CHAPTER 1

# ERP Components: The Ups and Downs of Brainwave Recordings

Emily S. Kappenman *and* Steven J. Luck

**Abstract**

This chapter provides a framework for understanding, interpreting, and using event-related potential (ERP) components in the broad domain of mind, brain, and behavior sciences. The first section defines the term *ERP component*, describing the neural events that give rise to ERP components and explaining how multiple components sum together to form the observed ERP waveform. The next section describes the problems involved in isolating individual ERP components from the observed waveform, which is often much more difficult than researchers realize. This is followed by a discussion of the challenges involved in linking an ERP component with a specific neural or psychological process and then using this link to answer broader questions about the mind and brain. The chapter concludes with a discussion of what types of questions are most easily answered with ERPs and the approaches that have proven effective in overcoming the challenges of the technique.

**Keywords:** event-related potential, ERP component, peaks, waves, reverse inference

The goal of this chapter is to provide a framework for understanding, interpreting, and using event-related potential (ERP) components in the broad domain of mind, brain, and behavior sciences. Researchers in other areas such as political science, economics, law, and medicine may also find this overview useful as a guide to a broad understanding of ERP components. Event-related potentials have been used for decades to uncover aspects of the sensory, cognitive, and motor processes that underlie human thought and behavior. The excellent temporal resolution of the technique provides a narration of neural processes as they unfold millisecond by millisecond, adding whole pages to the story of the mind that behavioral and imaging techniques leave blank. However, the ERP technique is not without limitations. As reflected in the title of this chapter, there are both advantages and limitations of the ERP technique, and we will explore both the ups and the downs of ERPs in this chapter.

The first section of the chapter is aimed at defining the term *ERP component*, describing the neural events that give rise to ERP components and explaining how multiple components sum together to form the observed ERP waveform. The next section describes the problems involved in isolating individual ERP components from the observed waveform, which is often much more difficult than researchers realize. This is followed by a discussion of the challenges involved in linking an ERP component with a specific neural or psychological process and then using this link to answer broader questions about the mind and brain. These challenges may seem insurmountable, but researchers have developed experimental and analytic approaches that can overcome them in many cases. The key to using ERPs effectively is to understand what questions can be answered by ERP experiments and how the limitations of the technique can be avoided. Indeed, despite its limitations, the ERP technique is often

the best one for answering certain types of questions. The chapter therefore ends with a discussion of what types of questions are most easily answered with ERPs and the approaches that have proven effective in overcoming the challenges of the technique.

Although a number of the issues we address are discussed elsewhere in the literature (e.g., see Luck, 2005), this chapter provides a comprehensive and concise overview of the nature and use of ERP components from a vantage point that is readily accessible to researchers from a wide range of backgrounds. Readers who have no familiarity at all with the ERP technique may wish to first read the more basic introduction provided by Luck (in press).

## The Nature of ERP Components

### *What Is an ERP Component?*

The ERP waveform appears on the scalp as a series of positive and negative peaks[1] that vary in polarity, amplitude, and duration as the waveform unfolds over time. However, the actual waveform is continuous, with no sudden transitions between one peak and the next, and division of the ERP waveform into discrete peaks is somewhat arbitrary. Indeed, this peak-centered view of the ERP waveform may reflect an intrinsic predisposition of the human visual system to use *minima of curvature* (places where orientation reverses direction) to define the parts of complex real-world objects (Hoffman & Richards, 1984). Although the peaks are visually salient, there is no a priori reason to believe that each peak reflects a specific brain process. However, early ERP researchers tended to make this assumption, and this has had a major influence on the terminology and analytical techniques used in ERP research. Sophisticated ERP researchers have recognized for decades that the peaks are somewhat arbitrary, and they make a distinction between *peaks* (local voltage maxima) and *components* (discrete intracranial sources of voltage that reflect specific neurocognitive processes, defined further below). Nonetheless, it is still common for researchers to assume that a peak in the observed ERP waveform is equal (or approximately equal) to an underlying ERP component. Perhaps the most important goal of this chapter will be to encourage readers to look beyond the visually salient peaks to the underlying components; it is the underlying components rather than the peaks that directly reflect the neural and psychological processes we wish to study.

To clarify the relationships among peaks and components, it is important to begin with some clear definitions. We can define the observed ERP waveform as *a depiction of the changes in scalp-recorded voltage over time that reflect the sensory, cognitive, affective, and motor processes elicited by a stimulus*. We can define an ERP peak as *a reliable local positive or negative maximum in the observed ERP waveform* (the term *reliable* allows us to disregard local maxima that result from high-frequency noise).

The term *ERP component* is more challenging to define. This term gets bandied about in the literature very frequently, but it is rarely defined or conceptualized beyond the peaks in the observed ERP waveform. In some sense, the term *ERP component* is analogous to the concept of attention: Just as "everyone knows what attention is" (James, 1890, p. 381), everyone knows what an ERP component is (at least everyone in the ERP world). Moreover, despite the fact that attention researchers all believe they know what attention is, they vary substantially in how they use the term *attention* (Luck & Vecera, 2002), and ERP researchers similarly vary in how they use the term *component*. Therefore, just as it is difficult to elicit agreement on the term *attention* in a room full of attention experts, it is no easy task to find a simple, concise, and widely accepted definition of the term *ERP component*. Furthermore, there is an important distinction between how these terms have evolved: although attention researchers frequently debate the fundamental nature of attention, ERP researchers rarely discuss the nature of ERP components.

There are, of course, counterexamples to this sweeping generalization about the nature of ERP components. For example, Manny Donchin has written extensively and explicitly throughout his career about ERP components and their existence beyond the peaks in the observed ERP waveform (e.g., see Donchin & Heffley, 1978). More recently, Luck (2005) provided a comprehensive discussion of the distinction between components and peaks. The concept of a component has also been discussed in the context of mathematical techniques for isolating components, such as principal component analysis (Donchin & Heffley, 1978) and independent component analysis (see Chapter 3, this volume). However, this important issue is often ignored in the ERP literature and warrants continued discussion.

In a general sense, we can define the term *ERP component* as *a scalp-recorded voltage change that reflects a specific neural or psychological process*. Although most researchers understand and use words such as *reflect* and *process*, such terms themselves refer to loose concepts without clear definitions. Consequently, it will be necessary to fill out

the details of this definition over the course of this chapter. However, this concise definition does provide a reasonable approximation of the way the term *ERP component* is usually used by ERP researchers. We will illustrate the relationship between the ERP waveform and the underlying ERP components in the following sections, first discussing the neural events that give rise to the observed ERP waveform and the process of isolating the ERP waveform from other electrical activity. We will then illustrate the differences between the peaks in the ERP waveform and the underlying ERP components through the use of simulated waveforms.

### *Where Do ERP Components Come From?*

Event-related potentials are voltage fluctuations in the ongoing electroencephalogram (EEG) that are time-locked to an event, such as the onset of a stimulus or the execution of a manual response. Electroencephalographic research began long before laboratory computers were available, and early researchers were able to observe only large ERPs that were visible on single trials (Davis, 1939) prior to the advent of computer averaging in the early 1960s (Galambos & Sheatz, 1962). However, most ERPs are rather small in comparison with the ongoing EEG activity and usually become visible only when multiple EEG epochs are combined together to form an average ERP waveform. This averaging process proved extremely beneficial to the field of ERPs and was the first occurrence in which signal averaging "revealed the existence of novel, previously unknown, neural processes" (Donchin et al., 1978, p. 349).

To understand the intricate mixture of signals we record on the surface of the scalp, we must first understand where and how these signals arise neurally. Although it is difficult to know with certainty how scalp-recorded voltage changes originate at the neural level, the following represents the best estimate based on our understanding of both biophysics and the properties of neural communication.

The changes in scalp-recorded voltage that give rise to the ERP waveform reflect the summation of postsynaptic potentials (PSPs) that occur simultaneously in large numbers of cortical pyramidal cells that are orientated in a similar manner with respect to the scalp (see Luck, 2005, chap. 1). These PSPs are a result of changes in electrical potential that occur when ion channels open or close in response to neurotransmitters binding with receptors on the postsynaptic cell membrane, which leads to the flow of ions into or out of the cell. When a PSP occurs at one end of a cortical pyramidal neuron, the result can be considered an electrical *dipole*, with positive on one end and negative on the other end. When PSPs occur simultaneously in many neurons that are spatially aligned, such that their dipoles point in the same direction, the dipoles sum together to form a large dipole known as an *equivalent current dipole*. If a sufficiently large number of spatially aligned neurons are simultaneously active, the equivalent current dipole is large enough to be reliably recorded on the surface of the scalp. This requires the simultaneous activation of thousands of neurons, due in part to the many layers of tissue that separate the scalp electrodes from the neurons. This is most likely to occur in groups of pyramidal cells in cerebral cortex, which are lined up together perpendicularly with respect to the cortical surface and are often active in unison. In other words, ERPs are almost always the result of PSPs in large groups of cortical pyramidal cells. It should be noted that, except in a few unusual cases, scalp ERPs do not reflect action potentials. Thus, ERPs represent the inputs to a group of neurons rather than the outputs of those neurons. Also, due to the necessity for such large numbers of spatially aligned neurons to be simultaneously active in scalp recordings, much of the neural activity in the brain that gives rise to cognition and behavior is not visible to an electrode placed on the scalp.[2]

For a given equivalent current dipole or neural generator source, the specific distribution of positive and negative voltages recorded on the scalp is determined by the position of the dipole in the head and its orientation with respect to the scalp (although it should be noted that the choice of reference electrode can also play a factor in the voltage distribution; see Luck, 2005, chap. 3). In other words, each equivalent current dipole will produce both positive and negative voltages on the head, with a band of zero separating the positive and negative voltage halves. This voltage reversal on the opposite side of the equivalent current dipole is often not very noticeable, because electrodes are not generally placed over the entire head, but the reversal is easily observed for some components (such as the N170; see Chapter 5, this volume). The positive or negative polarity of an ERP component at a given electrode site is related to several factors, including the orientation of the equivalent current dipole with respect to the electrode, and it is not usually possible to link the polarity to the type of neural processing (such as inhibition versus excitation). For a more detailed discussion of the factors that affect the polarity of an ERP, see Luck (2005, chap. 1).

Because electrical potential travels close to the speed of light, the transmission through the brain, meninges, skull, and scalp is essentially instantaneous. In other words, the voltages measured on the scalp at a particular time reflect synaptic activity at that particular instant, with no measurable delay. Thus, ERPs provide a direct and instantaneous millisecond-resolution measure of activity related to neurotransmission.

### *Summation of Components in the Observed ERP Waveform*

It is important to note that although the ERP waveform at a particular instant reflects synaptic activity at that moment, it does not reflect *only* the neural activity that *began* at that particular instant. Specifically, the PSPs that give rise to ERPs last on the order of tens or even hundreds of milliseconds.[3] Therefore, as new mental processes are unfolding, the previous neural activations persist. In other words, multiple groups of neurons are active simultaneously in different regions in the brain. If we think of this neural activity in terms of dipoles, this means that multiple equivalent current dipoles are active simultaneously. In fact, source localization studies have shown that as many as 10 separate equivalent current dipoles may be active at a given time (Di Russo et al., 2002; Picton et al., 1999). If we return to our conception of ERP components, in which we define an ERP component as a signature of an individual neural process, each equivalent current dipole is essentially a separate ERP component. In other words, when we say that multiple equivalent current dipoles are active simultaneously, this really means that multiple ERP components are generated simultaneously.

