanding of the mental and emotional computations carried out by the human brain. Cognitive neuroscience has the capability of influencing psychology in diverse areas from how children develop to how adults age; from how humans learn to how we imagine; from volitional control to psychopathologies.

In 1987, the James S. McDonnell Foundation held a summer institute at Harvard University on cognitive neuroscience. Stephen Kosslyn was the Director of the Institute, and under his leadership, an outstanding program of speakers was assembled. Michael Gazzaniga had persuaded John Brer, the President of the McDonnell Foundation, that it might be a good time to invest foundation funds in the creation of this new field. Michael Gazzaniga had already coined the name and had used it as the title of an institute that he had developed as part of the cognitive science initiative started a decade earlier.

There is a long history of investments from foundations creating new disciplines in scientific fields. For example, 2 decades earlier, the Sloan Foundation invested in a number of fields, concentrating in neurophysiology, and helped to create the field of neuroscience. In terms of the sheer numbers of scholars attracted to a field, this venture has to be one of the best investments ever made. When the Sloan Foundation decided to act again to create the new field of cognitive science, they included psychology, linguistics and neuroscience as subdisciplines. Of course, the field of language was of major interest, and one of the earliest meetings in cognitive science was organized by Jacques Mehler (Mehler, Walker, & Garrett, 1982) and was devoted almost entirely to this area. This meeting introduced European scientists to the broad range of work in cognitive science. At this meeting, one of us (Posner, Pea, & Volpe, 1982) presented what was perhaps the first paper with the general theme of an emerging discipline: "Cognitive Neuroscience: Toward a Science of Synthesis." Although cognitive science continues as an important enterprise, it has never had the explosive growth and influence of the earlier neuroscience initiative. Though the success of a field cannot be insured by financial investment alone, a powerful scientific agenda can attract the best minds into a new field and increase its chance of success.

We think that cognitive neuroscience has just such an agenda. As cognitive neuroscience enters the new millennium, it seems an appropriate time to reflect on its history and to consider where the field is headed. In this article, we attempt to outline some of the intellectual background and discoveries in cognitive psychology, neuropsychology, and neuroimaging that provided a foundation for the genesis of cognitive neuroscience. Because we are not historians, the treatment is biased toward events and views that relate more closely to our own thinking and areas of expertise; nevertheless, we hope that this overview captures the basic ideas and methods of the field.

Every psychologist who entered the field in the mid-20th century studied the history of efforts to localize higher level mental function in the brain. Students of the field were taught that mental function could not be localized. Karl Lashley (1929) had discovered that interference with maze learning in rats depended on how much tissue was removed and not on what part of the cortex was removed. Studies of patients following traumatic frontal lobe damage were also interpreted as showing that the brain operated as a whole during cognition. The principles of mass action and equipotentiality were used to explain failures to find any precise localization of cognitive function. Further, caution was extended to efforts to analyze cognitive processes into elements because the whole was then thought to be different than the sum of the parts. More specifically, it was thought that tasks were completely altered when any part was changed.

As George Miller noted (Miller & Lenn, 1978), it is often not obvious to psychologists why the issue of localization of function is of importance because knowing what part of the brain
is involved may tell little about what cognitive operations are occurring. Miller (p. 6) suggested that the issue of localization is related to other, more general assumptions necessary for understanding the relationship between the human brain and behavior. These are:

1. Complex brain functions can be decomposed into simpler, more general processes.
2. These components can be localized anatomically and studied in relative isolation.
3. The simpler brain processes can be correlated directly with their simpler behavior processes.

In the last several years, functional human brain imaging has validated many of these assumptions. To grasp the scope and significance of cognitive neuroscience, it is important to trace the developments in cognition and neuropsychology, together with brain imaging, that have produced this new synthesis. Table 1 outlines the overall goal of using mental operations, as defined in cognitive studies, to link cognition to neural systems.

We argue that development in the understanding of localization of function and the neural circuitry of the brain can influence theories and assumptions about higher cognitive processes. Specifically, we review changes in the conception of major psychological issues (such as priming, attention, and executive control) resulting from the growing integration of neuroscience data with cognitive data. Finally, we consider developments likely to occur in the near future.

Cognition

Almost from the inception of efforts to develop models of human cognition, investigators felt there would be a time when cognition would be linked to developments in the brain sciences. Herbert Simon (1981), whom many consider the founder of cognitive science, remarked in his Compton Lectures:

The main reason for the disembodiment of mind is, of course, the thesis that I have just been discussing. The difference between the hardware of a computer and the hardware of the brain has not prevented computers from simulating a wide spectrum of kinds of human thinking. It would be unfortunate if this conclusion were altered to read that neuropsychology has nothing to contribute to explanations of human behavior. This would be a ridiculous doctrine.

Table 1

<table>
<thead>
<tr>
<th>Level</th>
<th>Example</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive system</td>
<td>Language, attention, motor control</td>
<td>Verbal report</td>
</tr>
<tr>
<td>Elementary operation</td>
<td>Next, rotate, zoom</td>
<td>Artificial intelligence</td>
</tr>
<tr>
<td>Pathway activation</td>
<td>Facilitate, inhibit</td>
<td>Cognitive studies, neural networks</td>
</tr>
<tr>
<td>Neuro system</td>
<td>Parietal cortex, basal ganglia</td>
<td>PET, ERP, MRI</td>
</tr>
<tr>
<td>Microcircuits</td>
<td>Visual lamina, area V4</td>
<td>Cellular recording</td>
</tr>
</tbody>
</table>

Note. A framework for linking cognitive tasks of daily life (top row) to brain systems (bottom two rows). This article concentrates on the middle rows (rows 2 and 3) by discussing how elementary mental operations can be localized in brain tissue. PET = positron emission tomography; ERP = event-related potential; MRI = magnetic resonance imaging.

It is to physiology that we must turn for an explanation of the limits of adaptation. As our knowledge increases, the relation between physiological and information processing explanations will become just like the relation between quantum mechanical and physiological explanations in biology. They constitute two linked levels of explanation: the limiting properties of the inner system and showing up at the interface between them. (p. 97)

The level of mutual support between these two areas has proven much closer than Simon might have suggested. It is hoped that this article will illustrate both how closely linked the hardware of the brain is to the performance of cognitive and emotional tasks and the importance of environment and self-regulation to the operations of the human brain.

One of the initial bridges to creating a synthesis between cognition and brain function was developed by cognitive psychologists in the mid-to-late 1960s. It was demonstrated that a wide variety of tasks could be described in terms of mental operations (or subroutines) that took place over tens to hundreds of milliseconds (Newell, 1990; Sternberg, 1969).

The measurement of mental operations by reaction times (RTs) was always closely associated with the idea of serial processing in human performance. In the late 1960s, Saul Sternberg (1969) provided a general method for the isolation and investigation of serial stages called the additive factors method. In his task, participants determined if a presented digit was a member of a set of items that they had learned previously. His data showed that successful determination generally took place at a rate of about 30 ms per item. The speed at which people could identify even complex patterns made it seem unlikely that all tasks could be thought of as having the serial structure that this task demonstrated (Treisman & Klein, 1998). However, where such a serial organization does apply, the additive factors method remains a powerful tool for the analysis of mental processes and has been applied successfully to brain imaging data (Dehaene, 1996; Sanders, 1998).

In the mid-1980s, new developments paved the way for a connection from cognitive operations (or computations) to physical architecture (or neural structure). A number of researchers led by Rumelhart and McClelland (1986) developed ways to analyze behavior by representing psychological concepts (e.g. words) as a set of nodes connected by weights. This concept built on ideas of parallel distributed processing, for example, by representing a word as a hierarchy of features, letters, and words that were activated in parallel. Using these methods, it was possible to simulate the empirical finding that a single letter had a lower perception threshold if embedded in a word relative to being presented in isolation (Reicher, 1969). This powerful word superiority effect, which appeared so counterintuitive in the era of purely serial methods, was completely explicable if the stored information about words could be activated in parallel by single letters and thus reinforce the visual input to letter detectors.

