Integrating electrophysiology and neuroimaging in the study of human cognition

G. R. MANGUN and J. B. HOPFINGER
University of California, Davis, California

and

H.-J. HEINZE
Otto von Guericke University, Magdeburg, Germany

Noninvasive recordings of electrical and magnetic fields generated by neuronal activity have helped to characterize the temporal sequencing and mechanisms underlying human cognition. Progress is being made toward the goal of localizing the intracranial loci at which many important electromagnetic signals are generated through the use of new analytic techniques and of scalp recordings of electromagnetic activity in neurological patients and through related work in animals. Such methods alone, however, do not yet have the three-dimensional spatial resolution that is necessary in order to identify the intracranial anatomical structures that are involved in the generation of externally recorded activity and, thus, cannot yet inform us with precision about the anatomic substrates of neural events. In comparison, neuroimaging methods, such as positron emission tomography and functional magnetic resonance imaging, can provide higher spatial resolution information about which brain structures are involved in perceptual, motor, and cognitive processes. However, these imaging methods do not yield much information about the time course of brain activity. One promising approach is to combine electromagnetic recordings and functional neuroimaging in order to gain knowledge about the spatiotemporal organization of human cognition. Here we review how electrophysiology and functional neuroimaging can be combined in the study of attention in normal humans.

Neural computations operate on a millisecond to millisecond time scale within the brain, and the underlying neural systems supporting these computations have a complex anatomical organization. In order to arrive at a complete description and understanding of brain function, both the time course and the anatomy of sensory, cognitive, and motor processing must be understood. Numerous techniques exist to measure brain functioning in humans, but no one approach provides both the temporal and anatomical precision necessary to elucidate the mechanisms of human thought and action. However, through a multi-methodological integration, it is now becoming possible to view human brain function in real time. This paper discusses one such integrative approach and illustrates its progress and pitfalls in the context of studies of human attention.

Visual Selective Attention

Focusing our attention on locations in the visual world aids our perceptions of relevant events and objects. Psychophysical studies have demonstrated that human performance in target detection and discrimination tasks is facilitated for stimuli presented within the focus of attention (see, e.g., Downing, 1988; Hawkins et al., 1990; Luck et al., 1994; Posner, Snyder, & Davidson, 1980). The precise neurobiological mechanisms enabling these perceptual benefits have been studied in animals through the use of direct neuronal recording (see, e.g., Moran & Desimone, 1985; Motter, 1993, 1994) and in humans through the use of recordings of the electroencephalogram (EEG; see, e.g., Eason, Harter, & White, 1969; Eimer, 1994; Mangun, Hillyard, & Luck, 1993; Van Voorhis & Hillyard, 1977). The evidence to date suggests that selective attention to locations (i.e., spatial attention) is correlated with changes in neuronal activity in both cortical and subcortical brain structures. One model supported by the data presently available is that stimulus processing in the visual cortex is altered as a function of attention; some of the findings are consistent with a gain control view for spatial attention (see, e.g., Hillyard & Mangun, 1987). Despite growing evidence, however, we do not currently have a complete understanding of the component parts of attentional brain circuitry or of how these functional brain components...

**Time Course of Attentional Processing in Humans**

Event-related brain potentials (ERPs) have provided high-resolution temporal measures of sensory information processing within the ascending sensory pathways and cortical sensory areas. The ERPs are extracted from the ongoing EEG using signal averaging; they reflect the responses generated by large populations of neurons whose extracellular electrical fields are summed and conducted to the scalp, where they can be recorded. In the 1960s, Eason and his colleagues (1969) used ERPs to investigate visual processing as a function of spatial attention, and they found that certain of the sensory-evoked ERP components were modulated by visual attention. Subsequent studies have documented amplitude modulations of sensory-evoked components during spatial selective attention that begin as early as 70-80 msec after the onset of an attended stimulus (i.e., the so-called “P1” component)—peak latency of approximately 110 msec; see, e.g., Eason, 1981; Harter, Aine, & Schroeder, 1982; Heinze, Luck, Mangun, & Hillyard, 1990; Heinze et al., 1994; Hillyard & Münthe, 1984; Mangun & Hillyard, 1987, 1988, 1990; Mangun et al., 1993; Neville & Lawson, 1987; Rugg, Milner, Lines, & Philp, 1987; Van Voorhis & Hillyard, 1977; see Figure 1). Some reports have suggested that spatial attention effects can begin prior to the latency range of the P1 attention effect (see, e.g., Oakley & Eason, 1990), thereby affecting earlier components of the visual ERP, such as the N100 (50–80 msec). However, these effects have proved difficult to replicate in similar studies (see, e.g., Clark & Hillyard, 1996; Gomez-Gonzalez, Clark, Fan, Luck, & Hillyard, 1994; Johannes, Münthe, Heinze, & Mangun, 1995; Mangun et al., 1993). Thus, it is possible to demonstrate in humans that spatial selective attention modulates neural activity related to relatively early stages of visual information processing, but perhaps not earlier than the cortical level for the visual modality. When attention is selectively directed to other, nonspatial stimulus features, such as color or shape, or to higher order stimuli, such as complex objects, the early P1 attention effect is not observed—that is, the P1 modulation is specific to spatial selective attention. Nonspatial attention leads to other, longer latency effects in the ERPs. In this paper, we will limit our discussion to voluntary spatial selective attention.

