11 The Operation of Attention—Millisecond by Millisecond—Over the First Half Second

Steven J. Luck

11.1 Chapter Overview

This chapter reviews two decades of progress in mapping out the time course of visual attention over the first half second. In particular, the chapter will examine what kinds of attention-related processes operate at various times after the onset of a stimulus. Traditional behavioral methods do not typically directly reveal the precise time at which attention operates, and the chapter will therefore emphasize techniques that provide a continuous, millisecond-level measure of processing, especially the event-related potential (ERP) technique.

11.2 The ERP Technique

Because the ERP technique is not as well-known as some other techniques, it is worth spending a few words on the essential aspects of this technique, particularly with respect to its temporal properties. ERPs almost exclusively reflect postsynaptic potentials that occur when neurotransmitters bind with receptors, causing ion channels to open or close (see Allison et al., 1980; Nunez, 1981). This contrasts with single-unit and multi-unit recordings, which measure the action potentials that occur if the summed postsynaptic potentials across many synapses reach the neuron’s threshold for firing. Thus, ERPs reflect the inputs to a neuron, whereas single-unit and multi-unit recordings reflect the outputs from a neuron.

Postsynaptic potentials in neurons are accompanied by changes in the local electrical field, and a nearby electrode can easily measure these local field potentials. The electrical potentials elicited by different neurons combine together by simple addition, and if postsynaptic potentials occur in many similarly oriented neurons at the same time, the sum of the local field potentials will be large enough to record at a distance. Local field potentials can therefore be recorded from within the cortex or from the cortical surface in patients undergoing exploratory neurosurgery. The
brain is a highly conductive medium, and the electrical fields resulting from post-synaptic potentials at one location therefore spread throughout the entire brain, although they diminish in size as the recording electrode is moved away from the source. The skull has a higher resistance than the brain, and electricity follows the path of least resistance, so much of the electrical activity spreads laterally when it reaches the skull. However, some of the activity does pass through the skull to the scalp, where it can be recorded noninvasively.

The electrical potentials recorded on the scalp or the cortical surface are termed the *electroencephalogram* (EEG). The EEG reflects the summed activity of many different brain systems operating in parallel to perform various tasks, and when an experimenter presents a stimulus, only a small portion of the overall EEG reflects the processing of this stimulus. Consequently, if the scalp EEG following a single stimulus is examined, it is difficult to extract the response to the stimulus from the other, unrelated sources of electrical activity. The simplest way to extract the stimulus-related activity is to average together the EEG segments from a large number of stimulus presentations. Any consistent stimulus-evoked activity should be present on every trial, whereas any unrelated EEG activity will sometimes be positive and sometimes be negative. Consequently, the consistent stimulus-related activity will remain in the average, but unrelated EEG activity will cancel out and approach zero if enough trials are averaged together. In the experiments described in this chapter, at least 100 trials were averaged together for each condition in each subject. The resulting averaged waveforms are called *event-related potentials* because they are electrical potentials that are related to a specific event, such as the presentation of a stimulus.

Figure 11.1 shows an example of the ERP waveform elicited by a flashed visual stimulus. Note that time zero represents the onset of the stimulus and that, by convention, negative is plotted upward. A series of positive and negative peaks can be observed following the onset of the stimulus, reflecting the flow of information through the visual system and into higher cognitive and response systems. The peaks are also called *waves or components*, and they are typically labeled with a *P* or *N* to indicate whether the peak is a positive- or negative-going one, along with a number to indicate timing (e.g., *P1* for the first major positive peak, or *P115* to indicate its precise latency of 115 ms). For visual stimuli, an early component arising from area V1 is often observed, and this component is often called *Cl* because it may be either positive or negative, depending on the spatial location of the evoking stimulus (the variations in Cl polarity are caused by the distinctive folding pattern of area V1 in the human brain). The Cl component is followed by P1 and N1 components that are modality specific. That is, the P1 and N1 waves elicited by a visual stimulus are completely unrelated to the P1 and N1 waves elicited by an auditory stimulus (they have the same names across modalities only because of their
Figure 11.1.
Examples of early (A), intermediate (B), and late (C) effects of attention on the visual event-related potential (ERP) waveform. Each waveform shows how voltage varies over time. By convention, negative is plotted upward. Time zero represents the moment of stimulus onset.
temporal order and polarity, not because of any functional relationship). By the time of the P2 and N2 waves, a combination of modality-specific and modality-independent activity is present, and the P3 wave is largely modality independent.