Cognitive psychology proposed ways of dividing tasks into constituent operations and of using mental chronometry to measure operations in empirical studies. With the advent of parallel distributed processing models, there were theoretical structures in which to embed the findings of empirical studies. Now, the notion of distinct cognitive modules, each computing a different operation to achieve some behavior, led naturally to questions regarding localization of function. These developments provided important
tools to test the idea of whether these operations were localized within specific brain regions.

With the advent of neuroimaging, it has become very important to document in the millisecond range the time course of activity in neural substrates to distinguish purely sensory activation from top-down influences (Kosslyn, 1994). The importance of these time course measures is evidenced in processes involved in the perception of objects (Treisman & Kanwisher, 1998). Several types of representation are constructed by the visual system in parallel, allowing observers to recognize, prepare actions in response to, and construct preferences toward sensory events. Awareness of an object is critically dependent on the first 100 ms following input. The difference between the activation caused by novel and familiar objects can be shown in ventral visual areas within 155 ms after input (DiGirolamo & Kanwisher, 1995; Thorpe, Fize, & Marlot, 1996). Neuroimaging and cellular studies have shown that even the primary visual cortex (Posner & Gilbert, 1999) can show top-down influences; hence, a time course is needed to distinguish between sensory-driven and attention-driven activations that appear similar in hemodynamic measures (e.g., positron emission tomography [PET] or functional magnetic resonance imaging [fMRI]).

We turn to the study of neural function and localization in a semichronological order of development, and begin with lesion methods and single-cell recordings in the next section.

Neuropsychology

The term neuropsychology has been applied to both lesion studies of humans and animals and efforts to record electrical activity from individual cells (or groups of cells) in higher primates (including some studies of human patients). The development of the microelectrode had an enormous impact on neuropsychology. Starting in the 1960s, the arrival of a means of exploring the organization and responsiveness of the visual system of cats and monkeys opened up new areas of exploration into the physiology and function of brain processes (see, e.g., Hubel, 1987). In the mid-1970s, a number of researchers began recording cellular activity in alert-behaving monkeys (Evarts, 1973; Mountcastle, Lynch, Georgopoulous, Sakata, & Acuna, 1975; Wurtz, Goldberg, & Robinson, 1980). This methodology provided an important approach to understanding how the brain might carry out higher level cognitive functions like attention. In the following section, we trace studies of visual spatial attention from the development of mental chronometry for determining the cognitive operations of spatial attention to the neural areas involved in these operations, studies that were based on single-cell recordings in monkeys and lesions in humans.

Attention

The problem of attention is one of the oldest in psychology (James, 1890) and continues to be a vibrant and growing field. The dominance of behavioral psychology in the United States in the first half of the 20th century, however, postponed research into the internal mechanisms of selective attention. Integrity of the reticular formation to maintain the alert state provided anatomical reality for one aspect (arousal) of attention (Moruzzi & Magoun, 1949). The anatomical approach has led to many new findings relating the individual subcortical transmitter systems (e.g., dopamine, norepinephrine) to the computations underlying the selection of information (Marrocco, Witte, & Davidson, 1994). Other approaches to attentional selection arose in work on the orienting reflex (Kahneman, 1973; Sokolov, 1962). However, changes in the autonomic systems (e.g., skin conductance) as a measure of orienting occurred over 1–2 s and made it difficult to analyze the cognitive components and neural systems underlying orienting (see, e.g., Dawson & Schell, 1982).

Unlike sensory systems that can be studied in anesthetized animals, attention requires an alert animal interacting with the environment. In the mid-1970s, neurobiologists began to study information processing in alert monkeys (Wurtz et al., 1980). Because the visual system had been relatively well investigated using microelectrodes, much of this work involved orienting to visual stimuli. During the past 15 years, there has been a steady advancement in the understanding of the neural systems related to the visual orienting of attention. Converging results have come from studies using single-cell recording in alert monkeys (Chelazzi & Corbetta, in press; Motter, 1998), from neuroimaging (see, e.g., Corbetta, 1998; Corbetta et al., 1998), and from lesion data (see Rafal, 1996, for a review).

The initial findings demonstrated a relatively restricted number of neural areas in which the firing rates of neurons were selectively enhanced when monkeys were trained to attend to a location or object. At the level of the superior colliculus, selective enhancement was obtained only when the monkey also moved its eyes. In contrast, neurons in the posterior parietal lobe demonstrated selective enhancement even when the animal maintained fixation. An area of the thalamus, the lateral pulvinar, was similar to the parietal lobe in containing cells that demonstrated selective enhancement whenever monkeys attended (Colby, 1991). Later studies indicated that nearly all sensory and motor systems could show the influence of attention (Moran & Desimone, 1985; Roland, 1993).

Physiological studies of cats and rodents (Skinner & Yingling, 1977) suggested that selection of the relevant channels might involve a thalamic gating mechanism using the reticular nucleus of the thalamus controlled by prefrontal sites. Peripheral gating mechanisms still represent a potential source of selection that might be especially important in lower mammals. However, in human information-processing studies, it was clear that unattended information often was processed to a high level, as evidenced by an important message on the unattended channel interfering with the selected channel (Posner, 1978). More recent monkey and human studies suggest that thalamic mechanisms might work in conjunction with sensory-processing areas to gate information at cortical levels (LaBerge, 1995a). The ability to select input channels and the levels at which selection occurs has remained an active focus for current studies of attention.

One effort to integrate cognitive ideas of attention with those arising in neuroscience used the chronometric techniques with normal participants and patients who had lesions in those areas in which attentional effects had been found in alert monkeys (Posner, Cohen, & Rafal, 1982). To study how normal and brain-damaged participants orient, brief cues were used to provide information about the location of a subsequent target. In normal participants, it was found that a cue summoned attention very rapidly to its location and improved RT to targets that occurred there. If the
target occurred at an uncued location, RTs were slowed relative to the cued location.

Once attention moved from the cued location to a new location, recollection to the previously cued location was inhibited with respect to other locations in the visual field (inhibition of return).

These results seemed to make ecological sense, in that attention could be seen as moving toward a location of interest in advance of the eyes and was usually followed by foveation of the selected location. Inhibition of return appeared to be a mechanism to favor searching in novel locations.

In the 1970s, psychologists began to distinguish between automatic and controlled processes (Posner, 1978). Visual words could activate other words similar in meaning (semantic associates) even when a person had no awareness of the words' presence (Marcel, 1983). However, selecting a word meaning for active attention appeared to suppress the availability of other word meanings. Attention was viewed less as an early sensory bottleneck and more as a system for providing priority for motor acts, consciousness, and memory (Allport, 1980). These higher level mechanisms of attention were assumed to involve frontal brain areas and to provide an important means of coordinating cognitive processes. The semantic priming studies indicated that the parallel organization found for sensory information extended to semantic processing; further, both automatic and voluntary processes were involved in these higher level cognitive abilities (Neely, 1977).

In line with this work, it was determined that even spatial attention could be differentiated on the basis of automatic or voluntary aspects. For example, if a peripheral cue indicated that the target would occur elsewhere, facilitation was first found at the location of the cue and then at the location indicated by the cue. It appeared that attention could either be summoned automatically by a peripheral cue or be moved voluntarily and that the cognitive processes involved in these operations could be isolated using chronometric methods (Posner, Inhoff, Friedrich, & Cohen 1987).

To determine if the cognitive operations were related to the same brain areas as those that were modulated by attention, researchers tested patients with lesions of the parietal lobe, thalamus, and superior colliculus on this same orienting task. Each patient group showed a deficit that could loosely be called neglect of the contralesional side, but the details demonstrated remarkably different problems suggestive of the operations that involved orienting. For example, lesions in the parietal lobe primarily produced a deficit when attention was already engaged at a different location (dissociation). Lesions of the colliculus influenced the speed of movement (move), whereas thalamic lesions appeared to make it difficult to select a target among competitors even when attention was correctly oriented (engage). Mental operations appeared to be localized in specific brain areas. However, only the presence of an interconnected vertical network of areas (or neural circuit) could produce successful orienting. These results suggested a localization of operations, one in which several brain areas all worked in concert to perform a single cognitive task. This specific model is still in doubt, and there are many ideas about how spatial attention works. However, the idea was important in that it suggested both that cognitive computations were anatomically localizable and that most tasks involved a whole set of computations and hence, a number of separate neural areas. Exploring this idea required an ability to examine simultaneously many brain areas in normal participants while they performed cognitive tasks. However, no method present in the late 1970s could accomplish this task.