In some cases, neurons engaged in one mental process may be distributed in different areas of the brain, such as the simultaneous processing of a single auditory signal in both the left and right temporal lobes. This would essentially lead to two equivalent current dipoles. Should we consider these two dipoles as two separate ERP components or as a single ERP component? They are typically treated as parts of a single component under the assumption that both hemispheres are engaging in essentially the same mental process. However, this is a fine detail of the definition of an ERP component, with little practical significance for the use of ERP components. Furthermore, resolution of this issue would require a precise definition of what is meant by *mental process* in terms of the behavior of neurons, both individually and as a group. That is, how do we determine whether the same mental process is occurring in two individual neurons, and on a larger scale, in groups of neurons? This is a complex issue that remains to be resolved by future research.

The combination of multiple ERP components on the scalp leads to the *superposition problem*, which is depicted in Figure 1.1. When multiple ERP components are simultaneously active, the recorded voltage at the scalp is based on the sum of the voltages from all the individual components. This is a simple additive process. That is, if you knew the true waveform for each individual component, you could add all the component waveforms together to get the ERP waveform at each electrode site (scaling each component by a weighting factor that reflects the contribution of the component to the voltage measured at a specific electrode site). Unfortunately, the true waveform for each component is not known in real recordings, and it is quite difficult to reduce the sum of the components in the observed data to the individual components. However, understanding with simulated data how the voltage recorded at a particular electrode site reflects the various internal generator sources can help us understand the properties and intricacies of the ERP signals.

The propagation of voltage from a single generator site to a particular electrode site depends on the position and orientation of the ERP generator source with respect to the electrode, along with the conductance of the brain, skull, and scalp. This can be quantified with a weighting factor: The contribution of a given generator to the voltage recorded from a given electrode site is simply the waveform at the generator multiplied by the weighting factor (see Figure 1.1). There will be a separate weighting factor specifying the relationship between each electrode site and each internal neural generator source. Together, the set of weighting values between each source and each electrode site provides a *mixing matrix* that defines how the different components mix together at each site. Some mathematical techniques for recovering the underlying components work by computing an *unmixing matrix* that reverses this process, passing the observed data through the unmixing matrix to compute the component waveforms (see Chapter 3, this volume).

When multiple ERP components are simultaneously generated in different brain areas, the voltages from these components sum together. The voltage recorded at each site will therefore be the sum of each of the internally generated ERP components, with each scaled by the weight between that electrode site and each of the generator locations. The value at

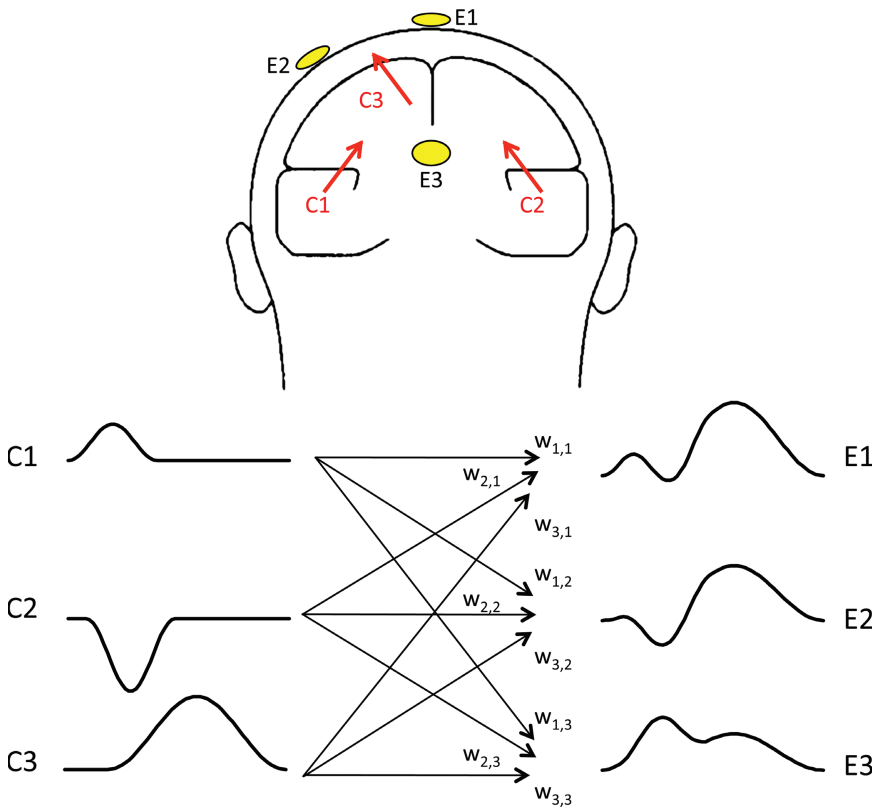


**Fig. 1.1.** Relation between the underlying component waveforms and the observed scalp waveforms. In this example, three components are present (C1, C2, C3), each of which has a waveform (shown at the bottom left) and a generator location (represented by the arrows in the head). The contribution of each component waveform to the observed waveform at a given electrode site is determined by a weighting factor that reflects the location and orientation of the generator relative to that electrode, along with the conductivity of the tissues that form the head. The observed waveform at a given electrode site (shown at the bottom right) is equal to the sum of each of the component waveforms, multiplied by the weighting factor between each component and that electrode site. The weights are indicated by the *w*'s on the arrows between the component waveforms and the observed waveforms (e.g., $w_{2,3}$ represents the weighting factor between component 2 and electrode 3).

a given electrode site at a particular moment in time is equivalent to the magnitude of each component at that time, scaled by the appropriate weighting factor and then summed together. Consequently, the ERP waveform at each electrode site contains information about all of the neural generators in the brain, not just the generator sources located close to the electrode (although nearby sources will usually have a greater weight).

The inability to relate the ERP waveform at a particular electrode site to the neural tissue directly below the electrode site is made even more severe by the properties of the head. Specifically, as electrical activity travels from the brain to the surface of the scalp, the activity must pass through layers of skull and scalp. Although these constituents of the head are sufficiently conductive to allow the electrical activity generated in the brain to appear on the surface of the head, they are not perfect conductors, and the high resistance of the skull relative to the low resistance of the underlying brain and overlying scalp causes the voltage to spread laterally as it travels. The signals are therefore blurred together by the head, which further distorts the relationship between the voltage at a particular electrode site and the cortex directly under that site.

Of course, anyone who has seen the ERP waveforms from multiple electrode sites knows that differences exist in the shape and size of the ERP waveform across electrode sites. In other words, although the waveform at each electrode site reflects neural signals from all over the brain, the summated signals are not identical at each site. It is tempting to use the scalp distribution information to estimate the location of the neural generator source by, for example, determining at which electrode site the signal is largest. However, the superposition of multiple components and the blurring of the voltages

across the head make it impossible to determine the locations of the generator sources solely from the observed waveforms. In fact, an infinite number of internal generator configurations could produce any observed distribution of ERP activity over the scalp (see Luck, 2005, chap. 7). Thus, there is no technique that can determine, with certainty, the locations of the sources and the waveform at each source without bringing in difficult-to-verify assumptions or other sources of evidence.

To summarize, the ERP waveform reflects ongoing synaptic activity related to mental processing as it unfolds millisecond by millisecond. However, because scalp-recorded signals require the simultaneous activation of large groups of spatially similar oriented neurons, only a portion of the neural activity that occurs in response to a stimulus will be measurable from electrodes on the surface of the scalp. Furthermore, the ERP waveform at a given electrode site reflects the contribution of many simultaneously active ERP components that overlap in time, and it is difficult to mathematically unmix the observed waveforms and determine the original component waveforms.

### *Other Approaches to Defining ERP Components*

In this section, we will consider the relationship between the definition of the term *ERP component* that we have proposed in this chapter and the way that components are defined by four other approaches: *source localization*, *principal component analysis* (PCA; see Donchin & Heffley, 1978), *independent component analysis* (ICA; see Chapter 3, this volume), and *time-frequency analysis* (see Chapter 2, this volume). We will concentrate on the spatial variants of PCA and ICA, in which components are defined on the basis of scalp distribution information (see Spencer et al., 2001, for a discussion of temporal and spatiotemporal PCA).

We will begin by considering the source localization, ICA, and PCA approaches. In these three approaches, a component is defined solely by its scalp distribution, which is assumed to remain stable over the course of a single experimental session (this is a reasonable assumption given that brain geometry is unlikely to undergo major changes within a few hours). As mentioned in the previous section, these techniques provide an *unmixing matrix* that reflects the estimated scalp distributions of the individual components; the waveform for each component is computed by passing the observed waveforms through this matrix. That is, rather than passing the component waveforms through the weights shown in Figure 1.1 to obtain the observed waveforms at each electrode (moving from left to right in the figure), these techniques pass the waveforms observed at each electrode site through an unmixing matrix to obtain the component waveforms (moving from right to left). Unfortunately, there is no unique solution to the problem of determining the underlying component waveforms from the observed scalp waveforms, and these three techniques use different assumptions to pick a single solution to this problem (without any guarantee that the correct solution will be found).

In source localization techniques, a component is equivalent to a neural generator source. These techniques use biophysical assumptions about the flow of current through the conductive tissues of the head to define the scalp distribution of each component (and thereby compute a unique unmixing matrix). To obtain a unique solution, these techniques must also rely on additional assumptions, such as a specific number of discrete dipoles or maximal smoothness in the distribution of current flow over the cortical surface. That is, these techniques find the set of single-component scalp distributions that can sum together to provide the best fit to the observed scalp distribution as it varies over time while also being consistent with a variety of assumptions (for a review and critique, see Luck, 2005, chap. 7). Thus, source localization techniques define a component as activity arising from a region of cortex, which is similar to our definition of an ERP component as reflecting a specific brain process (on the assumption that most brain processes occur in discrete areas[4]). However, our definition of the term *ERP component* goes further, because more than one brain process may occur in a given region of cortex. Moreover, source localization approaches differ considerably from the traditional approach to defining components in the procedures used to discover and define individual components. Whereas source localization techniques use a variety of assumptions to select a set of scalp distributions that together provide a quantitative account of the data from a given experiment, traditional approaches to defining components are based on using experimental manipulations to test hypotheses about the link between a voltage deflection and an underlying neural or psychological process (as discussed further in a later section).

Principal component analysis and ICA make no biophysical assumptions, but instead use the statistical properties of the data to derive the scalp

distributions of the components. That is, the observed scalp distribution changes from moment to moment and from condition to condition as the underlying components wax and wane, and the statistical relationships between the values observed at the different electrode sites are used to determine the scalp distributions of the individual components. In PCA, for example, two electrode sites will tend to contribute strongly to the same component if they tend to covary in voltage. Principal component analysis is designed to find an unmixing matrix in which a small number of components—each with its own scalp distribution—can sum together to explain most of the variations in the observed scalp distribution. It reduces a large and complex set of observed scalp distributions (for each time point, condition, etc.) to a small number of component scalp distributions. In contrast, ICA is designed to find an unmixing matrix that maximizes the independence of each component so that every individual component represents the largest possible amount of information. The scalp distributions of the components in ICA may be correlated with each other (as would be expected for two independent but nearby neural sources), but the strength of activation of each component varies independently of the strength of the other components over time points and over conditions. Whereas PCA attempts to lump as much information as possible into a small number of components, ICA attempts to split apart the information into different components (for a detailed comparison, see Chapter 3, this volume).