### 11.3 Spatial Resolution and Magnetic Recordings

Cognitive neuroscience techniques are often described along the dimensions of temporal resolution and spatial resolution. The ERP technique has relatively poor spatial resolution. One factor that contributes to this is the high resistance of the skull, which causes the electrical field to spread laterally. To avoid this, it is possible to record the magnetic field that accompanies the EEG, which is called the magnetoencephalogram. Electrical dipoles are always accompanied by magnetic fields, and changes in the magnitude of the electrical potential lead directly to changes in the strength of the magnetic field. Consequently, one can obtain essentially the same information by recording magnetic rather than electrical activity, and the same averaging approach can be used, yielding event-related magnetic fields (ERMFs) rather than ERPs. The advantage of ERMF recordings is that the skull is transparent to magnetism, eliminating the spatial blurring that is present in ERP recordings and providing improved spatial resolution. However, the magnetic fields are extremely small, especially in comparison to the Earth's magnetic field, and recording ERMFs requires extraordinarily sensitive and extremely expensive equipment. In addition, ERP/ERMF localization is much more indirect than, for example, localization of the blood oxygen level dependent (BOLD) signal in functional magnetic resonance imaging (fMRI) experiments. In particular, ERMFs and ERPs are localized by means of mathematical modeling techniques, and the models are evaluated by testing their ability to fit the observed pattern of data. The models typically have very large numbers of free parameters (tens, hundreds, or even thousands), and a given data set can potentially be fit by an infinite number of models. Consequently, ERP/ERMF localization models are not usually definitive. That is, they provide a useful suggestion about the likely generator source of a given signal, but it is not usually possible to provide a margin of error, a likelihood ratio, or some other quantification of confidence in the localization.

### 11.4 Temporal Resolution of ERPs

Although the ERP technique does not have high spatial resolution, it has very high temporal resolution. Electrical potentials pass through conductors essentially instantaneously, so fluctuations in voltage recorded at the scalp provide a precise index of fluctuations of neural activity within the brain. The bioelectric properties
of the brain, skull, and scalp presumably lead to some temporal distortion of scalp-recorded ERPs, but the temporal resolution can be in the submillisecond range under optimal recording conditions. In most attention experiments, the effective resolution is probably on the order of 5–10 ms. This lower resolution is caused by several factors, including the use of filters to attenuate high-frequency electrical noise. However, this level of temporal resolution is entirely sufficient to answer most questions about attention, and higher temporal resolution is possible when necessary.

Although many writers have praised the temporal resolution of ERPs, raw temporal resolution is not really the main advantage of this technique. After all, it would be trivial to record button-press reaction times (RTs) with submillisecond resolution. The key advantage of ERPs over behavioral and neuroimaging measures is that ERPs provide a high-resolution and continuous measure of processing. That is, whereas an RT measure provides no information about processing between the time of the stimulus and the time of the response, ERPs provide a continuous measure of processing between the stimulus and the response (and even after the response). Neuroimaging methods can also provide a continuous measure of processing, potentially sampling at relatively high rates. However, current methods primarily measure the consequences of changes in blood flow, and these changes are very slow, so the effective temporal resolution of current neuroimaging methods is typically on the order of a few seconds (although resolution on the order of a few hundred milliseconds may be possible under certain conditions). Consequently, ERPs and ERMFs provide the only mainstream techniques for providing noninvasive, continuous measures of processing with high levels of temporal resolution.

11.5 Limitations of the ERP Technique

Although the ERP technique has many advantages over other techniques, it also has some significant limitations. First, as already mentioned, ERPs do not provide a direct means of determining the neuroanatomical sites of activation. Second, ERP waveforms are almost always based on averages of tens or hundreds of trials, and any trial-to-trial variations in the operation of attention may be lost in the averaged waveforms. In a visual search task, for example, attention might shift in a different spatial pattern on each trial, and this would be difficult to see in an averaged ERP waveform (note, however, that this problem is not entirely impossible to solve—see Woodman & Luck, 1999, 2003b). In addition, activity that is evoked by the stimulus but varies randomly in phase will average to zero in the averaged ERP waveforms (but this can be addressed by spectral analysis methods—see, e.g., Tallon-Baudry et al., 1996). Third, the observed ERP waveform consists of a set of inter-
nal components that sum together, and it can be very difficult to determine exactly which ERP component is influenced by an experimental manipulation. For example, attended auditory stimuli elicit a more negative response than ignored auditory stimuli in the time range of the N1 wave, but it has been difficult to determine whether this effect reflects an amplification of the sensory-evoked N1 wave or the addition of overlapping endogenous brain activity (Hansen & Hillyard, 1980; Naatanen, 1975; Woldorff et al., 1993). Finally, like the BOLD signal in fMRI experiments, the relationship between scalp-recorded ERPs and specific neural events is somewhat indirect, and it is therefore challenging (although not impossible) to draw precise conclusions about the information-processing transactions represented by a given ERP effect.