Neuroimaging

The effort to link cognition to specific brain systems depends on having methods available to secure links between cognitive computations and neural processes (Sejnowski & Churchland, 1989). In 1978, a Scientific American article (Lassen, Ingvar, & Skinhøj, 1978) reviewed the work of Scandinavian researchers exploring the use of changes in cerebral blood flow during reading. Reading a story is a complex task and involves many mental operations; during this task, much of the brain was active, but the potential for further analysis was clearly present.

Cognitive Anatomy

Since the 1970s, under the general title of neuroimaging, many methods for exploring the parts of the brain active during cognitive tasks have become available (Posner & Raichle, 1994; Roland, 1993; Toga & Mazziotta, 1996). The parameters for classifying methods are based on the spatial resolution and temporal precision of the techniques. PET and fMRI are two methods with good spatial resolution that are used to measure brain activity in humans during cognitive tasks. To move beyond the localization of anatomical areas in humans, it is useful to use methods sensitive to the time dynamics of information processing. Cognitive studies have shown that to understand cognitive operations requires analysis in the millisecond range. The temporal precision of various methods is changing rapidly (see, e.g., Dale & Buckner, 1997), but imaging methods based on blood flow require changes in blood vessels that limit their temporal precision to the constraints of the hemodynamic lag (1–4 s). Event-related brain potentials (ERPs) and magnetoencephalograms (MEG) are two alternative methods that have precise temporal resolution. Currently, studies that combine localization methods (e.g., PET) and time-dynamic methods (e.g., ERPs) provide a convenient way to trace the rapid time-dynamic changes that occur in the course of human information processing (see, e.g., Heinze et al., 1994).

It would be ideal if noninvasive imaging methods were developed that provided the desired combination of high temporal and spatial resolution. However, all current technologies have their own limitations. For example, cellular recording is limited to animals and is associated with sampling problems. MEG is expensive and, unless many channels are used, must be repositioned for examination of each area. ERPs suffer from difficulty in localizing the neural generators of the electrical signal directly from scalp distributions; subdermal electrode recordings solve this problem, but opportunities for such recordings are uncommon. fMRI and PET are limited by the time for blood vessels to reflect brain activity (hemodynamic lag). Event-related fMRI uses averaging techniques similar to cognitive or ERP studies to improve on the time course resolution and rectifies the problem of having to block trial types (see, e.g., D'Esposito, Zatorre, & Aguirre, 1999; Josephs, Turner, & Friston, 1997). Each of the extant methods will lead to modifications that may address some or all of these problems, and new methods are becoming available all the time (see, e.g., Gratton & Fabiani, 1998).
**Experimental design.** A frequent design for brain imaging studies involves a comparison of an experimental condition with a control condition. In PET, for example, one often subtracts the data obtained in the control condition from the experimental condition, with the goal of determining what operations are uniquely involved in the experimental condition. Consider an effort to see what parts of the brain are involved in seeing motion. If the experimental condition is a set of moving dots and the control condition is the same dots without motion, a subtraction may be thought to involve motion-sensitive areas (Zeki, 1993). This type of design has been quite successful in localizing areas of the visual system that are related to processing motion, color, or form. A quite different design is to use the same stimuli but require the participant to attend to different aspects of the stimuli. Attention to motion in one condition versus attending to color or form in another condition has been shown to activate pretty much the same areas as when subtracting activation from moving versus non-moving stimuli (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990).

Things become more complicated when higher level cognitive processes are involved. It is certainly not possible to say that most experimental and control tasks differ in only one way. For example, consider a task in which one wishes to understand how a word association is generated; the experimental condition may involve generating the use of a noun (e.g., see hammer and say "pound"), whereas the control condition may involve merely reading the word aloud. The resulting subtraction might rid one of the sensory and motor components but leave intact a number of operations related to differences in attention and effort required between the experimental and control task, as well as the operation of interest (e.g., semantic processing). This issue has sometimes been called the problem of pure insertion, and it is generally agreed that it is not possible to guarantee successful control of all but one operation.

Fortunately, the isolation of cognitive operations can be achieved in different ways for imaging studies than for purely behavioral methods. Table 2 indicates the use of a hierarchical design based on cognitive studies that provide information about the constituent processes of word reading. By the subtraction of the control condition from the experimental conditions and the reverse subtraction, it is possible to see not only the increases in blood flow that take place because of generating a word but also which brain areas decreased in activity between conditions. It is also possible to test if the overall subtraction of the fixation condition from the generate condition yields the appropriate areas in each of the component subtractions. These are powerful checks on the subtractive method.

When participants retrieve a semantic associate (verb) during the generate-use task, a network of cortical areas is active relative to reading aloud. These areas include the anterior cingulate gyrus (thought to be related to the increased attention required by the generation task), left inferior frontal areas and left parietal-temporal areas (both thought to be related to aspects of getting the associated use), and the right lateral cerebellum. However, another frontal structure, the anterior insula, is actually more strongly activated during reading aloud than during generating uses. One explanation of this later result is that a different network of cortical areas is involved in the effortful task of generating the use of a word from the automated task of reading a word aloud. Indeed, there is evidence that when the same list is generated several times, the network changes from the one involved in generation to the one used for reading (Abdulhay & Posner, 1997; Raichle et al., 1994).

The subtractive method is not the only strategy for analyzing functional imaging data. Other methods have attempted to demonstrate activation of an area whenever the proposed operation of that area is required, regardless of the task (see, e.g., Wojcikull & Kanwisher, 1999), or to use parametric increases of the operation that is proposed for a particular neural structure. For example, studies of working memory have manipulated the number of items held in mind. As the number of items increases, there are systematic increases in the strength of activation of brain areas related to storage load (Jonides et al., 1997).

Because there is often interaction between cognitive variables and their effects on particular brain areas, factorial designs are sometimes used. For example, Dehaene (1996) proposed an additive model for the task of deciding if a single input digit is above or below the number 5. In his behavioral study, he obtained additivity between the proposed cognitive operations involved in this task (Arabic versus written digits, distance [i.e., close or far from 5], and output). These cognitive interactions provided the basis for findings from ERP and PET studies that demonstrated that each cognitive operation corresponded to a particular neural area active at a particular point in time. More methods and further advances are being developed as neuroimaging becomes a more established and employed science. Many of these methods are reviewed elsewhere (Friston, 1997; Toga & Mazziotta, 1996).

**Localization.** Early cognitive studies with PET involved a blur circle of activation encompassing an area of about 18 mm. However, even with these devices, it was possible to obtain differences between two locations activated separately when 3 mm or less separated them (Fox, Mintun, Raichle, & Miezin, 1986).

The development of new PET scanners and the use of fMRI have improved the ability to differentiate areas of activation in the range of about 1 mm. Cognitive data from PET often use a standardized brain space (Talairach & Tournoix, 1988) and relate activity in each individual brain to this common space. Coregistration of fMRI and structural MRI scans has provided more precise localization of activity with respect to the rather different anatomy of individual human brains. Because fMRI does not rely on radioactive tracers, it is possible to collect a large amount of imaging data from one participant; this should allow studies of individual variation.

---

**Table 2**

**Paradigm Design**

<table>
<thead>
<tr>
<th>Control state</th>
<th>Stimulated state</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixation point only</td>
<td>Passive words</td>
<td>Passive word processing</td>
</tr>
<tr>
<td>Passive words</td>
<td>Repeat words</td>
<td>Articulatory coding, motor programming</td>
</tr>
<tr>
<td>Repeat words</td>
<td>Generate uses</td>
<td>Semantic association, selection for action</td>
</tr>
</tbody>
</table>

*Note.* An example of a hierarchy of tasks to study single word processing. Each condition serves as an experimental or control state (after Petersen, Fox, Posner, Mintun, & Raichle, 1989).
Many of the PET studies of cognition have involved averaging data over a sustained period of 60 s, and most fMRI experiments have used at least a 10-s averaging window. Recently, evoked blood flow responses have been obtained by averaging trials at intervals following a stimulus (event-related averaging) in a way analogous to the time-locked averaging of scalp electrical data (Buckner, 1998; D’Esposito et al., 1999).