**A Mechanism for Spatial Selection**

How may one interpret the ERP correlates of spatial attention? The nature of the modulations of the ERP provides some clues. The effects of spatial attention on stimulus processing are observed to be changes in the amplitude but not in the latency of the P1 component. These amplitude modulations occur with little or no change in component waveshape or scalp distribution (see Mangun, 1995, for a review). This pattern has been interpreted as evidence that the neural generators of the early ERP components are activated by the sensory features of the stimulus, whether attended to or ignored, and that spatial attention modulates the activity of those neural structures. Hence, the conceptualization of these effects as a gain control over visual information processing has been proposed (see, e.g., Eason, 1981; Hillyard & Mangun, 1987). The idea is that descending neural projections from executive control circuitry in the frontal and parietal cortex modify the excitability of neurons in visual cortical maps. The precise mechanisms of such a system are not understood, but several versions of this type of model have been proposed (see, e.g., Harter & Aine, 1984; Hillyard & Mangun, 1987; Posner & Petersen, 1990), and some versions incorporate the idea that subcortical structures, such as the pulvinar nucleus of the thalamus or the superior colliculus, are involved in the overall system (see, e.g., Desimone & Duncan, 1995; LaBerge, 1990). One factor that has slowed the delineation of the neural circuitry of spatial attention is the possibility that spatially focused attention has different effects on stimulus processing at various stages of visual analysis. Thus, at one stage of processing, attention might act as a spatial gain control system for altering the signal-to-noise ratios of inputs from different regions of the visual field (see, e.g., Mangun, 1993), whereas, at another stage of analysis, the primary result of focal attention might be to change the tuning curves for selectivity of visual neurons for the features they code (see, e.g., Spitzer, Desimone, & Moran, 1988). Failing to appreciate the possibility of multiple, perhaps unrelated effects of attention makes formulation of a unified neural model an intractable problem.

**Intracranial Generators of Visual ERPs**

A key question that emerges from the ERP studies of attention is: Where in the visual hierarchy are the attention-sensitive ERP components being generated? In order to integrate scalp recording with the ever-growing knowledge from animal studies about the properties of specific visual areas and streams of processing within the visual hierarchy (see, e.g., Ungerleider & Mishkin, 1982; Van Essen & DeYoe, 1995), it is important to describe ERP effects in terms of functional anatomy. Although numerous studies in which multielectrode array were used have proposed that various visual ERP components prior to 150-msec latency were generated in either the striate or the surrounding extrastriate cortex (Clark & Hillyard, 1996; Gomez-Gonzalez et al., 1994; Mangun et al., 1993), until recently little consensus has existed as to the locus of generation of the major deflections of the visual ERPs. Indeed, significant controversy continues in the field with respect to this question.

Previous studies of the generators of visual ERPs have used combinations of stimulus manipulations and inverse dipole modeling (see, e.g., Clark & Hillyard, 1996; Gomez-Gonzalez et al., 1994; Johannes, Knalmann, Mangun, Heinze, & Münthe, in press; Mangun et al., 1993). Dipole modeling uses a computer model of a head and makes
Figure 1. Grand-average ERPs showing attended and unattended waveforms for upper left visual field stimuli and topographic maps at key time periods (from Mangun & Buck, 1998). The top panel shows a topographic voltage map for the time period (60–80 msec latency) during which the NP80 (sometimes called C1) component is maximal. The bottom panel shows a topographic voltage map for the time period (80–110 msec) during which the P1 component is maximal. On the right are shown the ERP responses from two electrode locations, OZ and OR, showing the NP80 and P1 components, respectively. The waveforms show an effect of selective spatial attention on the P1 component but not on the earlier NP80 (C1) component.