11.6 Measuring the Time Course of Attention

A straightforward way to study the timing of attentional processes is to examine the ERP waveform elicited by a given stimulus when it is attended and when it is ignored, determining the time point at which the waveform differs due to the allocation of attention (for classic examples, see Eason et al., 1969; Hillyard et al., 1973; Van Voorhis & Hillyard, 1977). As illustrated in figure 11.1A, attention might suppress the response to an ignored visual stimulus relative to an attended stimulus beginning in area V1, leading to a reduced C1 amplitude (and reductions in all of the subsequent ERP components). In contrast, attention may have no effect on the initial sensory response, influencing activity beginning at the time of the N1 wave (see figure 11.1B). Or, as proposed by late-selection theories of attention, attention may have no effect until the time of late, modality-independent ERP components, such as the P3 wave (see figure 11.1C). Given that visual information processing does not consist of a set of discrete, serially ordered stages, it is also possible that attention could influence some early components without influencing all of the later components.

In this manner, ERPs can be used to pinpoint the time at which attention begins to influence processing. Two caveats are necessary, however. First, the absence of an early effect does not by itself imply that attention only operates at late stages. Only a fraction of neural activity can be detected with ERP recordings, and it is always possible that an early effect is present but cannot be observed with ERPs (e.g., the individual local field potentials may be oriented in a manner that causes them to cancel out). A second caveat is that raw timing is of limited use unless it can be linked to specific processes. Given what we know about the time course of information processing in visual cortex (see, e.g., Schmolesky et al., 1998), it is usually safe to conclude
that any effects of attention arising from visual cortex within 100 ms of stimulus onset reflect a modulation of perceptual processing rather than modulations of higher level cognitive processing (although some prefrontal activity begins this early—see Foxe & Simpson, 2002). And given the use of brief stimuli, any effects of attention after 300 ms poststimulus can be assumed to reflect modulations of higher level processing and not perceptual modulations. However, this leaves a large ambiguous range (100–300 ms), and it makes more sense to treat the perceptual–cognitive distinction as continuous rather than categorical. Consequently, it is important to supplement raw timing information with other sources of evidence about the specific nature of the attention effects. The next sections will summarize what has been learned about the effects of attention in this type of experimental paradigm, in which attention is explicitly directed toward or away from a given stimulus source. We will begin with studies in which the stimulus source is defined by its location and then move on to other feature dimensions.

11.7 Visual-Spatial Attention: Timing and Neural Substrates

Many ERP experiments have examined visual-spatial attention using variations on the sustained attention paradigm illustrated in figure 11.2A. In this paradigm, subjects fixate a central point, and fixation is monitored to ensure that covert rather than overt attention is manipulated. Stimuli are flashed in a rapid, random sequence to locations in the left and right visual fields, and subjects are instructed to monitor one of the two locations for target stimuli for a period of a few minutes. Most of the stimuli are nontarget stimuli, and the targets are defined by a subtle difference in some dimension, such as size. Subjects make a button-press response when they detect targets in the attended location, and they are told to ignore all stimuli in the other location. They attend to one location on some trial blocks and attend to the other location on other trial blocks, making it possible to compare the response to a given stimulus when it is attended versus when it is ignored. Because the task is identical for the attend-left and attend-right conditions and the order of stimuli is unpredictable, the subject’s state of arousal is equated across conditions. Thus, any differences in the ERPs elicited by a given stimulus when it is attended compared to when it is ignored can be attributed to specific spatial attention effects.

This experimental paradigm is similar to the spatial cuing paradigm that is popular among behavioral attention researchers, but it differs in two key ways. First, after attention is directed to a specific location, only one stimulus is presented in the spatial cuing paradigm, whereas a long stream of stimuli is presented in the sustained attention paradigm (and attention must be sustained to the cued location through the entire stream). The use of a long stream of stimuli is helpful in ERP
Figure 11.2

(A) Schematic version of sustained attention paradigm that is commonly used to assess the time course of attention in event-related potential experiments. Subjects are instructed to attend to the left or right visual field at the beginning of each trial block and press a button when slightly shorter target stimuli are presented at that location. Subjects are also required to maintain fixation on the central fixation point at all times, and accurate fixation performance is monitored. (B) Typical results from this type of experiment (from the study of Gomez Gonzales et al., 1994). The C1 component is not influenced by the direction of attention, but the P1 and N1 components elicited by a stimulus are larger when that stimulus appears on the attended side than when that stimulus appears on the unattended side.
research because it allows many more stimuli to be presented in a given amount of time (i.e., less time is “wasted” on presenting cues). This is important for achieving reasonably low-noise averaged ERP waveforms. The second difference is that subjects in the spatial cuing paradigm are instructed to respond to targets in the uncued locations as well as targets in the cued location, whereas subjects in the sustained attention paradigm are instructed to respond to targets only in the cued location. Responses to uncued targets are simply unnecessary when ERPs are recorded. These differences do not seem to have a substantial impact on attentional processes, because the same basic pattern of results has been obtained in both types of paradigms (see, e.g., Mangun & Hillyard, 1991).