Because it is a "lumping" technique, spatial PCA by itself is unlikely to produce components that are related to individual neural and psychological processes. However, the essence of ICA corresponds well with a reasonable assumption about these processes. Specifically, for something to count as a unique process, it must be dissociable from other processes. This is largely identical to saying that the process must sometimes vary independently of other processes, and this is exactly the type of independence that ICA uses to define components. Thus, although ICA uses a mathematical approach rather than a hypothesis-testing approach to derive the components, it shares much with the definition of the term *ERP component* that we have proposed in this chapter. Moreover, the components isolated by ICA often have a scalp distribution that matches what would be expected for a single dipole, even though the technique makes no biophysical assumptions about dipoles (see, e.g., Figure 3.9 in Chapter 3, this volume).

There are, however, some practical problems associated with linking ICA components to ERP components as we have defined them here. First, ICA is applied to single-subject data, and it can be difficult to determine the correspondence between the ICA components obtained for the different subjects. The same problem arises when comparing components across experiments. Second, the ICA computational approach requires that the number of ICA components is always equal to the number of electrodes, and this means that multiple true components may be lumped together into a single ICA component or that a single true component may be distributed across multiple ICA components. It remains to be seen how well the traditional approach and the ICA approach to defining and isolating components can be combined.

The time-frequency approach is very different from the source localization, ICA, and PCA approaches (although it can be combined with them). In the time-frequency approach, the EEG is decomposed into the sum of a set of oscillations, and the power in each frequency band is estimated at each moment in time (with varying degrees of temporal precision; see Chapter 2, this volume, for details). The results of this approach can be related in conventional ERP components in two main ways.

First, if the oscillations vary randomly in phase from trial to trial, they will ordinarily disappear when the single-trial EEG epochs are averaged together; oscillations of this sort are completely invisible in conventional averaged ERP waveforms (for an exception, see Mazaheri & Jensen, 2008). In such cases, oscillations within a given frequency band are often considered as being analogous to ERP components, reflecting a specific neural or psychological process. However, many different processes might lead to oscillations in a given frequency band, so it is problematic to assume that power in a given frequency band in one experiment reflects the same process reflected by power in that same frequency band in another experiment (e.g., theta-band activity in one experiment may reflect very different processes than theta-band activity in a different experiment). Assuming that a given band reflects a specific process would be analogous to assuming that any positive deflection in the P3 latency range reflects a single process.

A second possibility is that a stimulus might perturb the phase of an ongoing oscillation, causing the phase to become consistent across trials during the period immediately after the stimulus. In such cases, the phase consistency across trials will allow

the oscillation to survive the averaging process (see Figure 2.2 in Chapter 2, this volume). When this happens, a component in an averaged ERP may actually consist of a portion of an ongoing oscillation rather than reflecting a discrete voltage deflection that is elicited by the stimulus.

## Challenges in Isolating ERP Components

We have defined the term *ERP component* as scalp-recorded neural activity that is associated with a particular neural or psychological process. It is the nature of the underlying process that we are seeking to uncover with ERP research; however, as discussed in the preceding section, the ERP waveform that we can record contains a mixture of many different ERP components. Deconstructing the ERP waveform into its ERP components is no trivial task. An infinite number of combinations of underlying components could sum together to give rise to a given ERP waveform. This section is devoted to illustrating the difficulty in assessing changes in a component from the observable ERP waveform. To illustrate these points, we will use simulated data for which the underlying ERP components are known and modifiable. This section is primarily aimed at pointing out the limitations of ERP component research. Although this section may make ERP research seem dismal, you should not become disheartened with ERPs. The final section of this chapter will provide some tools that have been successful for using ERPs to answer questions about the mind, brain, and behavior.

### *ERP Peaks ≠ ERP Components*

As discussed earlier, the ERP waveform looks like a succession of distinct and easily separable peaks, but these peaks do not map onto distinct ERP components in a simple one-to-one manner. The neural activation associated with each distinct mental process persists for tens or hundreds of milliseconds, which means that the ERP signature from one process will overlap with the ERP signature for subsequent processes either in part or in whole. Even if these neural processes occur in separate parts of the brain, the ERP waveform at a given electrode site will be the weighted sum of all of the underlying components. In other words, each peak in the waveform is usually determined by more than one, and often several, separate ERP components.

Much ERP research has centered on evaluating differences in the size or timing of an ERP component across conditions or across groups of subjects. Such changes can speak volumes about differences in neural processing. However, the problem of overlapping components makes it difficult to ascertain whether a change in a peak in the observed ERP waveform is due to a change in one component, a change in a different component, or changes in a combination of multiple components. In the language literature, for example, it is not always clear whether a putative increase in N400 amplitude might actually be a decrease in P3 amplitude, and a great deal of work was needed to determine that the P600 component elicited by syntactic anomalies was different from the P3 wave (see Chapter 15, this volume).

Figure 1.2 illustrates some of the measurement problems that arise due to the overlap of ERP components. In this simulated example, the observed waveform shown in Figure 1.2A is the sum of the three underlying components shown in Figure 1.2B. In other words, Figure 1.2B is the observed ERP waveform and Figure 1.2A shows the underlying components (which we cannot observe directly in real experiments). Looking at the observed waveform, the ERP appears to consist of a positive component from 0 to 90 ms, a negative component from 90 to 180 ms, and a positive component from 180 to 450 ms. However, the underlying components are much longer in duration, with the first positive component active from 0 to 200 ms, the negative component active from 50 to 325 ms, and the second positive component active from 100 to 450 ms. Thus, one cannot easily determine the duration of an underlying component from the duration of the peak in the observed waveform. The difficulty of assessing component duration from the ERP waveform is a problem in experimental contexts as well, particularly when a smaller component is preceded or followed by a much larger component. For example, it is difficult to assess the duration of the N2 component when it is followed closely by the much larger P3 component. Although it is often the case that evaluating the length of a peak in the waveform minimizes the apparent duration of a component, the waveform can also make a component seem longer in duration than it is in actuality. For example, the late positive potential (LPP) in the emotion literature appears as a single component that is hundreds of milliseconds in duration; however, the LPP may actually be composed of several distinct shorter-duration components (see Chapter 16, this volume). Therefore, the duration of peaks in the ERP waveform is often quite different from the duration of the underlying components.

Changes in the timing or size of components across experimental conditions or groups of subjects

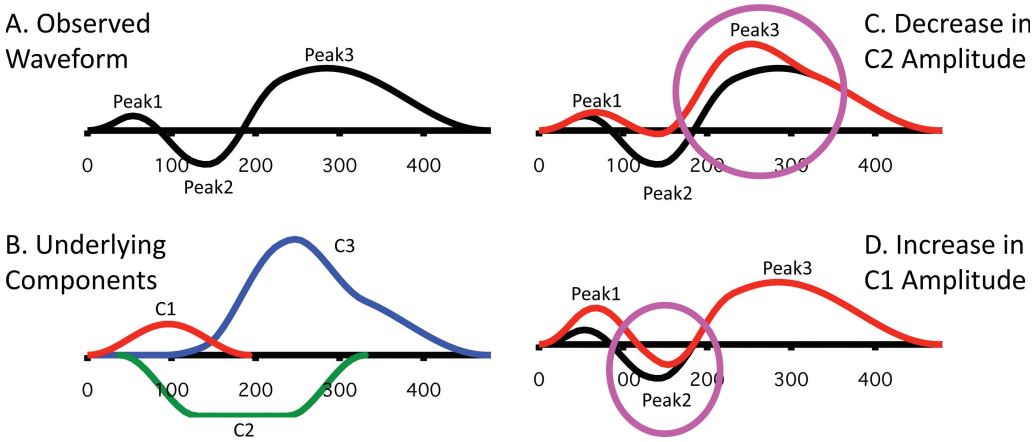


**Fig. 1.2.** Example of how the peaks in an observed waveform can misrepresent the underlying components. Panel A shows the observed waveform, and Panel B shows the underlying components that sum together to produce the observed waveform. Note that Peak 1 is much earlier than the peak of component C1, and the shape of Peak 2 is very different from the shape of component C2. Panel C shows the original waveform overlaid with a waveform in which the amplitude of component C2 has been decreased. Note that this change in C2 causes an increase in the amplitude of Peak 3, even though component C3 does not differ between these waveforms. Panel D shows the original waveform overlaid with a waveform in which the amplitude of component C1 has been increased. Note that this changes the amplitude and latency of Peak 2, even though component C2 does not differ between these waveforms.

can also be difficult to assess from the ERP waveform. Figure 1.2C shows the effect of an experimental manipulation that decreases the amplitude of the negative component. In addition to decreasing the measured amplitude of the negative peak in the observed waveform, this manipulation greatly increases the amplitude of the second positive peak (even though the manipulation did not change the amplitude of the second positive component). This is one clear example of how changes in the amplitude of one component (the negative component) can result in an amplitude change in a subsequent part of the waveform (the second positive peak). Based on a superficial evaluation of the waveform, these changes would lead to the erroneous conclusion that the difference between conditions was the result of modulations in two underlying ERP components; however, in this case, both peak modulations were caused by a change in a single underlying ERP component. Therefore, researchers may draw substantially incorrect conclusions if they assume that a change in the size of a peak reflects a change in the size of a particular component.

Similarly, Figure 1.2D shows the effect of a manipulation that increases the amplitude of the first positive component. In addition to increasing the measured amplitude of the first positive peak in the observed waveform, this manipulation decreased the measured amplitude of the negative peak. The manipulation of the amplitude of the first positive component also increased the apparent latency of the negative peak, even though no latency shift occurred for any of the underlying components. In other words, a change in the amplitude of one component can in some cases masquerade as a shift in the latency of a different component. Therefore, it is often difficult to determine whether a specific type of modulation of the ERP waveform is related to the same type of change in the underlying components. In other words, measured shifts in peak latency can sometimes be caused merely by changes in component amplitude, and measured changes in peak amplitude can sometimes result from shifts in component latency.

Although we have shown a few cases of the difficulty in linking changes in the ERP waveform with changes in particular underlying ERP components, this is by no means an exhaustive description of the ways in which changes in underlying components can affect the observed ERP waveform. We encourage anyone interested in exploring these effects to create simulated data and see how modulations in the underlying components affect various parts of the ERP waveform (this is easy to do in a spreadsheet program, such as Excel). Furthermore, it should be noted that the simulation shown in Figure 1.2 may actually underestimate the severity of the problem of measuring amplitudes and latencies from the ERP waveform, because modeling efforts suggest that 6–10 generators may be active

within a given 150 ms period (Di Russo et al., 2002; Picton et al., 1999), in contrast to the 3 neural generators used in the simulation shown in Figure 1.2. On the other hand, considerable information about the underlying component structure can often be obtained by examining the waveforms from multiple electrode sites, because different components will be weighted differently at each electrode.

### *Variability in ERPs*

Amplitudes and latencies are almost always measured from the average of multiple EEG segments but separately for each individual subject. In other words, all of the trials in a condition are averaged together for a given subject, and the amplitude and latency measures are computed for each subject from this average waveform. Each subject then contributes a value to the statistical test for differences across conditions or groups, with the variance across subjects contributing to the ability to detect a significant experimental effect. This process of signal averaging is incredibly important and integral to the utilization of ERPs; averaging across multiple EEG epochs reveals ERPs that are not visible on single trials, and data from multiple subjects provide a measure of variance that is important to assessing statistically significant changes. However, it is important to understand distortions that can be introduced by the averaging process.