simplifying assumptions about the nature of the underlying neural generators (e.g., dipolar current source), in order to simulate the locations in the brain from which scalp-recorded activity might be originating (see, e.g., Scherg, 1992). Such methods, however, are limited in their ability to establish precise three-dimensional (3-D) localization of the generators of scalp-recorded activity. This problem stems from several unresolved theoretical and technical issues having to do with the difficulty in inferring intracranial localization on the basis of scalp recordings alone—the so-called inverse problem (see Dale & Sereno, 1993). The incorporation of magnetic recordings—especially, perhaps, the new generation of whole-head magnetoencephalographic (MEG) devices—provides some added benefit in addressing the inverse problem (see Hämäläinen, Hari, Ilmoniemi, Knuttila, & Lounasmaa, 1993). But assumptions about generator configuration, head and brain structure, and conductivity necessarily limit the ac-
curacy of simple dipole methods. The ultimate accuracy of a dipole model of human brain activity will depend on various factors—including such things as (1) the orientation and depth of the neural source, (2) the degree to which the source is discretely localized in the brain, and (3) whether or not the neural source is active at the same time that other confounding electrical events are being generated. Certain technical considerations are also, of course, important (see, e.g., Zhang & Jewett, 1994). Nonetheless, these methods permit models of possible intracranial sources to be tested and, when combined with other information, may permit the elimination of some hypotheses.

Gomez-Gonzalez and colleagues (Gomez-Gonzalez, et al., 1994; see also Clark & Hillyard, 1996) recently modeled the intracranial generators of the visual ERPs in a selective spatial attention task. They used a three-shell spherical head model and multiple, time-varying dipole solutions—that is, the BESA algorithm of Scherg (see, e.g., Scherg, 1992). In their model, they identified several candidate neural generators for early components of the visual ERP, including a striate component and an extrastriate (i.e., P1) component. The putative striate generator was not affected by the direction of attention, whereas the later, presumably extrastriate generator was modulated. They assigned this latter activity to locations in the lateral occipital visual cortex and argued that, as proposed by Mangun et al. (1993), this reflected activity in the extrastriate cortex. Recent related studies have drawn the same general conclusions (see, e.g., Clark & Hillyard, 1996; Johannes et al., in press).

The simple model of an early (50–80 msec) ERP component (i.e., NP80 or C1) that is generated in the striate cortex and a subsequent P1 component (70–130 msec) that is generated in the extrastriate cortex is consistent with their latencies (Robinson & Rugg, 1988), scalp distributions, sensitivities/insensitivities to retinotopic stimulus manipulations (Clark & Hillyard, 1996; Mangun et al., 1993), and their differential responsivity to the spatial frequency of the evoking stimulus (Zani & Proverbio, 1995). Despite the consistency of these findings, however, the precision with which such methods can accurately localize the anatomical site at which a given ERP component is generated remains uncertain. This latter, rather pessimistic view is based on the inability to verify the accuracy of inverse dipole solutions under normal physiological conditions—that is, once a solution is obtained, how well will we judge whether it is accurate? In the absence of converging evidence, the best we can hope for with present methods is to be able to address the extent to which any solution is neuroanatomically plausible. However, this caveat regarding the use of inverse dipole modeling does not invalidate its use in hypothesis testing, because, even in the absence of precise localization of any particular ERP effect, such techniques may allow one to have higher confidence in distinguishing between different activities—especially when they are sufficiently well separated within the brain (see, e.g., the modeling of the striate and extrastriate ERP components in Gomez-Gonzalez et al., 1994).

Functional Neuroimaging

The limitations in localization with the use of scalp-recorded ERPs—or event-related magnetic fields (ERFs) that are obtained by signal-averaging the MEG—should not be construed as failures of these techniques but rather, perhaps, as goals that may yet be met through algorithmic and technological developments (see, e.g., Dale & Sereno, 1993; George et al., 1995). In contrast, 3-D functional neuroimaging methods, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), can provide relatively high-resolution views of active brain regions by indexing regional cerebral blood flow (rCBF) that is coupled to neuronal activity. A severe limitation with PET and fMRI, however, is the extent to which temporal information is lost with these approaches. The time course of the hemodynamic responses that are used in most functional imaging studies using PET and fMRI is slow and is presently in the time range of seconds at best. With PET, the necessity of acquiring data over many tens of seconds means that PET is effectively blind to the temporal aspects of information processing. Functional MRI may do somewhat better than PET on this front, but the rather sluggish hemodynamic responses that give rise to the signals of interest (i.e., regional blood flow modulations) present us with an upper limit on the temporal resolution that these methods can achieve. What this limit may actually be remains to be determined, but it is likely to be something on the order of hundreds of milliseconds at best (see, e.g., Kim, Richter, & Ugurbil, 1997; Menon et al., 1995), by which time even a slow observer can discriminate and fully respond to a complex visual target.