**Plasticity**

In neuroscience, the issue of plasticity in brain activity has been discussed mainly at the synaptic level. For example, correlated neural firing among neurons in contact with each other leads to a change in the probability of one neuron being able to induce firing in the other. This principle of learning, first discussed by Hebb (1949), has been shown to be a basic principle for synaptic plasticity.

The use of neuroimaging methods, however, has provided a different level of analysis of plasticity. Instead of individual synapses, the focus is on the question of how experience influences both the set of neural areas active within a task and the time course of their activations. This work has begun to allow scientists to consider possible neural mechanisms for many of the kinds of changes involved in learning and education. Table 3 indicates some of the ways in which experience or strategy might work to change brain circuitry on a temporary or more permanent basis. Although the time courses and mechanisms listed in Table 3 no doubt overlap, we hope they provide some idea of the wide range of important applications of this field to the study of learning and memory.

The top row of Table 3 refers to the finding that attention allows rapid changes in neural activity in local brain areas. As we have outlined above, neuroimaging methods are sensitive to changes in blood flow that accompany neural activity. When a brain area is being used to perform computations in high-level skills, it increases in activity (Corbetta, Miezin, Dobmeyer, Shulman & Petersen, 1991). Attention influences the priority of a given operation by amplifying the amount of neural activity within the area performing the computation (Posner & Pavese, 1998). Often, this is done voluntarily, as one tries to select a set of operations that seem most appropriate for a given task. This is what we discuss later as effortful control by attention.

Priming (row 2 of Table 3) can be produced by the mere presentation of a sensory event. Neuroimaging and cellular studies suggest that the number of neurons activated by a primed target is reduced relative to nonprimed target processing. The prime apparently tunes the neurons involved in the target event so that only those most appropriate to processing the subsequent target are activated (Ungerleider, Courtney, & Haxby, 1998).

Priming is the best understood of the forms of plasticity described in Table 3. After a stimulus is displayed, performance for a subsequent target identical to the prime can be enhanced by two different mechanisms: attentional modulation and automatic pathway activation (Posner, 1978). Although both attention and repetition improve performance, they do so by two contrasting effects in neuronal activity. Attentional modulation increases neuronal activation and increases blood flow to that region (Corbetta et al., 1991), whereas the automatic pathway activation following a repetition decreases neural firing and decreases blood flow to that region (Wiggs & Martin, 1998).

This reduction in neuronal firing may help to explain why the second presentation of an item leads to improved RT but impoverished recollection (e.g., in massed practice effects in memory studies). It is well known in the cognitive literature that a to-be-remembered stimulus benefits by repeated presentations if intervening stimuli occur between the repetitions, whereas a stimulus whose repeated presentations are spaced closed together is less likely to be remembered (the “spacing effect”; see, e.g., Hintzman, 1974). Evidence has accrued that suggests that attention-driven processing plays a key role in explicit memory, whereas automatic pathway activation may play a key role in implicit memory (see Cowan, 1995). These two processes usually act harmoniously to speed the identification and response selection for a stimulus. However, the fast processing and subsequent decreases in neuronal activation for stimulus-driven activation may play a detrimental role in explicit memory. Wagner, Maril, Dale, Rosen, and Schachter (1999) have recently demonstrated that stimuli whose repetitions appear further apart are likely to show increased activation in dorsolateral frontal areas of the brain. Frontal activity was positively correlated with the likelihood of the stimuli later being remembered. In contrast, stimuli whose repetitions occurred close together had reduced blood flow in object-processing areas of the brain. Most importantly, reduction in blood flow in posterior regions of the brain was negatively correlated with memory for that stimulus; that is, if the posterior regions showed reduced blood flow because of repetition, participants were less likely to remember that stimulus on a subsequent recollection test. Hence, memory encoding for repeated stimuli, the attentional-driven and stimulus-driven processes may work in opposition.

Practice on a set of already learned but not recently rehearsed associations (row 3 of Table 3) shows that automaticity can completely change the neural pathways used to accomplish a task. As discussed previously, Raichle et al. (1994) required people to generate a use for a read or heard noun (e.g., pound as a use for a hammer). When a new list of words was presented, there was activity in the left frontal and posterior cortex, the anterior cingulate, and the right cerebellum. Activity in the anterior insula was reduced over what was found in simply reading the words aloud. A few minutes of practice at generating an associated use shifted activation so that the left frontal and posterior areas important in generating a new use dropped away and the anterior insula, strongly activated during reading aloud, increased. When generating a given word became automated with practice, the same circuit

---

Table 3

<table>
<thead>
<tr>
<th>Time</th>
<th>Name</th>
<th>Mechanisms</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seconds</td>
<td>Priming</td>
<td>Automation</td>
<td>Raichle et al. (1994)</td>
</tr>
<tr>
<td>Minutes</td>
<td>Practice</td>
<td>Connections</td>
<td>Karni et al. (1998)</td>
</tr>
<tr>
<td>Months to years</td>
<td>Development</td>
<td>New networks</td>
<td>Posner &amp; Rothbart (1995)</td>
</tr>
</tbody>
</table>

**Note.** Mechanisms of brain changes (plasticity) obtained at different time scales and through different mechanisms.
was used as when skilled readers simply spoke words aloud. There appeared to be one neural circuit associated with the thought needed to generate a familiar but unpracticed use of a noun and a different neural circuit when the task was automated, as in reading aloud or generating a just-practiced association. The circuit used for thought includes attentional mechanisms involving effortful control, whereas an automated circuit does not involve attention.

In the study cited above (Raichle et al., 1994), people were dealing with already well-known associations, as, for example, the association between hammer and pound. Even when they have not been practiced recently, connections between hammer and pound are available because of their semantic connections from experience. However, it is often necessary to acquire entirely new associations, as in learning the words of a foreign language. This involves establishing new connections in the brain (row 4 of Table 3) and may require many weeks of practice. In one study of learning a new artificial language, it took 20–50 hours of practice on 40 lexical items before the learned words showed the same superiority in RT usually found for reading the native language (McCandliss, Posner, & Givon, 1997).

Even more complex than learning a few new associations is developing a whole system to carry out an important linguistic function (row 5 of Table 3). Studies using PET with literate adults have shown that areas of the visual system of the brain become active when strings of letters are possible words in English, whether they have meaning or not (Petersen, Fox, Snyder, & Raichle, 1990). This area of the brain is not active for nonsense strings like a series of consonants. It seems to represent English orthography and has been called the visual word form system (Petersen et al., 1990). This system appears to be a left posterior function that serves to group letters of a word automatically into a single chunk. This system appears to require some years to develop. Evidence suggests it is not present in 7-year-olds, even in those who know how to read, and can be found in 10-year-olds only to a limited degree. Moreover, once this system is developed, it appears to be strongly resistant to change (Posner & McCandliss, 1999).

The final row of Table 3 refers to changes in brain structures that develop over the early life span of the individual. We have in mind the several years apparently required before attentional networks are developed. We discuss this material in more detail in the section on Development.

Pathology

Knowledge of the normal networks of attention provides a basis for consideration of pathologies. Attentional deficits following either lesion or organic brain dysfunction (see, e.g., Knauper, 1921) have long been an area of study in psychology and neurology. As researchers have come to learn more about the cognitive processes and neural areas involved in attention, they have gained important insights into how dysfunction of attention manifest and, fortunately, how they may be rehabilitated (see, e.g., Robertson, Mattingley, Rorden, & Driver, 1998; Rossetti et al., 1998).