The process of averaging across multiple trials to form an average ERP waveform relies on several assumptions, the most important of which is that the timing of the signal of interest is the same on each trial. However, this is often not the case. Specifically, just as the behavioral reaction time varies substantially from trial to trial in an experiment, the timing of the underlying neural processes that give rise to the ERP components may also vary from trial to trial. The variability in the timing of a component across trials is known as *latency jitter*, and it can actually be quite problematic to the interpretation of an average waveform. When latency jitter is present for a component, as depicted in Figure 1.3, the average ERP waveform will contain a "smeared-out" version of the component. Specifically, the average ERP waveform will reflect both the *earliest* onset and *latest* offset times of the component, as opposed to reflecting the *average* onset and offset times. In addition, latency jitter can greatly reduce the measured peak amplitude (discussed more fully later in the chapter). Furthermore, although this variation in timing across trials is informative about the nature of the process reflected by the component, it can make the comparison of the size and timing of a component across conditions or across groups of subjects more difficult. Specifically, greater variability in the timing of a component across conditions may be incorrectly interpreted as a change in the size or duration of the component. For example, a comparison of the two conditions depicted in Figures 1.3A and 1.3B might lead to the erroneous conclusion of a smaller component in condition A than in condition B, even though the only difference between the conditions lies in the variability in the component timing across trials. Therefore, understanding how latency jitter

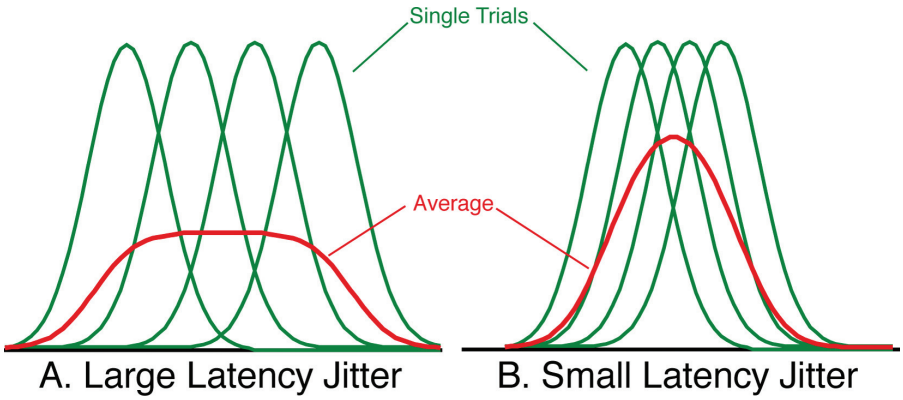


**Fig. 1.3.** Example of how differences in latency jitter (the amount of variability in component latency across trials) influence the average waveform. The green waveforms are the single-trial data, and the red waveforms are the averages across trials. The jitter in single-trial latency is greater in (A) than in (B), leading to a broader averaged waveform with a lower peak amplitude in A than in (B). That is, even though the amplitude of the single-trial waveforms is equivalent in (A) and (B), the peak amplitude of the averages differs between (A) and (B). In addition, the onset time and offset time of each average reflect the earliest onset times and latest offset times of the single trials rather than the average of the single-trial onset and offset times.

can impact the average waveform can be useful in interpreting experimental effects.

Measures of amplitude and latency are almost always taken from individual subject waveforms. By contrast, most ERP papers show the grand average ERP waveform across subjects, as opposed to each of the individual subject waveforms. Therefore, the characterizations we can make of the components in a particular experiment are generally taken from an average representation of all the subject waveforms in the study. It is tempting to think that the grand average ERP waveform would reflect the average of all of the individual subject waveforms that make up the average; however, just as the average of multiple EEG segments within a subject reflects the range of the epochs, a grand average across subjects actually reflects the *earliest* onset and *latest* offset times and not the average of the onsets and offsets of the components. In other words, if there is substantial variability in the timing of the components across subjects, the grand average ERP will reflect that variability.

One of the most salient factors when measuring the amplitudes and latencies of ERP components from the individual subject waveforms is the quite substantial variation in shape across waveforms. For example, consider the waveforms in Figure 1.4A. The bottom waveform is the grand average across subjects, and the other waveforms reflect 8 randomly selected subjects from the 20 individuals who contributed to the grand average. The highlighted portion of the figure corresponds to the time period one might select to measure the P2 wave, because it covers almost the entire duration of the wave in the grand average. However, the activity within this time window varies considerably across the individual subject waveforms. For some of the subjects, the first positive wave peaks prior to the

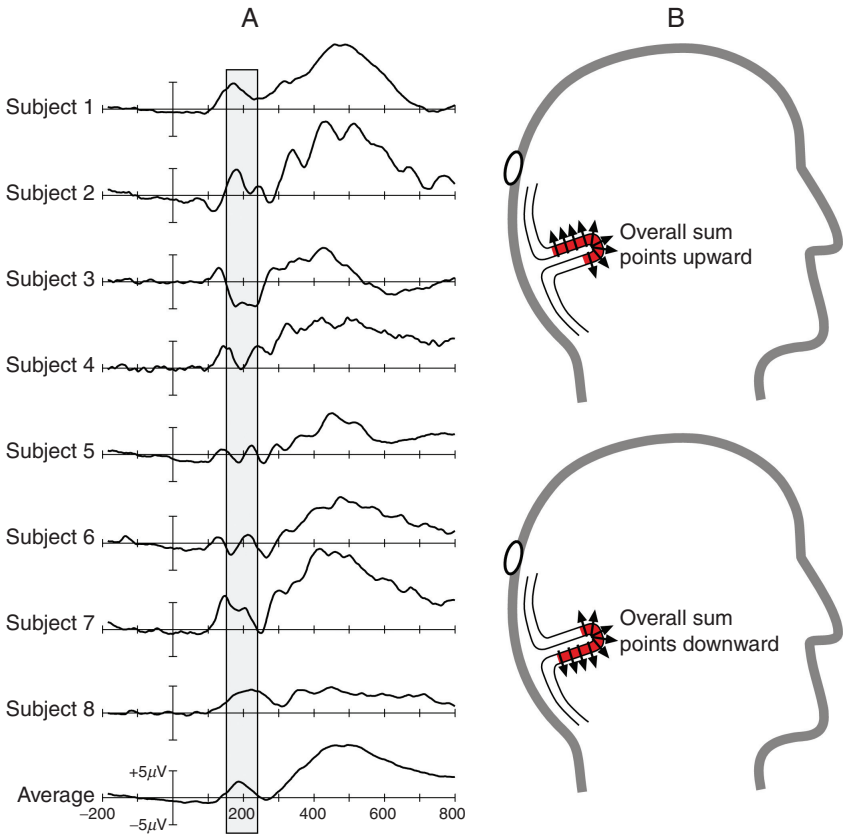


**Fig. 1.4.** (A) Single-subject ERP waveforms from 8 of 20 subjects in an oddball paradigm, along with the grand average of all 20 subjects (data from the study of Luck et al., 2009). (B) Example of how small differences between two subjects in the position of an active area of cortex within a sulcus could lead to opposite polarities at the electrode shown on the surface of the head. Each arrow represents the equivalent current dipole in a small patch of cortex, with positive at the arrowhead end and negative at the opposite end. Many of these dipoles will cancel each other, and the surface voltage will reflect the activity in the noncanceling dipoles.

beginning of the window (e.g., subjects 3, 4, and 7), and one subject's waveform is entirely negative during this window (subject 3).

The between-subject variations in the ERP waveform can be quite disconcerting when measuring a component from the single-subject waveforms. It is very unlikely that the same process reaches maximal activity at 145 ms in one healthy adult (e.g., subject 7) and at 220 ms in another (e.g., subject 6), so it does not seem appropriate to use a window that is broad enough to include peaks at such different latencies. And it is hard to understand how the negative deflection exhibited by subject 3 could represent the same functional brain activity as the positive deflection exhibited by subjects 1 and 2 in this same interval. However, as discussed above, peaks in the ERP waveform do not correspond directly to the underlying components. So, how problematic are these individual-subject waveform differences?

To understand whether the differences among individual-subject waveforms adversely affect our characterization of the components, we must first understand the source of the differences. For later periods of the waveform that reflect higher cognitive processes, differences in size and shape may reflect differences in the strategies subjects engage in during cognitive processing. Therefore, individual differences in the size and shape of the waveform may reflect actual processing differences. However, for the sensory processing that occurs within ~200 ms after the onset of the stimulus, it is unlikely that differences in waveform size and shape reflect differences in strategy or processing, at least in healthy subjects. Instead, the waveform differences most likely arise from differences across subjects in nonfunctional "nuisance" factors such as skull thickness and cortical folding patterns.

Just as fingerprints are unique to each individual, so is the intricate pattern of sulci and gyri in the human brain. Such changes in folding pattern could easily lead to differences in ERP waveforms across subjects like those illustrated in Figure 1.4A. For example, Figure 1.4B shows how a relatively small difference between two subjects in the location of an active strip of cortex within a sulcus could lead to opposite polarities for those two subjects at a given electrode site. More of the active region is on one side of the sulcus for one subject and more is on the opposite side of the sulcus for the other subject, leading to an overall equivalent current dipole pointing upward for one subject and pointing downward for the other subject. Consequently, the overall activity at a given scalp electrode will be positive for one subject and negative for the other.

Although this variability can be problematic for studies designed to assess individual differences, there is considerable similarity in the grand average ERP waveforms from different experiments that utilize similar tasks. This gives us some confidence that reliable conclusions can be drawn by comparing reasonably sized groups of subjects, even if the individual subjects within a group vary considerably in waveform shape. For example, there is great similarity across P3 oddball studies in grand average ERP waveforms, despite the fact that these waveforms are made up of different underlying individual-subject waveforms. Consider the ERP waveforms in Figure 1.5. The top left panel shows all 20 individual-subject waveforms from a P3 oddball task, subdivided at random into two separate groups of 10 subjects each. There is enormous variability between subjects in the amplitude and shape of the ERP waveform, with much larger P3 waves in some subjects than in others. However, as can be seen at the bottom of Figure 1.5, the grand averages across these two subgroups of subjects are quite similar in amplitude, timing, and shape, despite the large differences in the underlying individual-subject ERP waveforms that make up those grand averages. In other words, the individual-subject differences do not alter the overall experimental effect when the sample size is sufficient. However, it is important to remember that some measurement techniques may be more affected by this between-subjects variance than other techniques. We will address the issue of measurement later in the chapter. It should also be noted that statistical techniques can be applied that allow measurements to be made from grand averages rather than from single-subject waveforms, capitalizing on the stability of the grand averages (Kiesel et al., 2008; Miller et al., 1998, 2009).