If possible, the goal of human brain research is to achieve measures of brain activity that have both high spatial (i.e., on the order of a millimeter) and high temporal (i.e., on the order of milliseconds) resolution, in order to provide an understanding of the dynamics of those brain functions underlying human cognition. Unfortunately, there is no magic looking glass that provides this view of human brain function (however, see Fabiani, Gratton, Corbaliis, Cheng, & Friedman, 1998). Although intracranial recordings in neurological patients are very useful for some questions, their ultimate usefulness is limited because of their invasive nature and the necessary constraints imposed by medical considerations. Perhaps, then, the best approach at present is to combine methods in order to gain the desired picture of brain function (see, e.g., Fox & Woldorff, 1994; Walter et al., 1992). In recent studies of visual-spatial selective attention, we combined PET and ERP recording in normal human subjects in order to accomplish this highly desirable goal of viewing the time course of activity in localized brain regions (Heinze et al., 1994; Mangun, Hopfinger, Kussmaul, Fletcher, & Heinze, 1997).
Integrating ERPs and Neuroimaging

Our studies used both PET and ERP recordings to obtain high temporal (ERP) and high spatial (PET) resolution measures of brain activity from the same neurologically normal volunteers. In order to be able to relate the two recording modalities, we used identical stimuli in the separate PET and ERP sessions and compared identical conditions in each. The volunteers were required to focus their covert attention (eyes remained fixated on a central point) on one-half of a flashed bilateral visual symbol array. The PET 1H2O activation method was used to obtain measures of changes in regional cerebral blood flow as a function of selective attention.

The design of these studies was essentially the same as that used in several previous ERP studies we have conducted (see, e.g., Heinze et al., 1990; Heinze & Mangun, 1995). In separate blocks, the subjects were instructed to attend to either the right or the left half of bilaterally flashed stimulus arrays (Figure 2, left column) and to detect infrequent target stimuli (matching symbols) at the attended location; they also passively viewed the stimuli in separate blocks. Magnetic resonance imaging (MRI) scans were also obtained from each subject for use in spatial normalization of the PET images.

In the first experiment, we viewed changes in regional cerebral blood flow that occurred with the direction of visual attention by computing subtraction images (see, e.g., Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1991; Corbetta, Miezin, Shulman, & Petersen, 1993). The passive viewing condition was subtracted from the attend-left and the attend-right conditions. A significant increase in regional cerebral blood flow (activation) was obtained in the hemisphere contralateral to the attended visual hemifield (Figure 3). Thus, directing attention selectively to the right visual field produced a significant increase in blood flow in the left visual cortex, and attention to the left visual stimuli resulted in increases in the right visual cortex. These contralateral activations were on the ventral surface of the brain in the posterior fusiform gyrus of each hemisphere—an area that lies well outside the primary visual (striate) cortex and represents portions of the human extrastriate cortex. It has been suggested that this area represents a human homologue of V4 of the macaque monkey (see, e.g., Zeki et al., 1991) but may include V2 and V3 as well (Sereno et al., 1995). Macaque V2, V3, and V4 contain neurons that are sensitive to the direction of attention (see, e.g., Moran & Desimone, 1985; Motter, 1993, 1994; Spitzer et al., 1988). Interest-

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Luminance</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Symbol Image]</td>
<td>![Luminance Image]</td>
</tr>
</tbody>
</table>

**Figure 2.** Stimuli consisted of bilateral symbol arrays flashed in a rapid random sequence. The subjects fixated the central cross and attended left or right as instructed in different blocks—in this figure attention is indicated as dashed circles (attend right, shown). One task—the symbol discrimination condition (left half of figure)—was to detect symbol matches (defined as two of the same symbol, in the same orientation, on one side of fixation) at the attended location (Heinze et al., 1994, and Mangun et al., 1997). Another task—the luminance detection condition (right half of figure)—was to detect a small dot located within the area occupied by the symbols (Mangun et al., 1997).
ingly, however, in the macaque, attentional modulations of V4 neurons' transient responses to stimuli only occur when attended and unattended locations fall within the receptive field of a single neuron. This would not be the case in the present study, which used widely separate locations in the left and right visual hemifields. However, Moran and Desimone (1985) did show that attentional modulations with stimuli at widely separate locations were present in the inferotemporal (IT) cortex, a subsequent stage of processing in the ventral stream (Ungerleider & Mishkin, 1982). Moreover, V4 neurons do show changes in background firing rates when their receptive fields are in the focus of attention and prior to the arrival of the stimulus (Luck, Chelazzi, Hillyard, & Desimone, 1993). Thus, it remains unclear whether the fusiform gyrus effects obtained here are generated in a human homologue of V2, V3, V4, or IT and whether they reflect increased blood flow from enhanced background firing rates or enhanced transient-evoked responses (we will return to this latter issue later in this paper).

Significant PET activations were also obtained in the thalamus, right anterior cingulate gyrus, and left superior frontal gyrus. Interestingly, neither the striate cortex nor the parietal cortex was modulated by selective attention. Because there were thalamic activations, it is possible that visual sensory signals were actually modulated in the thalamus. However, the location of the PET activity does not indicate that this modulation reflected activity in the lateral geniculate enroute to the visual cortex. This possibility is also rendered unlikely because there were no effects in the striate cortex that received inputs directly from the lateral geniculate. If incoming visual signals were being affected at the thalamus, we should have observed these effects as they were passed upstream to the striate cortex.