Neuropsychology. Because more is known about the visual orienting system than any other attentional network, we first consider the neurological syndrome known as hemineglect. Unilateral strokes often produce a strong tendency to ignore information contralateral to the side of the lesion. The symptoms can involve very bizarre behavior; for example, a patient may deny that the left arm belongs to him or her, may only eat off one side of his or her plate, may comb only a part of his or her hair, may maintain he or she is not ill, and may exhibit inappropriately positive affect. Experimental studies have shown that at the core of this disorder is a deficit that includes a bias to events on the ipsilesional side of the lesion and a difficulty in disengaging from these stimuli to orient toward the contralesional direction. When the lesion is in the parietal lobe, this deficit appears to be permanent. Moreover, it does not appear to depend on the basis of the lesion. This deficit was first demonstrated in stroke patients, but it has now been shown following closed head injury and as a result of metabolic changes in the parietal lobe that occur with the onset of Alzheimer's disease (Parasuraman & Greenwood, 1998).

Cognitive studies have shown how this disorder interacts with tasks such as visual search and word reading to produce quite specific deficits in processing letter strings and looking for targets. It is possible to mimic these deficits in normal participants when orienting is manipulated by cuing participants to one side of the stimulus and presenting very brief flashes that do not allow time for disengagement from the cue. These results are important because they show that manipulation of the hypothesized component operations of attention mimics the deficit. This is an important strategy for linking normal and brain-damaged performance. Although an important core of the neglect syndrome appears to be in disengaging attention, many of the initial symptoms cannot be explained from this perspective, but they may involve the widespread disorganization that follows the initial damage (Rafal, 1998).

Recently, several studies have tried to build on the increased knowledge of neglect to ameliorate the symptoms associated with this damage (see, e.g., Ladavas, Menghini, & Umlit, 1994; Robertson et al., 1998). When damage is in the parietal lobe, it is possible to build some compensation through the use of symbolic cues that remind participants to orient to the neglected side. When the damage is more anterior, it is possible to use peripheral cues to draw attention to the side opposite the lesion. Alternatively, one can use other, intact attentional mechanisms to compensate for the damaged operation. For example, if the alertness of a neglect patient is increased by means of a warning sound prior to the appearance of targets in both visual fields, the patient is less likely to neglect targets in the bad visual field. This result demonstrates the ability of an intact attentional network (alerting) to compensate for a damaged network (visual orienting) even though these two processes do not interact in participants with intact brains (Fernandez-Duque & Posner, 1997).

Though neglect is a neurological disorder in which the lesion can be readily identified from MRI or CAT scans, its symptoms can be bizarre and seemingly arbitrary. In this way, it resembles many pathologies for which the brain systems are as yet unknown. It appears that it is possible to apply the lessons learned from the study of neglect to psychiatric disorders whose pathologies are even more diverse or whose etiologies are unknown.

Psychiatric. It is common to view schizophrenia as a higher order disorder involving attention. However, recent studies have begun to use the new imaging methods and ties between cognitive and anatomical concepts to move beyond this vague idea and provide detailed methods for studying the pathophysiology of the disorder (Frith, 1992; Frith & Dolan, 1998).
Imaging studies have demonstrated the involvement of specific cortical areas and subcortical areas. PET studies of never-medicated schizophrenic patients have shown evidence of a left globus pallidus (basal ganglia) abnormality (Early, Posner, Raichle, 1989a, 1989b). As the disorder progresses, there appears to be a widespread reduction of blood flow and metabolism in the frontal lobes that involves both the anterior cingulate and the dorsolateral prefrontal cortex.

Though the results of these studies have sometimes seemed in conflict, they all suggest dysfunction of a single network involving the attention and the parts of the brain used to represent information in absence (working memory). In cognitive studies, the term working memory usually involves both a storage component and an executive system that manipulates the stored information. It seems possible that the dorsolateral and midfrontal cortices are both part of such a system. Indeed, many of the deficits in cognitive tasks found in schizophrenics fit well with the disorder of an executive attentional function. Further, many of the disorders unique to schizophrenia seem to involve attentional networks including loose associations in language, thought-disorder, and aberrant spatial attention. These deficits are far more specific than might be thought at first. For example, the spatial attention deficit involves a specific difficulty in shifting attention to the right visual field. This deficit appears to occur only for schizophrenic patients displaying positive signs such as auditory hallucination and does not occur in nonschizophrenic monozygotic co-twins or in chronic patients (P. J. Purdon et al., 2000).

On the other hand, the deficit of the lateral prefrontal cortex seems to be more prevalent in chronic patients. These specific deficits cannot be attributed to general factors such as motivation or the ability to understand instructions or to general factors of the disorders. It should now be possible to select patients on the basis of specific abnormalities and symptomologies so that future studies can learn more about the progression of this disorder. Meanwhile, accumulating neuromaging data in schizophrenic patients has shown direct deficits in medial and lateral frontal structure (Fletcher, McKenna, Friston, Frith, & Dolan, 1999; Frith et al., 1995; McGuire et al., 1998) that are commonly active in normal participants during cognitive tasks such as the Stroop effect and the generate-use task. Pathological studies of schizophrenics (Benes, 1993) have shown cellular abnormalities in the anterior cingulate that have led to a conception of how the disorder might influence the ability to inhibit competing ideas. Indeed, in one recent neuroimaging study (Fletcher et al., 1999), direct evidence was found for a failure of the anterior cingulate to modulate temporal lobe areas.

Though there is no guarantee that knowledge of the specific pathophysiology of the attentional aspects of schizophrenia will lead to remediation strategies, at the very least it should allow a better understanding of the various forms of the disorder. One goal of understanding the specific cognitive and neural areas that are dysfunctional in patients with schizophrenia is to better isolate the particular neurochemistry that may play a key role. For example, in a beautiful integration of applied and basic cognitive neuroscience, Dolan and colleagues (1995) demonstrated that patients with schizophrenia showed impaired performance on the generate-use task and impaired activation of the anterior cingulate during this task. Following neuroleptics that modulated dopamine levels, patients with schizophrenia performed at the same level as participants with no history of psychiatric disorder and showed increased activation in the cingulate cortex. We hope that by isolating the cognitive operations and neural areas that subserve these operations, researchers can develop more specific psychotropic drugs to alleviate the symptoms associated with this disorder.

Cognitive Neuroscience

Cognitive neuroscience studies are beginning the task of integrating questions of human cognition from neurons through behavior as outlined in Table 1. This is no easy task, nor has it been completely accomplished in any domain. There are many problems in attempting this integration, ranging from methodological difficulties to theoretical issues arising with changes in levels of analysis. Our goal in this section is to indicate how an integrated view of cognition and brain function, even if necessarily very incomplete, helps to illuminate many topics of importance to psychology.

Functional Anatomy

A major achievement of brain imaging studies has been consistent localization of brain areas that perform particular functions. This achievement has gone a long way toward providing validation both of the decomposition of skills and of the mapping of these skills onto particular brain networks. The evidence is most impressive within the visual systems, where it has been possible to plot retinotopic maps that appear to be homologues of areas V1, V2, V3, V3a, MT/V5, VP, and V4 in macaque monkeys (Tootell et al., 1997). These areas seem to separate the processing of motion, color, shape, and other object features in the human brain much as they do in nonhuman primates.

For higher level cognitive tasks such as reading and listening, remembering, identifying faces, searching visual arrays, creating visual images, processing music, and learning sequences of spatial locations, the almost universal results have been activation of a small number of often widely separated neural areas. During diverse cognitive tasks, neural areas work in concert to form a neural circuit with particular areas active at particular times to perform the necessary cognitive computations in the appropriate order of the task demands (Posner & Raichle, 1994, 1998).

The study of attention has been one such area that has demonstrated networks of brain areas involved in attentional processing. Posner and Petersen (1990) used the neuropsychological and neuroimaging data to outline three networks involved in attention. Imaging data has confirmed the important role that areas within the posterior parietal lobe play in shifts of attention, especially between spatial locations (Corbetta, Miezin, Shulman, & Petersen, 1993). Although these studies have not separated disengaging attention from moving attention, they have suggested that activation in the parietal lobe is related to these two mental operations. Imaging studies (LaBerge, 1990; LaBerge & Buchsbaum, 1990) have also supported the pulvinar and associated thalamic areas as critical for engaging attention to a target in the presence of distractors.