In most studies using the sustained attention paradigm, the P1 and/or N1 components have been found to be larger for attended-location stimuli than for ignored-location stimuli, whether the stimuli are targets or nontargets. Some of these studies were also designed to evoke a C1 wave (from area V1), and these studies have demonstrated that the C1 wave is not influenced by attention. An example of this is shown in figure 11.2B (from the study of Gomez Gonzales et al., 1994). The P1 effect typically begins between 70 and 100 ms poststimulus, indicating that attention influences perceptual processing in this paradigm. The lack of a C1 effect suggests that attention does not influence processing in or prior to area V1, although it is possible that attention influences some sort of V1 response that does not contribute to scalp-recorded ERPs.

Heinze and his collaborators attempted to localize the P1 attention effect by combining ERP recordings with positron emission tomography (PET) imaging in a paradigm similar to that shown in figure 11.2A (Heinze et al., 1994). The PET data indicated an effect of attention in the fusiform gyrus, along the ventral surface of the occipital lobe. Mathematical modeling procedures demonstrated that this location was a plausible source of the P1 attention effect. Specifically, a pair of electrical dipoles located in the left- and right-hemisphere fusiform gyri could explain 96% of the scalp distribution of the P1 wave. More recent studies combining ERPs with fMRI have replicated this finding but have also shown that a portion of the P1 wave appears to arise from dorsal extrastriate areas (e.g., dorsal V3—see Martínez et al., 2001). Less work has focused on localizing the N1 attention effect (see Clark & Hillyard, 1996). This is due, in part, to the nature of mathematical modeling procedures: The N1 effect is overlapped by the P1 effect, and small errors in localizing the P1 effect could lead to larger errors in localizing the N1 effect.

Research using this paradigm has also been conducted with monkeys, using single-unit recordings (Luck et al., 1997a). Consistent with the human ERP studies, no effects of attention were observed in area V1. Attention effects were observed in areas V2 and V4. One key finding was that the effect of attention in V4 began at 60 ms poststimulus, which was the onset time of the sensory response in that area.
This is consistent with the proposal that visuospatial attention may serve as a sensory gain control, increasing or decreasing the feedforward transmission of information through extrastriate cortex (Hillyard et al., 1998). Another key finding is that a given neuron’s activity was modulated by attention only when both the attended and ignored locations were inside that neuron’s receptive field (typically separated by no more than 3° of visual angle). The human ERP studies have typically used very distant locations that would never fit within a V2 or V4 receptive field, suggesting that the P1 and N1 effects arise from a region such as inferotemporal cortex, where the receptive fields are very large.

Together, these studies indicate that visuospatial attention does not influence neural responses in area V1 but does change the gain of feedforward sensory transmission in extrastriate visual areas within the first 100 ms after stimulus onset. In contrast, several fMRI studies have found effects of attention in area V1 (e.g., Gandhi et al., 1999; Somers et al., 1999). There are two likely explanations for this discrepancy. First, it is possible that the fMRI studies used a task that was more appropriate for engaging attentional mechanisms in area V1. Second, it is possible that the fMRI effects reflect some kind of feedback signal rather than a modulation of feedforward sensory activity. Current evidence points to the second of these possibilities. First, a single-unit study using the curve-tracing paradigm has shown that attention effects are present in area V1 but that these effects begin more than 200 ms after stimulus onset (Roelfsema et al., 1998). Second, Martinez and her collaborators recorded both fMRI and ERP data from a group of subjects engaged in a spatial attention task similar to that shown in figure 11.2A, and they found an attention effect in area V1 in the fMRI recordings but no modulation of the C1 wave in the ERP recordings. Because the same task was used for both types of recordings, task differences cannot explain the discrepancy between the fMRI V1 results and the ERP C1 results. Third, additional studies using ERMs or a combination of ERPs and fMRI have confirmed the finding of no effect on the initial sensory response in area V1 but have shown a V1 effect beginning around 150 ms (Di Russo et al., 2003; Martinez et al., 2001; Noesselt et al., 2002). Thus, the evidence so far indicates that any attention effects in area V1 reflect a feedback signal of some unknown origin rather than a modulation of feedforward sensory activity. We are just beginning to understand the role of feedback in area V1 (see, e.g., Lamme, 1995), so the functional role of this feedback effect is not yet known.

11.8 Visuospatial Attention: Enhancement Versus Suppression

Although these data suggest that attention operates like a gain control within extrastriate cortex, this is probably an oversimplification. For example, it is not the case that all of the ERP components from the P1 wave onward are decreased by a
constant factor for ignored-location stimuli compared to attended-location stimuli. Indeed, it is possible to obtain a P1 modulation without an N1 modulation and vice versa. Some studies have therefore attempted to provide a more detailed characterization of the nature of the P1 and N1 attention effects. Several dissociations between the P1 and N1 effects have been described, and these differences led to the proposal that the P1 effect reflects a mechanism that suppresses activity at unattended locations to avoid interference, whereas the N1 effect reflects a limited-capacity discrimination mechanism that is applied to the attended location (see the review by Luck, 1995).