An alternative interpretation is that the present increases in regional cerebral blood flow in the thalamus reflect activity in another thalamic nucleus. Perhaps the activations were related to somatosensory activity that resulted from the fact that subjects were pressing a handheld button and pressing to targets that were infrequent in the attention condition but not in the passive viewing condition. This is possible, but we did not observe significant activation in either the motor or the somatosensory cortex in similar subtractions. It is likely that this was because the subjects were responding rather infrequently in the present design. One likely possibility is that our activations in the thalamus were in the pulvinar nucleus, which is known, from human lesions studies (Rafal, Posner, Friedman, Inhoff, & Bernstein, 1988), animal studies (Petersen, Robinson, & Currie, 1989), and neuroimaging studies (see, e.g., LaBerge, 1990), to be involved in various aspects of attentional orienting. Indeed, the present data are consistent with a current model that proposes that the pulvinar nucleus of the thalamus is part of a network that controls visual information processing in the extrastriate cortex during attention (see, e.g., LaBerge, 1995).

The ERP attention effects that were obtained in a separate session included a significant P1 modulation (80–130 msec) over occipital scalp regions contralateral to the attended hemifield. Figure 4A shows the waveforms and topographical voltage maps of this P1 attention effect. As in the finding for the striate cortex in the PET results, the ERPs also showed no evidence for attentional modulations in the latency range that corresponded to activity in the striate cortex (e.g., 50–80 msec). Given the high degree of correlation between the contralateral ERP and PET effects, one is tempted to infer that the fusiform PET activations and the P1 modulations in the ERPs are causally related. However, further evidence would be required for such a conclusion. Nonetheless, as described in a later section, the fusiform activations from PET and the P1 modulations in the ERPs are highly correlated (see, e.g., Mangun et al., 1997).

Reference Frames for Integration Across Methods

Here we describe the theoretical structure inherent in our experimental design. This essentially represents the design controls that permitted us to relate the ERP and PET attention effects. The ways that the PET and ERP data could be related were first considered from the perspective of how one would relate two more similar modalities such as PET and MRI, in which many common frames of reference exist (i.e., primarily spatial ones). As a result, we refer to this approach as the frames of reference logic for combined studies. The following describes the so-called reference frames that were common between the PET and ERP data in the present experiment.

In the present study, a common experimental frame of reference was used—that is, identical experimental conditions were used in both the PET and ERP sessions, and identical comparisons between conditions were used for the analysis of the data (i.e., attend left minus attend right—see below). In order to prevent any differences arising because of different stimulus parameters, we also used a common sensory reference frame by having identical stimuli in all recording sessions. A common biological frame of reference was established by using the same human volunteers in the ERP and PET sessions. As a result of the foregoing, we can argue that both the ERP and PET results are directly related to the same mental process—that is, spatial selective attention. At this first may appear to be a reasonable justification for concluding that the present effects (fusiform gyrus activity in the PET data and the P1 modulation in the ERPs) are reflections of the same neural process—that is, selective spatial attention. Unfortunately, although such a conclusion may not be unreasonable, it is not without shortcomings. Although it is true that the PET and ERP effects are both related to spatial selective attention—that much is very clear—we cannot rule out the possibility that the two effects reflect attentional modulations at different stages of visual processing.

A final reference frame, the spatial frame of reference, is needed in order to establish a more firm relationship between the P1 attention effects and the fusiform gyrus activations. In order to do this, we must utilize all possible in-
formation about the possible spatial localization of our P1 attention effect. This was done using inverse dipole modeling. As we noted earlier, the inverse solution from scalp-recorded electrical data has no unique solution, and, thus, it may appear paradoxical now to consider using it as an aid in answering the question of how PET and ERP effects relate. However, although modeling of ERP data alone cannot give one unique solution, the neuroimaging activations can be used as additional constraints on the modeling, which results in functionally and anatomically plausible solutions.

**Inverse Dipole Modeling**

**Constrained by Functional Neuroimaging**

For the first study introduced above (Heinze et al., 1994), the relationship between the fusiform gyrus PET activa-

---

**Figure 3.** PET activations from Heinze et al. (1994). Left column shows difference images from attend-left blocks minus passive-viewing blocks; right column shows difference images from attend-right blocks minus passive-viewing blocks. When the subjects attended to the left, activations were found in right fusiform gyrus (seen in the top row of images labeled basal). Attention to the right resulted in activations in the left fusiform gyrus. The second row of images (central) shows activation in the thalamus, and the bottom row (superior) shows anterior cingulate gyrus activation. From “Combined Spatial and Temporal Imaging of Spatial Selective Attention in Humans,” by H.-J. Heinze et al., 1994, Nature, 372, p. 545. Copyright 1994 by Macmillan Magazines Ltd. Reprinted with permission.