Meta-analysis based on combining data across studies and laboratories has proven an important vehicle for summarizing the almost overwhelming number of brain areas that have been activated in the last 10 years (Fox, Parsons, & Lancaster, 1998).
Despite the many difficulties that are discussed in Fox et al. (1998), the general degree of convergence across laboratories, methods, and modalities has been impressive. It is often possible to identify particular operations with individual areas of activity, but there are frequently disputes about the exact function of a particular brain area.

**Circuitry**

Hemodynamic responses lag behind neurons by a second or more (Raichle, 1998), whereas mental operations often take place in tens of milliseconds (Posner, 1978; Sternberg, 1969). For several decades, event-related electrical potentials recorded from the electroencephalogram in humans have been important in tracing the real-time activity of cognitive processes. However, with the advent of imaging techniques, it is now possible to relate the scalp activity to generators obtained from the blood flow measures in humans and neurophysiological measures in animals. Although the earliest parts of the ERPs are primarily driven by the sensory event, within 100 ms top-down attentional influences are clearly evident in human cortex (Hillyard, Teder-Sälejaervi, & Muente, 1998).

In high-level skills, the time course of cognitive processes is often constrained by the nature of the task. For example, a skilled reader uses only 275 ms for an individual fixation of a word; however, there is clear evidence that the meaning of the word influences the duration of the eyes before the next word is fixated (Rayner & Pollatsek, 1989). Thus, 275 ms sets an upper bound to achieving the word meaning sufficient to convey appropriate information to the saccadic eye movement system. Studies of visual word processing have shown specific areas of the occipital lobe and of the frontal lobe that are activated within the first 200 ms following input (Posner, Abdullaev, McCandliss, & Sereno, 1999). Further, these upper bounds are not constrained to just higher level tasks. The recognition of an object occurs extremely rapidly, and electrophysiological data have set the upper bounds for this process at 155 ms (DiGirolamo & Kanwisher, 1995; Thorpe et al., 1996).

One future goal of cognitive neuroscience is to integrate the wide of cognitive operations into the when of these same operations. Recent studies have begun to incorporate more exact time course measures (e.g., ERPs) with more precise localization methods (e.g., PET and fMRI) to understand the where and when of human cognition (see, e.g., Dale et al., 2000; Heinzle et al., 1994; Martinez et al., 1999). We hope that advances in methods and technology will allow for a combined time course analysis with localization of cognitive function. Priming, attention, and learning are all areas of psychological investigation whose main issues will benefit from knowing the where and when of brain operations. Many issues that are ambiguous when using cognitive methods can be illuminated by a time course and localization strategy (e.g., visual mental imagery and priming).

For example, during much of the research in the field of visual mental imagery, there was a long debate regarding the nature of a visual mental image. An image could be created using a generative language of propositional constructions (e.g., ball—box—on) or it could be created with a true one-to-one depictive representation of a ball on a box (for a review, see Tye, 1991). With the previous methodologies, this debate seemed unresolvable as each result that suggested that people created visual images in their mind’s eye was equally explainable by experimenter bias or participant expectancy. However, the advent of cognitive neuroscience methods, particularly, neuroimaging, helped to revise this debate.

In an important study in the field of visual cognition, it was demonstrated that creating a visual mental image activated the primary visual cortex (Kosslyn, Alpert, Thompson, & Maljkovic, 1993). In simple sensory processing, larger visual stimuli activate more anterior regions of the primary visual cortex, with smaller stimuli activating more posterior regions (Fox et al., 1986). Most persuasive for the notion that mental imagery has a visual component, activation in the visual cortex changes on the basis of the size of the images people create. In line with the perception results, the larger the visual mental image that participants created, the more anterior the activation in the visual cortex (Kosslyn, Thompson, Kim, & Alpert, 1995). This cognitive neuroscience result helped to clarify a long-standing question in philosophy (Plato, 1984) and psychology (Kosslyn, 1994): Mental imagery can and does use visual depiction.

Subsequent experiments (Kosslyn, 1994) demonstrated that both top-down and bottom-up processes are present during normal visual perception: Participants used many of the same imagery mechanisms to assist perception for occluded objects or objects at noncanonical views (e.g., a fence seen from above). This data also emphasizes the importance of time course data as higher level attentional or cognitive processes can drive activation in primary sensory areas.

Below, we concentrate on issues in attention that have been clarified with the advent of cognitive neuroscience methods and theories.

**Applications**

Do these new results modify thinking about traditional issues in cognitive psychology? The study of attention has been a central issue in psychology and continues to be a very active area in neuroimaging and in cellular studies of the primate brain (Motter, 1998). Below, we examine how the study of brain mechanisms has influenced questions of long-standing interest within the field of attention.

**Early and late selection.** One of the oldest issues in the field of attention is how early in processing attention can influence sensory input. Because the issue began during the information-processing period, it often took the form of whether attentional influences were largely changes in criteria in response selection (i.e., biases) introduced by the attentional cue or whether attention influenced the sensory evidence itself. Many empirical studies were done to determine if attentional benefits showed up as alterations in the beta parameter of a signal detection analysis or whether instead they involved changes in the d' parameter (e.g., Hawkins et al., 1990).

Neuroscience evidence from lesion studies, cellular studies, neuroimaging, and event related cortical potentials has been brought to bear on this issue. Lesions of the parietal lobe interfere with the ability to shift attention to the contralesional side and showed a very specific deficit in RT experiments (Posner, Walker, Friedrich, & Rafaile, 1987). Stimuli from the affected side appeared to have great difficulty in summoning orienting under conditions where the patient was already processing something on the ipsi-
lateral side of the lesion (disengage deficit). Although normal participants showed only a small deficit in RT when the target appeared in the opposite visual field from the cue (invalid trials), patients were often simply unaware of the target presented in the neglected field. This result seemed decisive evidence of a sensory influence because the participants with parietal lesions were unable to report the presence of the target.

However, further studies indicated that the stimulus is processed, even if unconsciously, and that awareness is still possible. For example, although patients might miss a contralesional target feature in isolation, they might report the stimulus if integration with material on the ipsilesional side completed an object, even without attention (Driver, Baylis, & Rafal, 1992; Mack & Rock, 1998). Moreover, implicit priming was also reported, indicating that many sensory operations were carried out on information on the contralesional side (McGlinchey-Berroth, Milberg, Vefaellie, Alexander, & Kilduff, 1993). In addition to the lesion evidence, both electrical recordings from the scalp and cellular recordings in monkeys in the object recognition pathway indicated that attention could influence portions of the visual systems in V4 and later.

At first, the cellular data suggested that V1 could not be influenced by attention. However, attentional modulation in V1 appears to occur when the visual field is cluttered (Motter, 1993). This finding shows how closely the findings from psychological method and the physiological results are integrated. Only in a cluttered field are extremely strong attentional effects in RT experiments in normal participants usually seen. Attentional effects in normals are rarely obtained in an uncluttered field, and only neuropsychological patients show a strong effect in an empty field. A cluttered field requires information about both the location of the target and its relation to the surrounding information, and it is then that effects in V1 have been reported (Motter, 1993, 1998; Posner & Gilbert, 1999).

The early/late question can be resolved into three somewhat interdependent issues. First, how early in the nervous system can attention influence stimulus input? The above results suggest that it can be as early as V1 under some conditions (Posner & Gilbert, 1999). Second, how quickly after input can attention influence information processing? The timing issue is of particular importance because activation of a particular brain area may be either along the input pathway or due to feedback from higher areas. It appears that the activation in V1 can be due to feedback from higher order visual systems (Martinez et al., 1999). Third, what does early selection mean for the processing of information (both attended stimuli and unattended stimuli)? Here, the answer is more complex. It seems to mean that certain aspects of complex scenes may be available for conscious report, whereas other aspects are only available if they succeed in producing recategorizing. Unattended objects, however, may still be processed to fairly high levels, and the processing itself may summon attention. The depth of cognitive processing of unattended objects and the possibility of attentional summoning suggest that early selection does not have the cognitive consequence originally implied. Selecting one stimulus over others does not mean that unselected items will not produce a recategorizing of attention (Mc Cormick, 1997) or still influence behavior (Neill, 1977).