### *How to Identify and Define an ERP Component*

Given how difficult it is to isolate a specific ERP component from the ERP waveform, you may be wondering how we even know that a specific ERP component exists. For example, how do we know that there is an N1 wave, a P3 wave, an N400, and so on? Of course, there is a voltage deflection in a broad time range corresponding to each of these components, but as we have already seen, there are usually multiple components active simultaneously in a given time range. So, how do we know that a voltage deflection is caused by a specific ERP

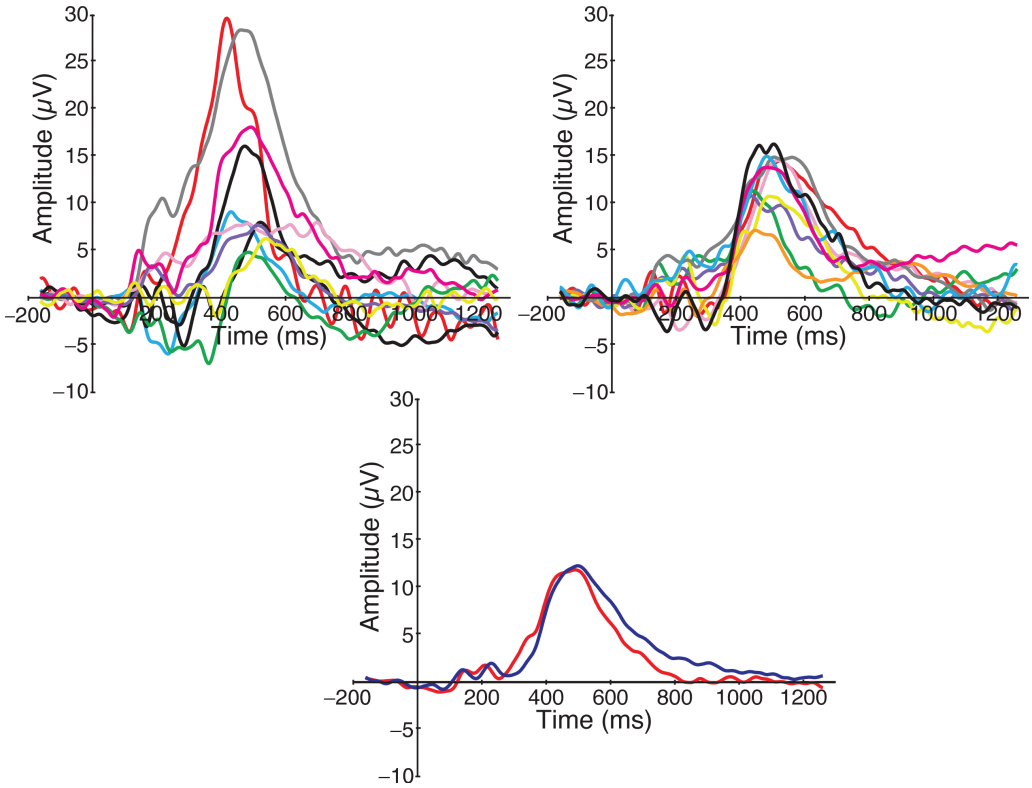


**Fig. 1.5.** Example of the similarity of grand average waveforms despite substantial differences among the single-subject waveforms. Waveforms from 20 subjects in an oddball experiment were randomly divided into two groups of 10. The single-subject waveforms for each group are shown at the top left and top right. Note the large variability in the amplitude and shape of the waveforms. The grand averages of these two subgroups of 10 subjects are shown at the bottom. Despite the large differences among the individual subjects, the grand averages from the two subgroups are quite similar.

component in one study, and how do we know that that same ERP component is active in subsequent studies? In other words, how do we *operationally identify and define* an ERP component?

Event-related potential components are often defined in terms of a combination of polarity, latency, and scalp distribution. This method of defining ERP components is evident from the common naming scheme in which ERP components are named in terms of polarity and latency (given either in milliseconds or as the ordinal position in the waveform). However, as we will see below, these dimensions describe the observed peaks and do not provide a stable and precise means of defining the underlying ERP components. That is, the factors of polarity, timing, and scalp distribution can vary from context to context, rendering them unstable representations of a component. We will explore each of these factors in turn and will end with some strategies for defining and isolating ERP components.

As discussed above, the timing of a neural process can vary across trials, subjects, and experiments. And because an ERP component is a scalp-recorded signature of a neural process, it stands to reason that the timing of an ERP component will vary across these same contexts. We can see timing variability quite clearly in studies of the P3 component, which can vary across conditions by hundreds of milliseconds, sometimes occurring before the manual response and sometimes appearing after the response. This is one reason that the moniker P300 is often shortened to P3, to eliminate the association with the time value of 300 ms. Although the timing of most ERP components is not nearly as variable as that of the P3, timing variability does occur for all ERP components. Visual sensory components, for example, increase in latency as stimulus brightness decreases for the simple reason that the amount of time required for information to reach cortex increases as brightness decreases. In addition, most components change in latency across early development (see Chapter 17, this volume) and across aging (see Chapter 18, this volume). Examining the variation in time windows over which the components

are measured in different studies makes the variation in component latencies across experiments quite obvious. Therefore, although a specific latency is often denoted by the name of an ERP component, this latency is approximate and often specific to the context in which the component was first identified, and latency cannot be used as a direct means of determining whether a component in a given study is the same as a component observed in previous studies, especially if the subjects, stimuli, or task differ considerably across studies.

Many ERP component names also make reference to the polarity of the component, but polarity may vary for a single component. For example, the C1 wave reverses in polarity for stimuli in the upper visual field compared with stimuli presented in the lower visual field owing to the cortical folding pattern of primary visual cortex. Both the positive- and negative-polarity C1 waves reflect the same underlying process and are therefore the same ERP component by any reasonable definition. Although other ERP components do not reverse polarity so dramatically, differences in cortical folding pattern across subjects might occasionally lead to polarity differences from one subject to the next at a given electrode site (see, e.g., subject 3 in Figure 1.4A). Furthermore, as discussed above, all ERP components are positive on one end of the dipole and negative on the other end, and all ERP components therefore reverse polarity at some place on the head.

If the polarity and timing information cannot be used to identify a component, what about the scalp distribution? Scalp distribution is often used to distinguish between components that have the same polarity and similar latencies, such as the "frontally distributed P3a" versus the "centroparietal P3b." In these cases, researchers often refer to a *family* of components (e.g., the N2 family of components) consisting of a set of *subcomponents* (e.g., the N2a, N2b, N2c, and N2pc subcomponents). Each subcomponent is actually a full-blown component, reflecting a different functionally and anatomically defined process, and the different subcomponents within a family are united only by their common polarity and similar timing.

Although adding the scalp distribution information can help to define a component, it will be ineffective if multiple subcomponents have similar scalp distributions (e.g., it seems likely that multiple different brain processes will produce a positive voltage deflection over the frontal lobes between 300 and 600 ms). Moreover, the scalp distribution for a single ERP component may vary across experimental contexts. For example, one subcomponent of the auditory N1 family arises from tonotopically mapped auditory cortex, and its scalp distribution therefore changes according to the pitch of the stimulus (Bertrand et al., 1991). Moreover, the scalp distribution in any given time range is influenced by all the components active in that range, which makes it difficult to determine the true distribution of a single component in a given experiment (unless that component has been isolated using one of the approaches described later in this chapter). Furthermore, the apparent scalp distribution can vary widely, depending on the choice of reference electrode (see Luck, 2005, chap. 3).

One additional variable that is often used to identify and define ERP components is their sensitivity to experimental manipulations or factors (see Donchin et al., 1978, for a thorough discussion). That is, what are the tasks, stimuli, timing parameters, and other factors that allow the component to be observed, and how do changes in these various factors modulate the timing, amplitude, and scalp distribution of the component? For example, the N2pc is observed for a target stimulus surrounded by distractors but not for a target stimulus presented in isolation (see Chapter 12, this volume). This dependence of the N2pc on the presence of distracting information in the display has played a large role in shaping various theories of the component. Furthermore, the N2pc has been shown to be largely unaffected by the probability of the target item (see Chapter 12, this volume), in contrast to the P3b. Therefore, sensitivity to experimental factors can help to identify the nature of a component and to distinguish among different components. However, just as discussed above with the variables of polarity, timing, and scalp distribution, sensitivity to experimental factors is not by itself a sufficient method for defining a component. For example, multiple ERP components may be sensitive to the same experimental manipulation, such as the similar dependence of P2 and P3b amplitude on the probability of the target stimulus. Furthermore, it is difficult to determine if an experimental manipulation has modulated the strength or timing of a specific component, or rather has resulted in a change in task strategy that has affected some other overlapping component. That is, it is difficult to assess whether an experimental manipulation has made an impact on a specific component, and it is also difficult to determine whether the experimental manipulation changed the strength, location, or timing of the neural process.

From these considerations, it should be clear that it is not appropriate to formally define an *ERP component* in terms of a combination of polarity, timing, scalp distribution, and sensitivity to experimental manipulations. These variables may be associated with a given component, but they do not define the component. We have instead argued that the term *ERP component* is best defined in terms of the scalp-recorded activity generated by a specific neural or psychological process, which in turn produces the polarity, latency, and scalp distribution of the component (which vary as that process varies), along with the sensitivity of the component to experimental manipulations. Unfortunately, our preferred definition is not very useful as an *operational* definition (i.e., a definition that describes the operations necessary to determine whether a specific voltage deflection reflects a specific component), because it is not usually possible to determine from the observed waveforms the voltage that is attributable to a specific known process.

Thus, in practice, the best way to identify a specific component is to take a *converging evidence* approach that intelligently combines various factors (including but not limited to polarity, latency, scalp distribution, and sensitivity to experimental manipulations) that would be expected to be true of a given process in a given context. For example, imagine that an oddball task was used in a study of elderly individuals, and a large positive voltage with a parietal maximum was observed to peak at 500 ms for the oddball stimuli, with a much smaller voltage observed for the standard stimuli. Four pieces of evidence converge on the conclusion that this voltage consists predominantly of the P3b component: (1) the voltage is positive at sites where the P3b is typically positive; (2) the latency is what we would expect given that cognition is typically slowed in elderly individuals; (3) the scalp distribution is consistent with previous studies of the P3b; and (4) the voltage shows the typical dependence on target probability. Now consider an example in which 5-year-old children are asked to passively view pictures of same-race faces and pictures of different-race faces, and a greater positive voltage is observed for the different-race faces with a peak latency of 325 ms. Imagine also that the voltage for both same-race and different-race faces was largest at parietal electrode sites, but the difference in voltage between same-race and different-race faces was largest at central sites. Is this a P3b component? A superficial analysis might lead to the conclusion that a larger P3b component was observed for the different-race faces, because the voltage was positive, peaked near 300 ms, and was maximal at parietal electrode sites. However, 325 ms would be an unusually early latency for a visual P3b component, especially in 5-year-old children. Moreover, even if a P3b were present in this latency range, the difference between conditions had a more central scalp distribution than is typical for the P3b component. Thus, it would be unlikely that this experimental manipulation primarily influenced P3b amplitude.

When this converging evidence approach is taken, it is important to consider both the strength of the evidence that a given component has a specific property and the degree to which other components might have that same property (this is essentially an application of Bayes's theorem). For example, although the N400 component is almost always present between 300 and 600 ms (see Chapter 15, this volume), many other components are also active in this latency range, so the finding that a given voltage deflection occurs in this latency range is not strong evidence that the deflection is an N400 component. In contrast, the lateralized readiness potential (LRP; see Chapter 9, this volume) and the N2pc component (see Chapter 12, this volume) have distinctive lateralized scalp distributions that are not present for many other components; the presence of this distinctive scalp distribution therefore provides strong (although not infallible) evidence that an LRP or N2pc was present.

With this approach, one is never completely certain that a specific component has been identified, and the strength of a conclusion will depend on both the number of pieces of converging evidence and the strength of each piece. Although it may be disappointing that one can never be certain that a specific component has been identified, this kind of uncertainty is common in all fields of science. Moreover, as discussed in the latter part of this chapter, it is sometimes possible to use *component-independent experimental designs* in which the conclusions of a study do not depend on identifying a specific ERP component.

## Linking Components with Processes: The Problems of Forward and Reverse Inference

Up to this point, we have assumed that we already know what neural or psychological process is reflected by a given ERP component. In this section, we consider how one might create this link (which we call the *problem of forward inference*) and how one might use this information to draw conclusions in new experiments (which Poldrack, 2006,

called the *problem of reverse inference* in the context of neuroimaging).