---

**Figure 4.** ERP and dipole model results from Heinze et al. (1994). (A) Topographic voltage maps of the P1 attention effect for attend-left minus passive (left column) and attend-right minus passive (right column) conditions. ERP waveforms are shown from the electrode sites where the effect was maximal. Gray shadings on the waveforms indicate the time period over which the P1 was measured and plotted as topographic maps (80–110 and 110–130 msec). (B) At left is a coronal section through the brain that shows the PET activations in fusiform gyrus from the attend-left minus attend-right difference image. The topographic voltage map for the attend-left minus attend-right ERP difference waveform is plotted on the right. The dots superimposed on the PET activation image represent the locations of the best-fit inverse dipole solutions (when the model was unconstrained and the dipoles could move anywhere) for the attend-left minus attend-right ERP voltage topography. (C) Left column shows the topographic map generated by using the dipole locations from the best-fit inverse dipole solution over the 80–130 msec time range, which closely matched the recorded ERP effect (seen above—8, right column). Right column shows the topographic map generated by seeding the model with dipoles located at the center of the PET activations (dipoles were put at the center of the PET activations and not allowed to move, only to change orientation). Only minor differences were observed between the best-fit and PET-seeded dipole generated maps. From “Combined Spatial and Temporal Imaging of Spatial Selective Attention in Humans,” by H.-J. Heinze et al., 1994, Nature, 372, p. 545. Copyright 1994 by Macmillan Magazines Ltd. Reprinted with permission.
tions and the P1 attention effects in the ERPs was investigated using the Brain Electric Source Analysis (BESA) program for inverse dipole modeling (see, e.g., Scherg, 1992). First, however, in order to obtain maximal signal-to-noise in the ERPs and to maintain the logic of the experimental frames of reference, we did this modeling on the attention difference effects obtained by subtracting the attend-right conditions from the attend-left conditions for both PET and ERPs, using the grand-average data (collapsed across all subjects). This subtraction controls for effects of nonspecific arousal on PET and ERP recordings, because it compares two conditions that are equated for overall arousal, task performance, and attentional effort. The effects presented earlier that were derived by subtracting the passive condition from each active condition do not control for such effects. Even though the shift in the PET and ERP effects from one hemisphere to the other as a function of attending right versus left were useful indices that the effects were not merely based on behavioral arousal, the effects in the frontal cortex, thalamus, and anterior cingulate remain suspect in those subtractions. Indeed, the attend-left minus attend-right subtractions yielded only extrastriate activations as statistically significant (Figure 4B).

The modeling was constrained by using the functional anatomical information from PET to place the model dipole(s) within the 3-D volume of the head: it was then compared to the models obtained when those location constraints were removed (cf. Dale & Sereno, 1993). A first step was to use information from the PET activations to specify how many dipoles would be used in the model. The answer was two, one each for the left and right fusiform gyrus activity. We then used the loci of PET activations in the fusiform gyrus to specify where the dipole should be located—one in the center of each fusiform gyrus activation. These models are referred to as the seeded forward models. The dipoles were, however, permitted to reorient and adjust dipole moments during fitting because the PET activations in these data could tell us little about the orientation or strength of an electrical signal. Approximately 96% of the scalp voltage patterns could be explained by using the seeded forward models for the time range from 110–130 msec, when the P1 attention effect was largest (Figure 4C and Figure 5).

In a subsequent step of the modeling, two dipoles were placed in the left and right fusiform gyri but were totally unconstrained. Hence, they could move and adjust location, as well as orientation and strength. Approximately 98% of scalp voltage patterns could be explained this way (Figure 5), but the dipoles moved slightly anterior and laterally from the center of the fusiform PET activation. However, they did remain in the same general region and were closest to the PET activity for the time period corresponding to the P1 attention effect—peaking at 110–130 msec latency (Figure 5). As a first approximation, these results support the conclusion that the P1 attention effect is generated within the fusiform gyrus for upper visual field stimuli. Recent evidence from fMRI of human visual cortex suggests that, if our PET and ERP effects are generated in the human homologues of V2, V3, or V4, there might be a shift in the maximal localization to more dorsal scalp regions for lower field stimuli (Sereno et al., 1995). Although we did not test the lower visual field in this study, Woldorff et al. (1997) have recently found such an effect to lower field stimuli by using combined PET and ERP methods.