The proposal that the P1 effect reflects distractor suppression whereas the N1 effect reflects target enhancement was based, in part, on studies in which a neutral attention condition was included (Luck & Hillyard, 1995; Luck et al., 1994). In one study, for example, attention was directed to one of four locations on each trial by a central arrow cue, and neutral trials were included in which all four locations were cued. The P1 wave was found to be suppressed on invalid trials, in which one location was cued and the target appeared at a different location, compared to neutral trials, in which attention was distributed over all four locations. However, when compared to the neutral trials, no enhancement of the P1 was observed on valid trials, in which one location was cued and the target appeared at the cued location.

Another important difference is that P1 attention effects are observed primarily when the target discrimination might be impaired by distracting information, whereas N1 effects are observed when a discrimination is being made, regardless of whether distraction is likely. For example, the P1 effect was observed in a task that required subjects to combine two features of a target that were individually present in the distractors (Luck et al., 1993), but this effect was eliminated in a nearly identical task in which the target was defined by the presence of a single feature (Luck & Hillyard, 1995). This fits with the idea that attention is necessary to combine features but not to detect them (Treisman & Gelade, 1980). The N1 effect was present in both of these conditions, but it was eliminated when subjects made a simple detection response rather than a discriminative response (Magnus & Hillyard, 1991). Further studies have shown that the N1 wave is larger when subjects perform discrimination tasks than when they perform detection tasks, even if all of the stimuli are presented at a single, attended location (Ritter et al., 1979; Vogel & Luck, 2000) and that this effect arises from ventral-lateral regions of occipito-temporal cortex (Hopf et al., 2002).

11.9 Nonspatial Features

Spatial location plays a distinctive role in visual representation. Beginning with the lenses of the eyes and continuing through area V1, the visual system uses special
mechanisms and considerable resources to coding space precisely and explicitly. Space has also been accorded a special status in several theories of attention (e.g., Logan, 1996; Nissen, 1985; Treisman & Gelade, 1980). ERP studies have supported a special status for space in visual attention, because spatial attention effects typically begin earlier than nonspatial attention effects.

Early studies of attention to nonspatial features used tasks that were analogous to the spatial attention paradigm shown in figure 11.2A, except that some other feature dimension replaced the dimension of spatial location (see, e.g., Anllo-Vento et al., 1998; Harter & Guido, 1980; Harter & Previc, 1978). In the study of Anllo-Vento et al. (1998), for example, subjects were instructed to monitor a stream of foveal stimuli, some of which were red and some of which were blue, attending to one color and ignoring the other color. The ERPs elicited by a given color when that color was attended were then compared to the ERPs elicited by that color when it was ignored. Feature-based attention effects have not been found to modulate the P1 wave; instead, the effects typically begin in the latency range of the P2 and N2 components. The most common pattern is a more negative-going waveform that begins around 150 ms poststimulus and lasts for 100–300 ms (this effect is often called a selection negativity).

Hillyard and Münte (1984) combined spatial and nonspatial attention within a single experiment to further define how they are related. Subjects attended to a particular color at a particular location, making it possible to examine the responses elicited by a stimulus that was attended along both dimensions, along only one dimension, or along neither dimension. Space was found to dominate color. Specifically, attended-location stimuli were found to elicit larger P1 waves than ignored-location stimuli, regardless of whether the color was attended or ignored. In contrast, a selection negativity was observed for attended-color stimuli compared to ignored-color stimuli, but only for attended-location stimuli. Unattended-location stimuli of the attended and ignored colors elicited nearly identical ERP waveforms. Thus, attention appeared to operate hierarchically, with location-based selection occurring first and color-based selection occurring only for attended-location stimuli. However, this pattern may obtain only if attention is strongly focused in advance of stimulus onset: As will be discussed later, feature-based attention may often be used to guide spatial attention to likely target objects.

11.10 The Visual Search Task

In sustained attention and trial-by-trial cuing paradigms, attention may be focused on a location prior to stimulus onset, making it possible for attention to modulate the initial feedforward wave of sensory activity evoked by the stimulus. In many
real-world situations, however, attention shifts among items that are already present, guided by feature information. Imagine, for example, that you are eating a meal in a restaurant and that you would like to eat a green bean. To select a single bean as the target for your fork, you will shift first the spatial focus of attention to the pile of green beans to determine which bean to eat. In this situation, you will likely use both the color of the green beans and your knowledge of their general location on the plate to shift attention appropriately.

In the laboratory, visual search tasks are used to approximate this sort of real-world situation. Subjects are told to search for a target defined by a particular set of features (e.g., a long green line), and then an array of objects is presented that may or may not contain the target. In most cases, subjects press one of two buttons to indicate whether the target was present or absent. Feature information is typically used to guide attention to likely target objects (e.g., subjects might shift attention to each green item to determine whether it is the long, green target). Spatial information can also be used to guide attention (Chun & Jiang, 1998).