### *The Problem of Forward Inference*

It is more difficult than one might think to demonstrate that a given ERP component (or any other physiological measure) reflects a specific neural or psychological process. The challenge arises from the fact that we are looking for a neural measure of a given process because we do not fully understand the process and wish to use the neural measure to study the process. Because we do not fully understand the process, it is difficult to design unambiguous tests of the hypothesis that a given component reflects this process. For example, imagine that component A is hypothesized to reflect the encoding of information in verbal working memory. We could test this hypothesis by comparing the ERPs in a condition in which subjects are asked to encode words in working memory and a condition in which they passively view the same words. However, it is possible that working memory encoding is fairly automatic and would occur in both conditions; thus, the absence of a difference in component A between conditions might not be strong evidence against the hypothesis that this component reflects working memory updating. Moreover, if component A is found to differ between conditions, this could reflect some other process that differs between these conditions (see Shulman, 1996, for an interesting discussion of a related set of issues in the context of neuroimaging).

This problem could potentially be solved with a bootstrapping approach (the term *bootstrapping* refers to "pulling oneself up by one's bootstraps"). In this approach, one begins by trying the most obvious and unassailable manipulations of a given process to see if the component is present under the conditions in which everyone would agree that the process should be present. If the hypothesis survives multiple tests of this nature, it is tentatively accepted. The component is then used to test new hypotheses about the process it is thought to reflect. If these experiments yield results that are broadly consistent with evidence from other approaches, then confidence in the link between the component and the process continues to grow. If discrepancies arise, then researchers must reappraise the link between the component and the process.

As an example, consider the N2pc component (for a detailed discussion, see Chapter 12, this volume). Luck and Hillyard (1994) proposed that this component reflects an attentional filtering process that is used to suppress inputs from distractor objects surrounding a potential target. This was initially tested with the most obvious possible manipulations, such as removing the distractors to see if the N2pc component would disappear. A second set of experiments tested more refined manipulations based on findings from monkey single-unit experiments (Luck et al., 1997). The results of these experiments were consistent with the proposed link between N2pc and attentional filtering, and subsequent experiments assumed that this link was true and used it to test hypotheses about attention. For example, one study asked whether the same putative filtering mechanism was used by targets defined by different types of features (Girelli & Luck, 1997), and another series of experiments asked whether this mechanism was applied in parallel or in serial (Woodman & Luck, 1999, 2003b). However, later evidence demonstrated that the N2pc does not reflect filtering of the distractors per se, instead reflecting operations that must be applied to the attended object itself when distractors are present (Hickey et al., 2009). This is a modest change in the process thought to be reflected by the N2pc, but it was enough to slightly change the conclusions that can be drawn from the previous studies.

### *The Problem of Reverse Inference*

Once the problem of forward inference has been solved and a given component has been linked with some certainty to a given process, it is desirable to use this component as a measure of the presence, magnitude, and timing of that process in new experiments. This leads to the problem of reverse inference: If a component is present at a particular time, can we conclude that the process was present at that time? In Poldrack's (2006) analysis of this problem in the context of neuroimaging, the question is framed as follows: If brain activity has previously been observed in area X when process P is active, can we use the presence of activity in area X in a new experiment as evidence that process P was active in that experiment? As an example, Poldrack cited experiments using differences in activity in the dorsal striatum across conditions, which had previously been associated with reward processing, as evidence that reward mechanisms were differentially active in these conditions.

However, one must be cautious about using reverse inference. Reverse inference is actually a case of the well-known logical error of *affirming the consequent*. If the presence of P (e.g., reward) leads to the occurrence of X (activity in the striatum), this does

not mean that the occurrence of X necessarily entails the presence of P. For example, sleeping (P) causes the eyes to close (X), but eye closure (X) does not necessarily mean that someone is asleep (P). Reverse inference is valid only when it is possible to say that X occurs if *and only if* P occurs (i.e., X never occurs without P). In functional magnetic resonance imaging (fMRI) this standard is difficult to meet, because it is likely that the thousands of neurons in a given voxel and the millions of neurons within a cortical area are involved in multiple processes (e.g., the same neurons in visual cortex that are involved in perception are also involved in working memory). Consequently, it is not usually possible to assert that activity in a given voxel occurs if and only if a single process occurred.

Fortunately, an if-and-only-if condition is not as difficult to achieve for ERP components, because scalp ERPs represent a subset of the activity occurring within a given brain area. As described earlier, ERPs reflect the synchronous activity of cortical pyramidal cells, and many processes that occur within a given brain region will not lead to an ERP signature on the scalp. Consequently, whereas almost any process within a given brain region will change metabolic activity and therefore change the blood oxygen-dependent (BOLD) activity, only a subset of processes within a given region will produce a measurable ERP on the scalp. This makes ERP components more likely than BOLD responses to be tied to a specific process, and makes it less likely that a change in a given ERP component reflects different processes in different experiments. In other words, it is more plausible that a specific ERP component will be present if and only if a given process is present than that a BOLD response in a specific voxel will be present if and only if a given process is present.

For example, the evidence to date indicates that the N2pc component is present if and only if attention is allocated to an object in the presence of distractors. Of course, future research may demonstrate that the N2pc component can sometimes be elicited under conditions that do not involve this attention process, but it is at least plausible that this component might be present if and only if this attention process occurs. For example, when Luck and Ford (1998) found that an N2pc was present for conjunction targets and not for feature targets, they were reasonably justified in using reverse inference to draw the conclusion that a specific mechanism of attention was allocated to the conjunction targets and not to the feature targets. In contrast, there is no area of the brain in which one could reasonably assume that the presence of an increased BOLD signal necessarily reflected the allocation of attention.

Two main problems must be solved for reverse inference to be used with a given ERP component to draw strong conclusions. First, it is necessary to conduct a comprehensive set of experiments testing the hypothesis that the component of interest is present if and only if the corresponding process occurs. This is the problem of forward inference, and it is made difficult by the fact that we do not usually know enough about the process that a component hypothetically reflects to know whether this process is present or absent in a given experimental condition. Second, once the problem of forward inference has been solved, new experiments that attempt to use reverse inference must solve the problem of component identification. That is, one must be able to demonstrate that voltage deflections observed in the new experiments represent the same component observed in the earlier studies that established the link between the component and the process.

These two challenges are sufficiently difficult that it may never be possible to use reverse inference with complete certainty. However, as Poldrack (2006) discussed in the context of neuroimaging, one can use a Bayesian approach to draw probabilistic inferences on the basis of reverse inference. This involves assessing the probability that the ERP component would be present even if the corresponding process was not active and the probability that the corresponding process would be active without eliciting the ERP component. These probabilities are difficult to calculate, so this Bayesian approach is usually used informally. For example, we do not know the probability that an N2pc component would be present without the allocation of attention, and we do not know the probability that the allocation of attention may occur without an N2pc component. Thus, we cannot provide a precise probability for the claim that the variety of attention indexed by the N2pc component is needed for conjunction targets but not for feature targets (based on the presence of an N2pc for the former but not the latter). However, given that several experiments support the contention that N2pc is observed if and only if this particular mechanism of attention is present, and given that the N2pc can be isolated quite well from other components because of its distinctive contralateral scalp distribution, we can say something informal such as "The finding that an N2pc was present for conjunction targets but not

feature targets provides good evidence that the attentional processes that were present in prior N2pc experiments are needed for the detection of conjunction targets but not for the detection of feature targets."

Interestingly, the logic of reverse inference may sometimes allow stronger conclusions to be drawn from the absence of an ERP component than from its presence. If we can say that a given physiological measure X is always present when process P occurs—without the if-and-only-if restriction—then we can use the *modus tollens* argument from classical logic. This argument says that if we know that the presence of P entails X, then the absence of X entails the absence of P. That is, if previous experiments demonstrate that process P always leads to physiological measure X, then the absence of physiological measure X in a new experiment can be used to deduce that process P was not present. For example, Vogel and colleagues (1998) assumed that working memory encoding leads to the occurrence of a P3 wave (for supporting evidence, see Chapter 7, this volume). They found that this component was absent under conditions that led to an "attentional blink," and from this they concluded that no working memory encoding occurred for stimuli presented during the attentional blink. This is a logically valid conclusion. However, its truth depends on the validity of the initial assumption that working memory encoding leads to a P3 wave, which is not certain. Nevertheless, this general approach is less problematic than the typical use of reverse inference, which is based both on the assumption that a component is present when the corresponding process occurs and on the further assumption that the component is absent when the process does not occur. Of course, it is important to ensure that the absence of a voltage deflection in a given condition truly reflects the absence of the component of interest rather than cancellation by an opposite-polarity component, latency jitter, poor signal quality, low statistical power, and so on.

## Solving and Avoiding the Problems Associated with ERP Components

We have now seen how difficult it can be to associate changes in the observed ERP waveform with changes in an underlying ERP component. You may find yourself rightfully wondering, so what is this technique good for? In this section, we explore methods and strategies that have proven successful in using ERP components to answer questions about the mind and brain.

Event-related potentials provide a unique window into ongoing processing in the brain, serving as a continuous play-by-play of processing as it unfolds over time. It is this high temporal resolution of ERPs that makes them so desirable as a measure of brain processing. With ERPs, we can see processing before, during, and after the execution of behavioral responses, providing us with additional insights that cannot be gained with behavioral measures alone. However, the limitations of the ERP technique discussed in the previous sections mean that ERPs are only well suited for answering certain types of questions. Understanding the types of questions that can be readily answered with ERPs is essential for the successful application of the technique, and the remainder of the chapter will focus on describing several types of questions that ERPs have proven useful in answering.

The domains covered here may not encompass every current or potential use of ERPs; for example, ERPs may be useful as potential biomarkers in mental illness (Javitt et al., 2008; Luck et al., in press). However, the topics covered here provide a broad overview of the ways in which ERPs have been most commonly used to make scientific progress. These can be broadly divided into four domains, which we will explore in turn below: (1) determining which cognitive or neural process differs across conditions or across groups (e.g., perception, attention, response selection); (2) determining whether and when the brain has completed some set of processes; (3) uncovering new mental processes and subdividing known processes; and (4) covert monitoring of processing in situations in which overt behavior is difficult to measure or interpret (e.g., coma, infancy). We will examine each of these areas, providing specific examples of how ERPs have been used to expand our understanding in each domain.

### *Using Specific Components to Index Specific Processes*

One of the most notable and widely used applications of ERPs is to determine which specific neural or psychological process is affected by the factors of interest in the experiment. In other words, does a particular manipulation affect process A or process B or alternatively, do two groups of individuals differ in process A or process B? Using ERPs in this manner usually requires that (1) the precise neural or psychological process indexed by a component is known and understood and that (2) the component can be successfully isolated from the surrounding

and overlapping ERP components. As discussed earlier in the chapter, both of these requirements are difficult to meet; therefore, this branch of research typically relies on a number of assumptions concerning the specific nature of the ERP component of interest. These assumptions about the nature of the component are usually based on a wealth of previous research on the component and ideally include both studies in which the experimental manipulations that alter the component are explored and studies that are specifically aimed at elucidating the functional nature of the component (termed *ERPology* by Luck, 2005). We will first give an example of using components in this process-dependent manner to make the main issues facing researchers in this domain concrete, followed by some tips on how to successfully isolate and measure an ERP component.

Imagine that we wanted to understand why schizophrenia patients show prolonged reaction times (RTs) across a wide variety of behavioral tasks, an effect that has been observed for decades (see the review by Nuechterlein, 1977). In other words, which stage or stages of processing are slowed in schizophrenia patients, producing the slowing of behavioral RTs? We can address this question by examining whether particular ERP components are affected in the patient group compared to healthy controls. That is, is the scalp-recorded signature of a particular cognitive process delayed in latency or decreased in amplitude in the patients compared to the controls? This general approach has been used in studies of schizophrenia to examine abnormalities in many components, including the mismatch negativity (MMN), the P1 wave, the N2pc component, the P3 wave, the lateralized readiness potential, and the error-related negativity (Bates et al., 2002; Butler et al., 2007; Javitt, 2000; Jeon & Polich, 2003; Luck et al., 2006, 2009; see Chapter 19, this volume, for a review).