In the Heinze et al. (1994) study, the task of matching the ERP and PET effects was made more simple by the fact that there was only one significant PET activation in each hemisphere in the visual cortex. Considerable difficulties would be presented if the task had been to match the ERP effect to more than two PET activations. In the first study, the effects were highly circumscribed in both modalities of recording, but there are no a priori reasons to hope for such simplicity. Indeed, given the fact that the visual cortex contains multiple maps of the visual world and separate streams of processing, it is rather curious that only two spatial attention-related activations were obtained in PET. Indeed, in the second study using 3-D PET acquisition (Siemens ECAT EXACT scanner), significant activations in the fusiform gyrus were replicated, and additional activations were found in the middle occipital gyrus of the hemisphere contralateral to the attended hemifield (Mangun et al., 1997).

Determining the ERP Generator Given Multiple Brain Activations

Given more than one PET activation, it is not possible to simply infer that activity in one recording modality is the same activity as that recorded in the other modality. One approach would be to proceed as described previously, using dipole models. In this case, one could seed the dipole model with four dipoles (two in each hemisphere) and solve for their orientations and strengths over time. As we did before, one could then compare the results of this model to the case in which the dipoles would be allowed to move their positions and observe which model best explains the recorded ERP results. This approach is logical for dipole modeling, but, as with all such modeling exercises, must be undertaken with caution.

Here we describe one additional means of relating ERP to PET or fMRI activity. This involves the idea that PET and ERP effects that are generated by the same neural and computational processes should covary with one another as a function of stimulus and task demands. In this study, we chose to manipulate the perceptual load of the target task. We did this by comparing two different task conditions within the framework of the spatial attention design. In separate blocks, the subjects performed two different tasks, but for (virtually) the same physical stimuli. One task was identical to that of Heinze et al. (1994) and involved a difficult symbol discrimination task (subjects had to respond to matching symbols at the attended location—Figure 2, left), whereas the other task was a simple luminance detection task (subjects had to respond to a small dot appearing within the confines of the bilaterally flashed symbol arrays—Figure 2, right). The same conditions were used in both tasks—attend right and attend left—as well as the same passive viewing and fixation conditions, but we only compare and describe the active attention con-
Dipole-Fitting

Residual Variance [%]

PET Seeded
Best Inverse

0 10 20 30 40
60 80 100 120 140 160 180 200 220
Time (msec)

Figure 5. Residual variance and dipole fits, compared with PET activation in fusiform gyrus from Heiske et al. (1994). Top: graph depicting the scalp voltage variance not accounted for when measuring the goodness-of-fit between the observed ERP scalp distribution of the attend-left minus attend-right difference waveform and the dipole models. Fit was measured over 20-msec time windows between the time period of 50–230 msec. Top trace shows the fit to the dipole model that was seeded with the PET activation locations. Bottom trace shows the fit to the unconstrained two-dipole model. Lower values of residual variance represent closer correspondence between the ERP recorded activity and the dipole model. At bottom are shown sagittal (right hemisphere) and coronal (occipital) slices showing the PET fusiform activation and the locations of the right hemisphere dipole at different time periods. The best-fit dipole locations in the P1 time range (gray circles) were shifted slightly from the center of the PET activation but were still in the region of the PET activation—these dipoles explained 98.4% in the 110–130 msec interval, corresponding to the peak of the P1 effect. From “Combined Spatial and Temporal Imaging of Spatial Selective Attention in Humans,” by H.-J. Heinze et al., 1994, Nature, 372, p. 545. Copyright 1994 by Macmillan Magazines, Ltd. Reprinted with permission.

ditions here. Although a significant main effect of attention in the ERPs was obtained for the P1 in both tasks, there was also an interaction between attention and task. This resulted from the fact that the more difficult discrimination task induced larger attention effects in P1 amplitude than did the detection task (Figure 6). In the PET data, we found significant main effects of attention (left vs. right) in the posterior fusiform gyrus and middle occipital gyrus contralateral to the attended hemifield (Figure 7A). We then proceeded to analyze the interaction of task and attention in the PET data. An analysis of the interaction of attention by task revealed that only the fusiform gyrus activation was significantly modulated in a manner similar to the P1 attention effect, as shown in SPM maps of the interaction effect in Figure 7B. This variation between the P1 effect and the fusiform gyrus PET effect strengthens our conclusion that the stage of visual processing indexed by the P1 component occurs in the extrastriate cortex in the posterior fusiform gyrus. This method of utilizing different task parameters in order to analyze covariances between attention effects observed in different recording modalities can serve as a strong test of hypothesized relationships, especially when the measures have been recorded from the same group of human subjects.

Relationship of Electrical to Blood Flow Measures

In the present context, it is important also to consider some caveats to the approaches currently employed. A fundamental question that still needs to be addressed is the
nature of the mapping between electromagnetic recordings and measures of regional cerebral blood flow (rCBF). That is, what is the relationship between activity measured by PET or fMRI versus that indexed by ERPs or ERPs?