11.11 The N2pc Component

In ERP studies of visual search, the focusing of attention is associated with a component called N2pc (N2-posterior-contralateral), which occurs in the N2 latency range (200–300 ms) and has a posterior, contralateral scalp distribution. This component was first observed by Luck and Hillyard (1990), and it was also observed in two other studies published in the same year (Heinze et al., 1990; Luck et al., 1990). The evidence linking this component with the focusing of attention was provided by several subsequent studies (Luck et al., 1997b; Luck & Hillyard, 1994a, 1994b).

The typical paradigm for elicits an N2pc component is shown in figure 11.3A. Subjects are instructed to attend to one color at the beginning of each trial block, and for each stimulus they must indicate whether the item drawn in that color is an upright T or an inverted T. Two distinctively colored items are present in each array, one in each hemifield, and subjects attend to one of these colors on each trial block. This makes it possible to examine the ERPs elicited by identical stimuli but with differing directions of attention. When left-hemisphere ERP waveforms are examined, the voltage from approximately 200–300 ms is found to be more negative for right-hemifield (contralateral) targets than for left-hemifield (ipsilateral) targets (see figure 11.3B). When right-hemisphere ERP waveforms are examined, the voltage is more negative for left-hemifield (contralateral) targets than for right-hemifield (ipsilateral) targets (see figure 11.3C).

The stimuli and tasks used in N2pc experiments are usually somewhat different from those used in behavioral experiments. There are two reasons for this. First, in
Figure 11.3
Experimental paradigm (A) and results (B–C) from a typical N2pc component. Subjects are told to attend either to red or to green at the beginning of each block of trials, pressing one of two buttons on each trial to indicate whether the item drawn in the attended color is an upright T or an inverted T. The positions of the items are randomized from trial to trial, so the subjects do not know which side will contain the attended color prior to stimulus onset. ERPs are recorded from electrodes over left and right posterior cortical areas. For left-hemisphere electrode sites (B), the response is more negative going (i.e., less positive) when the target is in the right hemisphere than when the target is in the left hemisphere. For right-hemisphere electrode sites (C), the response is more negative going when the target is in the left hemisphere than when the target is in the right hemisphere.

ERP experiments it is important to avoid low-level sensory confounds by comparing the responses to physically identical stimuli and vary only the instructions (e.g., attend red vs. attend green). Second, in many behavioral search tasks, the target is difficult to distinguish from the distractors, leading to a great deal of variability in the amount of time required to focus attention onto the target. This latency variability can be problematic for ERP recordings, because the ERPs are time locked to the stimulus rather than to the moment at which attention is focused on the target. Latency variability results in a temporal smearing of the N2pc component, making it look like a low-amplitude, long-duration deflection (even though it was presumably a large-amplitude, short-duration deflection on the individual trials).
Consequently, most ERP studies of visual search use stimuli similar to those shown in figure 11.3A, which allow subjects to rapidly and reliably shift attention to the target location. However, it is possible to observe an N2pc component with the types of stimuli used in typical behavioral experiments (see, e.g., Luck & Hillyard, 1990).

Several sources of converging evidence support the hypothesis that the N2pc component reflects the focusing of attention. First, a large N2pc is observed for non-targets that require careful scrutiny to be distinguished from the target, but little or no N2pc is observed for non-targets that can be rejected on the basis of salient feature information (Luck & Hillyard, 1994a, 1994b). Second, the N2pc component is larger for targets defined by conjunctions of multiple features than for targets defined by the presence of a single feature (Luck et al., 1997b), which corresponds to the greater attentional requirements of conjunction-defined targets (Treisman & Gelade, 1980). Third, even though an N2pc component is often observed for targets defined by a single feature, the process reflected by the N2pc component is not necessary for accurate feature detection but appears to be necessary for accurate conjunction discrimination. Specifically, when an attention-demanding central task is used to discourage subjects from focusing attention onto a concurrently presented visual search target, the N2pc component is no longer present for feature targets but is still present for conjunction targets (Luck & Ford, 1998). Fourth, the N2pc is larger when distractors are presented close to the target than when distractors are more distant (Luck et al., 1997b), which corresponds to behavioral observations of greater attentional demands when distractors are presented close to the target (Cohen & Ivry, 1991). Fifth, the N2pc component is larger in tasks that require target localization (Hyun & Luck, submitted; Luck et al., 1997b), which corresponds with the hypothesis that target localization requires focused attention (Treisman & Gelade, 1980).