This approach—as typically applied—requires that previous experiments have already linked a component to this process, and it requires determining that a newly observed difference between patients and controls reflects a change in this specific component and not some other component (see the earlier sections on forward and reverse inference). For example, the N2pc component was used to assess whether prolonged behavioral RTs are accompanied by delays in the allocation of covert visual spatial attention in schizophrenia patients (Luck et al., 2006), which relied on previous work demonstrating that the N2pc is a scalp-recorded signature of covert shifts of visual attention and on the ability to isolate the N2pc from the surrounding ERP activity (which was achieved by using contralateral-minus-ipsilateral difference waves, as described in more detail below). Additionally, we can use ERPs to assess whether multiple stages of processing are affected in a patient group. For example, is the RT slowing exhibited by schizophrenia patients caused by a generalized slowing of all cognitive and neural processing or a combination of some subset of processes?

## METHODS FOR ISOLATING AN ERP COMPONENT

As described above, the ability to use ERP components as indexes of specific processes (reverse inference) depends on the ability to successfully isolate the component of interest from the surrounding ERP components. This is not an easy task. It may even seem impossible. However, there are a number of tricks that can be used to isolate a particular ERP component of interest from all of the other ongoing activity. Although the specific methods will depend on the specific task, ERP component, question of interest, and so on, the following strategies have proven successful in a number of different contexts.

One strategy is to focus the experimental design on ERP components that are large compared to the surrounding components. For example, the P3 wave is often >10 microvolts, making it easy to distinguish from the much smaller surrounding and overlapping ERP components. A second strategy is to focus the task design such that only one or two ERP components differ across conditions. When the design focuses on a small number of ERP components, it is easier to avoid significant component overlap, making the measurement of a specific component much easier. A third strategy involves subtracting out overlapping ERP components by creating difference waves between conditions or between electrode sites. For example, the lateralized readiness potential (LRP) is a difference wave created by subtracting the voltage at sites ipsilateral to the response hand from the activity at sites contralateral to the response hand. This subtraction process effectively isolates only the activity related to response selection, subtracting away the many other processes that do not differ between the contralateral and ipsilateral hemispheres; indeed, any brain activity that differs between the contralateral and ipsilateral electrode sites (relative to the hand that responds) must

be generated during or after the process that determines which hand should respond (see Chapter 9, this volume). Similarly, by computing a rare-minus-frequent difference wave in an oddball paradigm, it is possible to isolate probability-sensitive ERP components such as the P3 wave (see, e.g., Luck et al., 2009; Vogel et al., 1998).

Although difference waveforms can be an effective tool in isolating specific ERP components, they are not a panacea. First, a difference waveform is effective in isolating a specific ERP component only when all or most other components do no not vary across the two conditions used in the subtraction. Second, when a difference wave varies in amplitude across groups or across conditions, it is difficult to know which of the two waveforms used in the subtraction actually varies. For example, the LRP is decreased in schizophrenia patients relative to control subjects (Luck et al., 2009), but this could reflect less activation over the contralateral hemisphere or more activation over the ipsilateral hemisphere. Third, activity in a difference wave could reflect latency differences between the two original waveforms rather than a difference in amplitude.

An additional class of strategies uses scalp distribution information to isolate components. A simple version of this strategy is simply to measure a given component at an electrode site where this component is relatively large and other components are relatively small. A somewhat more sophisticated approach is to use a *vector filter*, which combines the data across all scalp sites in a manner that reflects the scalp distribution of a given component (see, e.g., Gehring et al., 1992). Event-related potential source localization techniques go one step further, providing a source waveform for each estimated generator site. In addition, ICA and PCA can use scalp distribution information to isolate the time course of each component.

When evaluating these different approaches, it is important to remember that, just as every researcher has his or her own individual limitations, each technique used to isolate ERP components is limited in its own special way. No technique—despite what its proponents may shout loudly from the research pulpit—is without its shortcomings, flaws, and limitations. Successfully using any of the techniques at our disposal requires that we know and understand the limitations of the method. Before using source localization, ICA, or even simple difference waves, one must be careful to fully understand how the technique works and when it might fail.

## METHODS FOR MEASURING AN ERP COMPONENT

Once a component has been successfully isolated from the overlapping activity, some quantitative assessment of the component must be made in order to compare it across conditions or across groups of subjects. The most widely used quantitative characterizations of ERP components include amplitude and latency assessments. Despite the inherent difference between peaks and components described above, it is common for ERP researchers to quantify ERP results by measuring the amplitude and latency of the peaks. Peak amplitude and peak latency measures are generally computed by choosing a time window surrounding a peak in the waveform and finding the most positive point in that time window (or the most negative point for a negative-going peak). The amplitude at this point is used as a representation of the magnitude of the component, and the latency of this point is used as a representation of the timing of the component. Historically, peak measures were employed because, as Donchin and Heffley (1978) so aptly stated, "it requires nothing but an x-y plotter, a ruler, and enough time" (p. 557). These were often all that a typical ERP researcher had at his or her disposal in the early days of ERP research, but researchers today have computers capable of performing much more advanced algorithms than those that a ruler can accomplish, and we are no longer limited to such simple measurement techniques.

Is there anything special about the amplitude or timing of the peaks in the observed ERP waveforms? As Figure 1.2 illustrates, the amplitude and timing of the peaks in the observed waveform may be quite different from the amplitude and timing of the underlying components that sum together to produce the observed waveform. And as Figure 1.3 illustrates, factors such as latency variability can strongly influence peak amplitude. Moreover, it seems simplistic to assume that a process that extends over hundreds of milliseconds can be quantified by the value of a single time point. In addition, when the values are measured at multiple electrode sites, it makes no sense to use the peak at each electrode site to measure a single component: The peak will occur at a different time at each electrode site, but a given component necessarily has the same time course at each electrode site (because of the instantaneous transmission of voltage). Peak measures have other shortcomings as well (summarized in Luck, 2005, chap. 6), and there is a clear

trend away from peak measures among sophisticated ERP researchers.

How, then, can one better quantify the magnitude and timing of an ERP component? The first step is usually to isolate the component by computing some kind of difference wave that subtracts away most of the other components. As an example, consider the MMN data shown in Figure 1.6. In this experiment, subjects were presented with a frequently occurring *standard* pitch or a rare *deviant* pitch every 1000 ms (see Chapter 6, this volume, for details). When the deviant pitch was sufficiently different from the standard pitch, the ERP waveform was more negative for the deviant pitch than for the standard pitch from approximately 100 to 200 ms poststimulus. If we attempted to quantify the magnitude of this effect by measuring the amplitude of the most negative peak between 100 and 200 ms, we would face two serious problems. First, because the overall waveform contains a P2 peak during this interval, there is no negative peak to be measured in many of the waveforms shown in Figure 1.6 (especially in the waveforms elicited by the standards). Second, even if we could find a negative peak, the voltage at this peak would reflect a combination of this P2 wave, the MMN, and any other components that were active during this period.[5] Thus, it is better to quantify the magnitude of the MMN from the deviant-minus-standard difference wave.

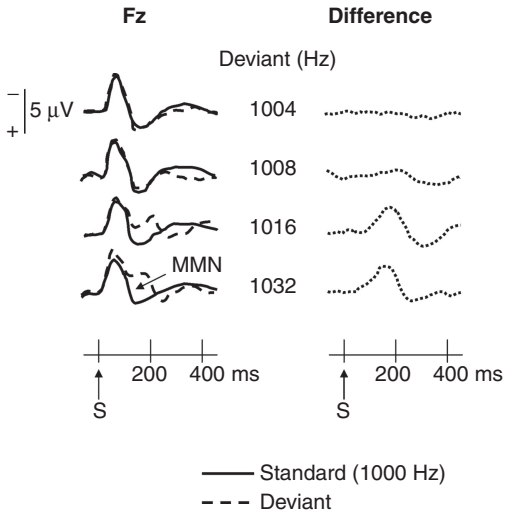


**Fig. 1.6.** Example of the use of difference waves in the context of MMN. The left side shows the waveform elicited by a 1000 Hz standard tone that occurred on 80% of trials, overlaid with deviant stimuli that differed in pitch from the standard by varying amounts and occurred on 20% of trials. The right side shows the deviant-minus-standard difference waves. Note that this is the same as in Figure 6.1 in Chapter 6, this volume.

By measuring amplitude or latency from a difference wave, the contributions of the overlapping peaks are reduced or eliminated. Of course, this will work well only if the other components are equivalent across the two waveforms that are used for the subtraction so that they are eliminated in the difference wave. One could use a peak amplitude measure to quantify the amplitude of a component in the difference, and this would certainly be an improvement over measuring peak amplitude from the two original waveforms used in the subtraction. However, there is still no particular reason to choose this one point as a reflection of the magnitude of the underlying process. If one is interested in the overall magnitude of a brain response, it is usually more reasonable to measure the area under the curve or the mean voltage over the duration of the component (these are nearly equivalent: mean is simply area divided by duration). An important exception arises, however, when one is trying to measure the amplitude of a component that varies in latency across conditions or across groups; in this case, it may be necessary to use a method that finds the peak and then measures the amplitude at (or around) this peak.

Peak latency is also a poor measure in most cases, because the latency of the peak is not usually a particularly interesting time point. Quantifying the latency of an ERP component by finding the peak is analogous to quantifying RT by finding the mode of the RT distribution for each subject. Instead, it is sometimes possible to quantify the midpoint of a component by finding the time point that divides the area under the curve into two equal portions. This is called the *50% area latency* measure, and it is closely related to median RT (see Luck, 2005, chap. 6). In addition, theories of cognitive processes often make predictions about the onset or duration of a process rather than the midpoint. Kiesel and colleagues (2008) provide an excellent comparison of the different methods that can be used for the onset of a component, and these methods can be easily extended to measure the offset and duration of a component.

### *Assessing the Time Course of Processing*

The temporal resolution of ERPs makes them an excellent tool for determining the time course of a neural or psychological process. The simplest way to do this is to measure the latency of a given peak in two different conditions or two different groups and

use this as a measure of the amount of time required for this process to occur in the two conditions or two groups. However, this approach is not usually very powerful, because it does not isolate a specific component and because it uses the peak as a measure of timing. A more powerful approach is to compare the waveforms from two conditions or from two groups of subjects to ascertain the point in time at which the waveforms begin to diverge. For example, ERPs have been used in the emotion literature to determine when, after the onset of a stimulus, processing differs between emotion-evoking and neutral stimuli (see Chapter 16, this volume). There are advantages and limitations to using ERPs in this manner, and we will explore both of these through some examples below.

Let's consider the emotion example mentioned above, in which we wish to know by what point in time processing related to the emotional content of a stimulus has begun. In other words, by what point in time has the brain distinguished between emotional and nonemotional stimuli? We can answer this question by comparing the ERP waveforms elicited by neutral stimuli (e.g., a picture of a landscape) and emotion-eliciting stimuli (e.g., a picture of a mutilation). We can use the time point at which the waveforms begin to diverge as a measure of when the brain has distinguished between the neutral and emotional stimuli. That is, the waveforms between an emotional and a nonemotional condition cannot diverge until the brain has begun to distinguish the emotional content of the stimulus (provided that all other factors, including physical stimulus factors, are matched between the conditions). The advantage of this approach is that, although specific ERP components may differ between the conditions, the conclusions about timing do not rely on isolating a specific ERP component. That is, the presence of a difference between conditions at a given time indicates that the brain has distinguished between the two conditions by this time, regardless of which component was responsible for this difference. This approach is one case of what are called *component-independent experimental designs* (see Luck, 2005, chap. 2).