The ERP responses that we study reflect the electrical current generated directly by postsynaptic activity in neurons (see Nunez, 1981). By recording ERPs time-locked to a stimulus, we measure a transient-evoked response that is composed of one of two types of neural activity: exogenous (stimulus driven) responses and endogenous (non-stimulus driven but observed to be time-locked to a stimulus or behavioral response) responses. Another type of electrical activity recorded with ERPs are slow potentials, such as the contingent negative variation (CNV), and other slow shifts in scalp voltage. The latter activities are not signs of stimulus-evoked activity but reflect changes in
cortical excitability. Whereas ERPs directly measure neuronal activity, measures of regional cerebral blood flow such as ¹⁵O-PET view this activity indirectly by measuring the blood flow changes that are coupled to neuronal activity. It remains unknown whether these recorded blood flow changes primarily reflect transient stimulus-evoked activity, changes in cortical excitability that is not stimulus-evoked, or both. Although ERP and PET measures are reflections of different physiological signals, they may nonetheless arise as the result of identical information processing activities occurring in the neuronal population. However, various possible mappings between the ERP and PET effects must be considered.

The possible mapping of ERPs/ERFs to PET/IMRI (rCBF) with respect to attention (as well as more generally) are: (1) the ERP and rCBF measures may both be reflections of identical neuronal activity—for example, both may result from either (a) stimulus-evoked activity differing with attention or (b) cortical excitability that is not stimulus-evoked per se but differs with attention; (2) the ERP and rCBF measures may be reflecting different aspects of neuronal activity but in the same neurons and as the result of the same attentional influence—for example, the sensory-evoked ERPs could be transient-evoked activities, whereas the rCBF changes could be more closely associated with changes in excitability and background firing rates but in the same neuronal population; (3) the ERP and rCBF measures are both reflections of differential processing with attention but in entirely separate neuronal populations, because each is either reflecting different types of neuronal activity or similar types of activity segregated in different visual cortical areas. In principle, another possibility exists—that they have no relationship whatsoever. However, if the experiment has been designed under the constraints of the frames of reference approach discussed earlier, this latter possibility makes little sense, because the effects in electrical and rCBF measures are by definition related via the experimental circumstances under which they were revealed—that is, same stimuli, same tasks, same experimental comparisons, and same subjects.
Which of the foregoing possibilities is most likely to be the case? In our view, at present they are all equally likely, and, indeed, all are probably going to be true to varying extents, depending on the process and brain system under investigation. They key question is: How correlated are the two measures in any given brain region or for any given behavioral or cognitive task? For example, it may be that the PET and ERP measures in the visual areas are reflections of stimulus-evoked activity, whereas, in the frontal cortex, PET and ERP effects might be more likely to be the result of changes in neuronal excitability. Studies in monkeys have clearly demonstrated that both tonic and transient-evoked neuronal modulations are present during spatial selective attention (Luck et al., 1993). Thus, the position we take is not to assume perfect correspondence but instead to build an experimental framework in which the effects are related both to a specific mental operation (as defined by the study) and to specific comparisons, and then to ask, through modeling and experimentation, just how the measures might be related for each brain activity and cognitive process under investigation.

Conclusions

These data demonstrate the logic of using combined electromagnetic and functional neuroimaging methods to study brain processes. The general logic of the frames of reference approach permits the activities of two very different recording methods to be related through their common tie to a specific experimental task and theoretical question. The use of electrical modeling to assist in carrying out this task is a natural one. Indeed, one may consider the present methods as merely additional means for constraining the ill-posed inverse problem by using information from functional neuroimaging as an extension of prior attempts to constrain such exercises by using only anatomical information (see, e.g., Dale & Sereno, 1993). In addition, the analysis of covariance between effects from different recording modalities across multiple tasks adds to the potential power of combining these approaches. Indeed, we suspect that, in many cases, the most powerful approach for relating electromagnetic to blood flow measures will simply be to demonstrate that they covary over a wide range of manipulations.

It is also abundantly clear that, given the unique sensitivities of each recording modality, there is no a priori reason to assume that every effect observed in one modality should necessarily have a correlate in recordings done using another method. Thus, the idea that the problem is merely how to match rCBF activations to electromagnetic effects is incorrect. Experimental designs must be formulated with the specific goal in mind of integrating these distinct measures; analyses must be executed in such a way as to make the results comparable; and procedures must be utilized that allow the combination of the results to more rigorously test hypotheses about the spatiotemporal dynamics of mental processes. Although the promise of integrating different physiological measures of brain activity is immense for human cognitive neuroscience, significant unresolved issues remain that will continue to challenge us as we pursue this goal. Of one thing we remain confident: As more of these issues continue to be resolved, these integrative approaches will lead to ever more exciting research into the basis of human cognition.

REFERENCES


sustained and transient attention to spatial locations. Neupropsychologia, 33, 889-908.