In addition to the N2pc’s being linked to attention in general, other evidence suggests a link to the more specific process of filtering out distractor stimuli so that they do not interfere with target identification. One piece of evidence for this more specific hypothesis is that the N2pc component is eliminated when no distractors are present (Luck & Hillyard, 1994b). Eimer (1996) demonstrated that an N2pc component is present if only one distractor is present, even when it is on the other side of the display from the target, and a similar result was obtained by Luck et al. (1997b). However, neurons at high levels of the visual system have very large receptive fields that would have encompassed both the target and the distractor in these experiments, so these results are not really evidence against the hypothesis that the N2pc component reflects a distractor-suppression process. Two additional findings provide support for this hypothesis by demonstrating that the N2pc component is eliminated if the distractors are not a real source of interference. First, if all the items in the search array are identical, the items do not conflict with each other and
no N2pc is observed (Luck & Hillyard, 1994b, experiment 4). Second, if the target is defined as the one item that differs from the other items in the array, such that the distractors are actually necessary for determining that a target is present, no N2pc component is observed. Specifically, the target could be either a single vertical item among several horizontal distractors or a horizontal item among several vertical distractors. Consequently, identifying the orientation of the discrepant item was not useful, but determining that this item actually differed from the surrounding items was necessary, so filtering the surrounding items would have been counterproductive. No N2pc was observed under these conditions (Luck & Hillyard, 1994b, experiment 2).

The N2pc component may be an ERP manifestation of attention effects that have been observed in single-unit recordings from extrastriate and inferotemporal areas of monkey visual cortex (Chelazzi et al., 1998; Chelazzi et al., 2001). When monkeys perform visual search, neurons in the object recognition pathway initially respond solely on the basis of the stimulus properties of the visual search array, but the presence of a target begins to influence activity around 175 ms poststimulus (which is quite close to the typical onset time of the N2pc component). For example, a neuron that is selective for red stimuli will initially respond if the search array contains a red item inside the neuron’s receptive field, whether or not the target is red. Beginning around 175 ms, however, the neuron’s response will fade away if the target for that trial is not red, but it will be sustained if the target is the red item. This single-unit effect is larger for complex target–nontarget discriminations than for simple discriminations, just as the N2pc component is larger for conjunction targets than for feature targets. The single-unit effect and the N2pc component are also both increased when a distractor item is located near the target, and both are increased when subjects must make an eye movement to a target rather than making a manual response (for details, see Luck et al., 1997b). In addition, an ERMF study demonstrated that the magnetic field distribution corresponding to the N2pc component is consistent with a generator source in ventral occipitotemporal cortex (Hopf et al., 2000).

11.12 Using Features to Guide Spatial Attention

How are spatial and nonspatial forms of attention related? As discussed earlier, the first nonspatial attention effects typically occur around 150 ms poststimulus, whereas the first spatial attention effects typically occur between 70 and 100 ms poststimulus. However, the early spatial attention effects are observed only when subjects know which location is to be attended prior to stimulus onset. In most visual search tasks, subjects know in advance what features are to be attended (i.e., the features
that define the target) but not the location of the target. It would therefore make sense that the feature-based effects that occur around 150 ms poststimulus reflect the processes that guide spatial attention to the target location around 200 ms poststimulus (as indexed by the N2pc component).

This possibility has been supported by single-unit and ERP/ERMF experiments. In a single-unit study, Motter (1994) found that color-selective neurons in extrastriate cortex responded differentially beginning at approximately 150 ms poststimulus, depending on whether the color of the stimulus inside the neuron's receptive field matched or mismatched the target color. Importantly, this effect occurred even when spatial attention was directed to a location outside the neuron's receptive field (see also Treue & Martinez Trujillo, 1999). A similar pattern of results has been observed in an ERP/ERMF study (Hopf et al., 2004). In this study, the presence of nontarget items containing target features led to a change in the neural response at approximately 140 ms poststimulus with an estimated neural origin in ventral occipitotemporal cortex, even when these features were at an ignored location. This was followed by an N2pc component at approximately 170 ms that was contralateral to the location of the actual target. Thus, feature-based attention precedes—and is presumably used to guide—spatial attention in visual search tasks.

11.13 Effects of Attention on Sensory Processing

Once attention is focused onto the target location, how is the processing of visual information changed at that location? To address this issue, Luck, Fan, and Hillyard (1993) flashed task-irrelevant probe stimuli at the location of the target or a distractor; the probe was presented 250 ms after the onset of the search array, during the period of the N2pc component. The ERP elicited by this probe stimulus was used to assess the status of processing at the target location relative to the distractor location at the time of the N2pc component. The probe-elicited P1 wave was found to be enhanced when the probe appeared at the location of the target compared to when it appeared at the location of the distractor, indicating that attending to the target causes a change in sensory responsiveness at the target location relative to the distractor location. Follow-up experiments (Luck & Hillyard, 1995) demonstrated that this effect was due to a suppression of responsiveness at the distractor location rather than an enhancement of responsiveness at the target location, consistent with the proposal that the N2pc component reflects a distractor-suppression process.
11.14 Serial Shifts of Attention

In the visual search experiments described so far, the target contained a distinctive feature and could therefore be localized very rapidly. In many behavioral visual search experiments—and in many real-world situations—the target is more difficult to localize and attention does not shift rapidly to the target. For decades, attention researchers have argued about the nature of the processing that occurs in these situations (for a review, see Chelazzi, 1999). One possibility is that attention shifts serially from one object to the next—identifying each attended object—until the target is found or the subject decides that no target is present. Another possibility is that all of the objects are attended and identified in parallel, but that this is a limited-capacity process that becomes slower as more objects must be processed. Behavioral experiments have provided evidence that parallel processing occurs under some conditions (Eckstein, 1998; McElree & Carrasco, 1999; Mordkoff et al., 1990), but this does not rule out the possibility that serial processing is used under other conditions.