The early versus late selection debate in cognitive psychology has been enormously advanced by the addition of information from cognitive neuroscience.

Site and source. An issue basic to understanding attention concerns whether the mechanisms that constitute the brain's attentional system are separate (James, 1890; Posner & Dehaene, 1994) or whether there are no mechanisms of attention different from sensory or response processes (Neisser, 1967).

Like the issue of early or late selection, this debate stemmed from cognitive studies of attention. If there is a general attention system, then a wide range of tasks with apparently different demands should interfere with one another. It was apparent that interference would result between tasks requiring common input or output systems (one cannot look forward and backward at the same time, or both flex and extend a muscle). However, some evidence suggested that interference occurred in many cognitive studies even when the modality of input and the motor output were quite separate. There have been many cognitive studies related to these issues. However, despite evidence on both sides (Bourke, Duncan, & Nimmo-Smith, 1996; Duncan, Humphreys, & Ward, 1997), no clear resolution has been possible.

From an anatomical perspective, one can distinguish two types of brain areas related to attention. By the source of attention, we mean those anatomical areas that seem to be specific to attention rather than being primarily involved in either sensory or response processing. However, when attention operates during task performance, it operates at the neural areas (site of attention) where the computations involved in the task are usually performed. Thus, when observers attend to the color, form, or motion of a visual object, they amplify blood flow in various extrastriate areas (Corbetta et al., 1990). These areas that are significantly more active when observers attend to a particular feature of an object are the same areas involved in the passive sensory processing of these same features. Most brain areas, especially cortical areas, show attention effects in this sense, although they are not part of the brain's attentional system.

Most discussions of the site versus source issue have occurred in the study of orienting (Cohen, Romero, Servan-Schreiber, & Farah, 1994; Luck & Ford, 1998; Posner, 1992). Evidence from neglect patients cited in the last section suggests that orienting in the visual system is achieved by orchestrating changes in the orienting network (involving the posterior parietal lobe, pulvinar, and colliculus) to influence input into the ventral object recognition pathway. Thus, visual search would involve changes in the parietal lobe that reflect disengagement, activation in the superior colliculus involved in moving attention, and the provision of a signal into V4 via the pulvinar that enhanced information at attended locations (LaBerge, 1995b). In favor of this view, PET studies showing strong activity in parietal, pulvinar, and frontal areas during attention shifts (Corbetta, Shulman, Miezin, & Petersen, 1995; LaBerge, 1990) and ERP data indicate early attentional enhancements similar to those found in the cuing studies cited above (Luck & Hillyard, 1994).

However, a different view suggests that attentional effects are primarily due to competition within the object recognition pathway (largely V4 itself (Duncan et al., 1997). This difference may be more one of emphasis than discord as Desimone and Duncan (1995) allowed for top-down influences to be involved in the decision about which part of the scene should be selected. In recent fMRI studies, they suggested that this top-down effect may be mediated by the same set of neural areas that we have been
Neuroimaging studies of the Stroop effects have demonstrated activation of a midline attentional system in resolving the word and color conflict (Bench et al., 1993; Bush et al., 1998; Carter, Mintun, & Cohen, 1995; Carter, Mintun, Nicholas, & Cohen, 1997; Derbyshire, Vogt, & Jones, 1998; George et al., 1994; J. V. Pardo, Pardo, Janer, & Raichle, 1990). Though the correspondence is not exact, all studies have found a midline cingulate activation when the word and color disagree (for a review, see Posner & DiGirolamo, 1998). Surprisingly, the congruent condition (relative to a neutral) sometimes also activated the cingulate cortex (Bench et al., 1993; Carter et al., 1995). Posner and DiGirolamo (1998) have suggested an analysis of the Stroop effect that postulates comparable selective attentional processes to resolve conflict in both the congruent and incongruent conditions (see also Carter et al., 1995). In support of this hypothesis, ERP waveforms for the incongruent and congruent conditions diverged from the neutral condition at about 250 ms but did not diverge from each other until near response (Posner & Rothbart, 1998). Moreover, dipole modeling of the congruent and incongruent conditions produced matching neural solutions consistent with the cingulate gyrus generator seen in the blood flow studies. These results suggest the activity of an executive attentional system during both congruent and incongruent conditions. Color word trials require a selection between information from the word and information from the color, independent of the congruity of the trial type. These data demonstrate that selection of one dimension of a stimulus that has multiple dimensions presents a situation of conflict that requires executive control mechanisms. The timing of the cingulate activation suggests that it is involved in the resolution between the anatomically separate systems involved in processing the ink color and color word name.

Neuroimaging studies have suggested that the cingulate cortex is also involved in attentional processing during dual-task situations. When participants are learning a sequencing task or performing the verb generate task, the cingulate cortex and dorsolateral frontal areas are highly active. In contrast, once performance asymptotes and attention is no longer required, the cingulate and dorsolateral frontal areas are no longer active (Jueptner et al., 1997; Raichle et al., 1994). Dual-task interference also decreases when either of the tasks becomes well learned. During dual-task performance of two tasks that share neither input nor response, performance deteriorates compared with either task alone, and there is increased activation in the cingulate cortex. Passingham and his colleagues (for a review, see Passingham, 1996) have suggested that interference occurs centrally, not in either perceptual or response systems; further, they suggested that one of the neural areas in which the central bottleneck occurs is the cingulate cortex. In support of this notion, when participants are asked to attend to a well-learned sequence, increased activation is found only in the cingulate and dorsolateral frontal areas, and dual-task interference increases. These areas are the same areas active during the original learning of the movement; these areas come back on-line when participants attend to the movements following practice (Jueptner et al., 1997).

To investigate the central executive of Baddeley's (1992) working memory model, D'Esposito and his colleagues (1995) also compared performance in single- versus dual-task conditions. As in the previous study, activation in the cingulate cortex and dorsolateral frontal regions significantly increased (relative to single-
task activity) during the concurrent task condition. D’Esposito and his colleagues argued that the cingulate cortex and dorsolateral prefrontal cortex compose part of the neuroanatomical circuit of the central executive.

Immediately following an error, people are often conscious of making a mistake. When participants were aware of making an error in speeded tasks, they showed negativity over midfrontal channels on the scalp following the key press (Dehaene, Posner, & Tucker, 1994; Gehring, Gross, Coles, Meyer, & Donchin, 1993). Using the Brain Electrical Source Analysis algorithm, Scherg (1989) found that localization of this error-related negativity suggested activation from the cingulate (see also Badgaiyan & Posner, 1998). Errors can be either slips (incorrect executions of a motor program) or mistakes (selection of an inappropriate intention). Dehaene et al. (1994) have shown that the negativity generated by the cingulate only followed a slip when the participants knew they had made an error, but not if they mistakenly selected an incorrect response. Activation of the cingulate following an error has now been shown in blood flow studies as well (Carter et al., 1998). Monitoring of the selected response and awareness of the response actually made is crucial for coherent and successful behavior. Cingulate activation is directly related to awareness of one’s own planned behavior.

The presence of frontal activity, particularly in midline structures in these conflict tasks, suggests that there may prove to be a common anatomy to at least some forms of executive control. The limbic nature of the cingulate cortex and evidence that activity is related to forms of emotional processing (Drevets & Raichle, 1998; Whalen et al., 1998) suggest a possibly important link between attention in cognitive and emotional tasks (Bush, Luu, & Posner, 2000). In this section, we have suggested that these frontal structures are related to subjective feelings of voluntary control of one’s thoughts and feelings. In the next section, we turn to evidence from development that may help to relate the cognitive and emotional influence of these systems.

Development

Infants and Children

The human infant has the longest period of dependence on caregivers of any mammal. Though much of the nervous system develops prior to birth, a long period of postnatal development is needed for primates and particularly humans to complete the basic formation of the brain. During this period, infants gain control of their behavior and mental state so that, as adults, they can exercise a degree of central control over their emotions, thoughts, and actions.