It is extremely difficult to distinguish between serial processing and limited-capacity parallel processing on the basis of behavioral methods (Townsend, 1990). The problem is that behavioral measures cannot directly index the spatial focus of attention at each point in time following the presentation of a search array. In contrast, the contralateral distribution of the N2pc component makes it possible to use this component to assess which hemifield is attended, millisecond by millisecond, during visual search. That is, if attention shifts rapidly from an item in the left visual field to an item in the right visual field, then the N2pc component should shift rapidly from the right hemisphere to the left hemisphere. This is exactly what was found by Woodman and Luck (1999, 2003b).

To measure shifts of attention from one item to another using the N2pc component, it is necessary to know the order in which items are attended. Otherwise, trials with different patterns of attentional shifts will be averaged together and cancel out. To avoid this problem, Woodman and Luck (1999, 2003b) developed tasks that biased subjects to search the stimulus arrays in a known order. In one experiment, each display contained four distinctively colored items, one of which was the target on 75% of trials (called the $C_{75}$ color) and another of which was the target on 25% of trials (called the $C_{25}$ color). If search is serial, then subjects should attend to the $C_{75}$ color first and then shift attention to the $C_{25}$ color; consequently, the N2pc component should rapidly switch from the hemisphere contralateral to the $C_{75}$ color to the hemisphere contralateral to the $C_{25}$ color. However, if search is parallel, then subjects should simply devote more attention to the $C_{25}$ color and less attention to the $C_{25}$ color, with an overall larger N2pc contralateral to the $C_{75}$ color and no change over time. When the $C_{75}$ and $C_{25}$ colors were in opposite hemispheres, the N2pc
component was found to appear over the hemisphere contralateral to the C_{25} color from 200–300 ms poststimulus and then to shift to the hemisphere contralateral to the C_{25} color from 300–400 ms, consistent with a serial search process with a dwell time of approximately 100 ms.

In another set of experiments in this study, two potential target items of the same color were present in each array, but one was close to fixation and the other was more peripheral (size was scaled according to the cortical magnification factor to equate discriminability). Previous studies have shown that subjects tend to search near items before searching far items (Carrasco et al., 1995; Wolfe et al., 1998); accordingly, the N2pc component was observed contralateral to the near item from 200–300 ms and then shifted to the hemisphere contralateral to the far item from 300–400 ms. These experiments provide direct evidence that attention shifts serially from object to object under some conditions.

11.15 The Time Course of Attention in Visual Search

Taken together, the results of these visual search experiments indicate that visual search unfolds over time in the following manner. When a visual search array is presented, the first feedforward wave of activity in the visual system simply reflects the physical properties of the stimulus array. If an item containing target features is present, neural activity is increased beginning 120–150 ms after stimulus onset. This activity is used to guide spatial attention to the location of the potential target item, which is reflected by the N2pc component (and changes in single-unit activity) beginning 175–200 ms poststimulus in occipitotemporal areas of visual cortex. This makes it possible to determine whether the attended item is actually the target, which will trigger a motor response. If the attended item is not a target, then attention may be shifted to another potential target item after a delay. The amount of time from the onset of attention to the first item and the onset of attention to the second item presumably depends on task difficulty and was found to be approximately 100 ms for a fairly difficult target–nontarget discrimination. Presumably, attention keeps on shifting from item to item until the target is found or the subject decides that no target is present.

11.16 Summary and Conclusions

This chapter has summarized the time course of attentional selection over the first half second. ERP studies—in combination with ERMF, neuroimaging, and single-unit studies—have revealed the following picture of the timing of attentional processing.
When attention is directed to a location before stimulus onset, spatial attention influences the feedforward transmission of sensory information through extrastriate cortex beginning 60–100 ms poststimulus. In contrast, attending to nonspatial features does not appear to influence feedforward sensory transmission in this manner; instead, feature-based attention begins to influence activity beginning approximately 150 ms poststimulus (for a potential counterexample, however, see Valdes-Sosa et al., 1998). But when the target location is not known in advance (as in many real-world situations), these feature-based attention mechanisms are used to guide spatial attention to objects that contain relevant features. When the target contains a salient feature, the salient feature is detected within 150 ms poststimulus and spatial attention is directed to its location 25–50 ms later. When the target does not contain a unique and distinctive feature, attention may be shifted to one or more nontargets before it shifts to the target. Under such conditions, attention sometimes operates serially, focusing on one object for approximately 100 ms before shifting to another object (although the dwell time may depend on stimulus discriminability).