The past quarter century has produced great improvements in the ability to study the behavior and cognition of even the youngest infants. However, until recently, most methods for the study of brain development have involved either the study of nonhuman organisms or the analysis of brain anatomy following death. Neither of these methods has been satisfactory for understanding higher level cognitive and emotional controls. Recently, approaches that were previously applicable only to adults have begun to be available for the study of human brain–behavior relationships during development. For example, MRI can be used to examine the size and shape of particular brain areas as a function of age (Casey, Trainor, Giedd, et al., 1997), and functional MRI can be used with older children (Casey, Trainor, Orendi, et al., 1997).

It appears that in the first year of life, maturation of the parietal lobe is accompanied by the ability of infants to disengage and move attention voluntarily from location to location (Ruff & Ruff, 1996). These achievements in orienting lead to the ability to learn to place attention and the eyes at particular locations. Various forms of novelty preference also develop in this period.

However, the higher forms of attention, such as the ability to resolve conflict between simultaneous stimulus events, seem to undergo strong development in the third year of life and beyond (Posner & Rothbart, 1998). There is evidence of development in the ability to resolve stimulus conflict in Stroop-like tasks between 2 and 4 years (Posner & Rothbart, 1998). There is also some direct evidence that the anterior cingulate is developing during later childhood in a way that might support the data on conflict resolution. In children ages 5.3 to 16 years, there is a significant correlation between the volume of the area of the right anterior cingulate and the ability to perform tasks relying on focal attentional control (Casey, Trainor, Giedd, et al., 1997). Moreover, in a recent fMRI study, performance of children ages 7 to 12 years and adults were studied when performing a go/no-go task in which they were required to withhold pressing to an X while responding to non-Xs. This condition was compared with control tasks where participants responded to all stimuli and thus never had to withhold a response. Both children and adults showed strong activity in the prefrontal cortex when required to withhold responses. Moreover, the number of false alarms made in the task was significantly correlated with the extent of cingulate activity (Casey, Trainor, Orendi, et al., 1997).

Aging

In a recent critical review of the literature on the neuroanatomy, neurophysiology, and neuropsychology of aging, West (1996) concluded that older adults are disproportionately disadvantaged on tasks that rely heavily on executive control processes that are supported, in large part, by the frontal and prefrontal lobes of the brain. Indeed, there is a good deal of evidence to suggest that morphological and functional changes in brain activity do not occur uniformly during the processes of normal aging. Researchers have reported substantially larger reductions of gray matter volume in the dorsolateral and medial frontal regions than in sensory cortical regions (see, e.g., Coffey et al., 1992). Studies of functional brain metabolism have reported similar trends, with frontal regions showing substantially larger decreases in metabolic activity than do sensory areas of cortex (see Madden & Hoffman, 1997, for a review).

These data are consistent with numerous reports of large and robust age-related deficits in the performance of tasks that are supported by the frontal and prefrontal regions of the cortex as compared with relatively small age-related deficits on non-frontal-lobe tasks. Indeed, any of the tasks subserved; in large part, by the frontal lobes involve processes associated with executive control functions such as the selection, control, and coordination of computational processes that are responsible for integrating perception and action. For example, large age-related deficits have generally
been reported when adults are required to rapidly shift between tasks (Kramer, Hahn, & Gopher, 1999). The ability to switch task sets is considered one of the primary examples of executive control, because it requires resolution of conflict between the two task sets and the internal goal and response sets necessary to switch between them (see Allport, Stiles, & Hsiieh, 1994; Norman & Shallice, 1986; Rogers & Monsell, 1995). Functional imaging studies have shown enhanced activation of regions of the prefrontal and frontal cortex when two tasks are performed together but not when they are performed separately (D’Esposito et al., 1995; Passingham, 1996). Lesion data have also implicated the frontal lobes in this task as damage to the frontal lobes impairs the ability to switch between tasks and often leads to perseveration of action (Duncan, 1995; Duncan, Emalfie, & Williams, 1996).

The Future

We expect both the methods of and the results from neuroimaging to influence many areas of psychology. The methodology should continue to improve in accuracy and ease of use in the future. New techniques such as optical imaging, as well as new methods of using magnetic resonance, should open up opportunities to explore issues of brain development and changes with learning.

The results to date provide much promise of revealing the neural basis of psychology that have not been generated by neurodes and researchers. In the course of this review, we have had occasion to refer to studies that involved cognitive, emotional, clinical, and developmental issues. As Miller (Miller & Lennenberg, 1978) pointed out, localization is not just about determining where things occur; it is also a method for determining if the brain respects a proposed decomposition or analysis of mental operations. Because all of psychology seeks analytic strategies to understand the tasks and persons being studied, diverse areas of study from basic, applied, and developmental research can all benefit from this general approach.

In this article, we have sought to show how localization and its implied decomposition can illuminate some of the issues that have been with psychology since the inception of the field. We have concentrated on topics of attention both because it is familiar to us, and because, we believe, attention and volition are at the very heart of the psychological enterprise. Further, many imaging studies have been conducted in the hopes of understanding issues in attention. However, we expect the general approach outlined in this article to be useful for many other questions in psychology.

One area that remains to be characterized is the reason that certain computations occur in particular brain areas. There has been spectacular progress in efforts to relate functional activation of visual areas in humans to the visual maps obtained from cellular recording in primates (Tootell et al., 1997). This information could illuminate the understanding of why particular areas are used to perform particular functions. An important example is the St. Louis PET group’s somewhat controversial finding of a brain area that chunked visually presented letters into units that the researchers termed the visual word form area (Petersen, Fox, Snyder, & Raichle, 1990). Clearly, any such brain system must be organized by the specific learning experiences that children have during their early education. Indeed, these areas develop in childhood by responding to specific words the child has learned, and only after training do they become related more to the orthography of the language than to the familiarity of the learned exemplar (Posner & McCandliss, 1999). This area in the left ventral occipital lobe appears to be in or near areas active during the processing of color. It is known that the color areas are late visual adaptations and that they have input from the high spatial frequency parvocellular systems of the geniculate nucleus and primary cortex. If one were to follow the exact computations performed by homologous areas in higher nonhuman primates and observe the role of spatial frequency and learning (Ivry & Robertson, 1998), one might have an empirical approach to understanding the affinity of this area for the computations needed to organize visual letters.

A more distant goal for the study of imaging, one that requires careful thought, is the development of general principles of how closely related computations are expressed both in brain tissue and in performance. Kinsbourne and Hicks (1978) proposed that the more densely connected two brain areas were, the more interference they would show when simultaneously active and the more priming possible between the two local areas. At the time this theory was proposed, scientists were not able to test these ideas except by experiments using different motor outputs. Now, if one assumes that being closer in brain space relates to connectivity, there are many opportunities to test links between performance and neural distance.

In neuroscience, a very fundamental unit is the cortical column, which is thought to form the basic computational unit of the brain (Mountcastle, 1976). No such principle is yet available for saying how similar cognitive operations should be mapped within the brain. The mechanisms that allow for covert shifts in orientation of attention (spotlight) are based in the parietal lobe, but how do these relate to selection of the scale of attention (zoom function)? One might expect two such related computations to be localized close together. Although mechanisms within a single cognitive domain may sometimes be close together (e.g., visual color and form), it often appears that both frontal and posterior cortical areas are involved in the same domain (e.g., language and working memory). It is possible that no general large-scale principles of brain organization will emerge. However, we must continue to probe the masses of data for common theoretical organization in diverse fields of study (e.g., language, attention, etc.).

The progress made in mapping cognitive functions in the human brain has been swift and startling. Though researchers are still a long way from exploiting these methods to derive general principles of the neural function of thought and behavior, they now have the tools to pursue the dream of Freud, Wundt, and generations of other psychologists. The future of cognitive neuroscience provides a means of understanding human cognition and emotion in terms of the neural basis of mental events.

References


Duncan, J., Brulsit, H., & Williams, P. (1996). Intelligence and the frontal


Received June 11, 1999
Revision received November 10, 1999
Accepted November 16, 1999