Because this method does not require isolating a specific component or linking a component with a specific process, it generally requires fewer assumptions than using ERPs in a component-dependent manner. However, there are some limitations to this approach. For example, it is important to note that this method provides an upper bound on the timing of an effect. Because many processes may be invisible in scalp ERP recordings, the brain might make a distinction between two stimuli long before the first point at which the scalp-recorded signals differ. Therefore, one can use ERPs to say that a particular effect has occurred *by* a particular time point, but one cannot use ERPs to conclude that an effect did not begin *until* a particular time. In our emotional content example, one could conclude that the brain has begun to process information related to emotional information by the point at which the waveforms diverge. However, one could not say that emotional processing did not begin until that time point, because the effect could have begun earlier in brain areas that did not give rise to a scalp-recorded ERP. Generally speaking, this technique is valuable in providing evidence that an effect happens early in the processing stream, but it cannot be used to prove that an effect does not happen until late in the processing stream.

The limitations in the conclusions that can be drawn about timing from ERPs may seem debilitating to the technique, but using ERPs in this manner has answered many important questions about cognitive and neural processing. For example, ERPs were able to end a long-standing debate in the attention literature about whether attention operates at an early stage or a late stage of processing (for reviews, see Hillyard et al., 1998; Luck et al., 2000). It is difficult to determine from behavioral studies whether the effects of attention on response speed and accuracy arose from changes in perceptual processing or changes in a postperceptual stage of processing. However, because ERPs provide a continuous measure of processing between the stimulus and the response, they can indicate whether the attention effects begin early or late in the processing stream. That is, the *locus of selection* can be assessed directly by asking whether the ERP waveforms for attended and ignored stimuli diverge early in time (e.g., within the first 100 ms after stimulus onset) or late in time (e.g., more than a few hundred milliseconds after stimulus onset). Research using this approach has shown that—at least under some conditions—attention influences sensory processing within the first 50 ms after stimulus onset for auditory stimuli and within the first 100 ms after stimulus onset for visual stimuli (see Chapter 11, this volume). These ERP results provided key evidence in favor of early selection models of attention, helping to answer a fundamental question that could not be easily addressed using behavioral techniques.

This time-based approach is often combined with the process-specific approach described in the

previous section, in which the effects are linked with specific components. For example, researchers have argued that the early ERP attention effects consist of modulations of specific sensory-evoked ERP components (see, e.g., Di Russo et al., 2003; Woldorff et al., 1993). This has been difficult to establish with complete certainty because of the many difficulties associated with trying to identify specific components, as discussed earlier in the chapter. However, the *converging evidence* approach described earlier in this chapter has been used to provide substantial support for the hypothesis that attention influences specific ERP components. Even more important, the simple fact that the waveforms for attended and unattended stimuli diverge at an early time provides very strong evidence that attention can influence perceptual processing.

### MEASURING PROCESSES THAT OCCUR PRIOR TO A COMPONENT

A related approach uses an ERP component to assess the processes that must have occurred *prior* to the ERP component. The advantage of this approach is that it does not require that we first determine a solid link between an ERP component and a specific process (i.e., we do not need to solve the forward inference problem). Instead, we can use simple assumptions about the processes that must have occurred prior to the ERP component to draw inferences about these processes.

As an example, consider the N400 component, which countless studies have shown is larger for words that mismatch the current semantic context than for words that match this context (reviewed by Kutas, 1997). For example, the word *nurse* will elicit a larger N400 if it is preceded by an unrelated word such as *cup* than if it is preceded by a related word such as *doctor*. Substantial controversy surrounds the question of exactly what process the N400 component represents (see Chapter 15, this volume). However, it is safe to assume that this difference in N400 between words that match and mismatch a semantic context could not occur unless the words were perceived. Thus, if we see that a given word elicits a larger N400 when the preceding word was related than if it was unrelated, then we can be certain that the words were perceived. This logic has been used to show that, under certain conditions, attention does not influence sensory processing and that words are fully perceived even when unattended (Luck et al., 1996; Vogel et al., 1998, 2005). That is, although attention influences sensory processing under some conditions, modulating the early sensory-evoked components, under other conditions attention only influences postperceptual processes that follow word identification (see the reviews by Luck & Hillyard, 1999; Luck & Vecera, 2002). Under these latter conditions, attention has no impact on the difference in N400 amplitude for words that match versus mismatch the current semantic context.

As a second example, consider the P3b component, which every ERP researcher knows is larger for infrequent target stimuli than for frequently occurring standard stimuli. However, an important implication of this probability dependence often goes unnoticed. Specifically, the onset of the difference in P3b amplitude between rare and frequent stimuli cannot occur until the brain has at least begun to determine whether the eliciting stimulus belongs to the rare category or the frequent category. This implication was spelled out very clearly by Kutas and colleagues (1977), who framed it in terms of the then-popular idea that the P3b component was elicited by surprising stimuli: "before a stimulus can surprise it must be identified. As P300 commonly appears as a discriminative response to specific stimuli within a series, its elicitation must be preceded by an adequate evaluation of the stimulus at some level of processing" (p. 792–793). This idea is commonly described by saying that the latency of the P3 wave reflects *stimulus evaluation time*, but this is a somewhat vague description. It is much more precise—and powerful—to say that the onset of the difference between the waveforms elicited by the rare and frequent stimuli reflects a time at which the brain has begun to determine whether the stimulus belongs to the rare or the frequent category. That is, the waveforms between these two conditions could not differ until the brain has determined whether the stimulus belongs to the rare or the frequent category, indicating that by that point the brain has begun to categorize the stimuli.

We have applied this more precise framing of P3b latency to understanding why behavioral RTs are slowed in patients with schizophrenia (Luck et al., 2009). Each stimulus in this experiment was a digit or a letter, with one category rare ($p$ = .2) and the other frequent ($p$ = .8). Subjects were asked press a button with one hand for digits and a button with the other hand for letters, and patient RTs were approximately 60 ms slower than control RTs. As shown in Figure 1.7, the voltage in the P3 latency range was larger for control subjects than for patients, for both the rare and frequent stimulus categories, but the latency of the P3 peak was similar across groups. However, given that many different processes

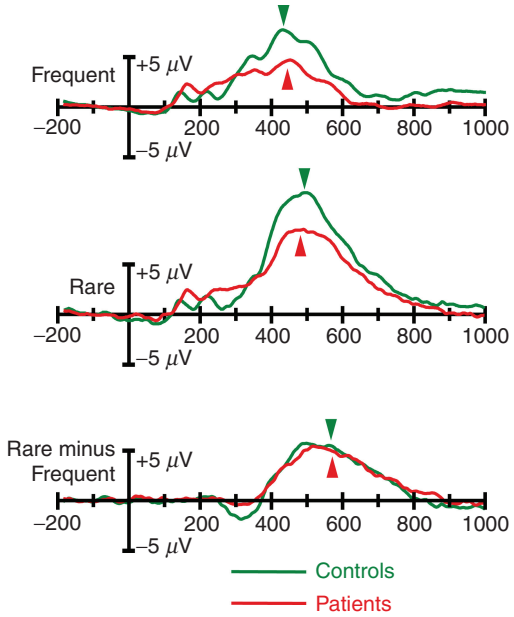


**Fig. 1.7.** Grand average ERPs recorded at the Pz electrode site from schizophrenia patients and control subjects (from the study of Luck et al., 2009). The patient and control waveforms are overlaid for the frequent stimuli, the standard stimuli, and the rare-minus-frequent difference wave. Triangles show mean P3 latency for each group, quantified as peak latency for the rare and frequent stimuli and 50% area latency (the point that divides the area under the curve into two equal portions) for the difference wave.

presumably overlap during the P3 time range, it is difficult to draw firm conclusions on the basis of the time of the peak voltage in this time range. More precise conclusions can be drawn by examining the rare-minus-frequent difference waves in each group (Figure 1.7, bottom). These difference waves reflect the differential processing of the rare and frequent stimulus categories, and any nonzero voltages in these difference waves must be a consequence of a preceding process that determined the category to which a stimulus belonged. The only difference between patients and controls in these difference waves was a reduction in amplitude in the time range of the N2 wave in patients. The difference waves were nearly identical across groups in the P3 time range, and the midpoint of the deflection in this wave (the time that divided the area under the curve into equal halves) was nearly identical across groups. Thus, no delay was observed in the brain's differential responding to rare versus frequent stimuli in the patients compared to the controls, despite a 60 ms slowing of the behavioral response in patients. This suggests that the slowing of behavioral responses was not caused by a slowing of the processes that lead up to the categorization of the stimuli, but was instead caused by postcategorization slowing. This conclusion was further supported by a reduction in the amplitude and a slowing of the latency of the lateralized readiness potential in the patients compared to the controls.

It is important to note that, in both of these examples, conclusions were drawn about the processes that logically must have preceded the component being measured rather than the process the component directly reflected. That is, the N400 was used to assess the perceptual processes that must occur before the brain can distinguish between semantically related and unrelated words, and the P3b was used to assess the perceptual and categorization processes that must occur before the brain can determine whether a stimulus belongs to a rare or a frequent category. An important advantage of this approach is that we do not need to know with certainty what process produces a given ERP component. Instead, we can make very straightforward assumptions about what processes must occur for a component to differ across conditions. In many cases, it does not actually matter which component differs across conditions; the mere presence of a difference indicates that the brain has made a specific discrimination by a given point in time. Thus, this is another example of a component-independent approach. This does not mean that components are irrelevant in the design of the experiment. Instead, it means that the conclusions do not depend on whether a specific component has been identified in the results.

### *Uncovering and Subdividing Mental Processes*

Event-related potentials have also been useful in identifying new, previously unknown mental processes and subdividing known processes into multiple separate subprocesses. From behavioral measures, it is difficult to ascertain how many mental processes intervene between the occurrence of a stimulus and the execution of a behavior. However, ERPs provide a continuous measure of processing before, during, and after the execution of the behavior. Therefore, it is possible with ERPs to identify processes that were previously unknown.

For example, error-related negativity (ERN; see Chapter 10, this volume) occurs after the execution of a response and therefore reflects a process that behavioral measures cannot directly measure. Although previous studies had pointed to the existence of processes related to detecting and correcting

errors (e.g., Laming, 1979; Rabbitt, 1966), no one had hypothesized a process with the timing of the ERN. The ERN helped to focus research on the processes occurring within 100 ms of an error response, which has led not only to numerous studies of processes related to error detection, but also to a large literature on response-conflict monitoring.

Similarly, ERPs can be used to determine whether a given behavioral effect is the result of a change in a single process or of multiple separable subprocesses. Almost every experimental manipulation that produces a behavioral effect leads to differences between conditions in multiple ERP components, and this naturally leads to the idea that the behavioral effect reflects changes in more than one process. Consider, for example, manipulations of attention. It is parsimonious to assume that any experiment in which behavioral responses are faster or more accurate for attended stimuli than for unattended stimuli reflects the operation of a single mechanism of attention, and most behavior-inspired theories of attention have taken a monolithic view of attention. However, ERP studies have demonstrated that different manipulations of attention influence different ERP components, demonstrating that different mechanisms of attention operate to produce the observed behavioral effects under different conditions (see Chapter 11, this volume). These ERP studies have inspired behavioral studies demonstrating that the details of the behavioral attention effects are indeed best explained by the existence of multiple mechanisms of attention (see, e.g., Vogel et al., 2005). Thus, the ability to monitor multiple processes with ERPs makes it possible to provide empirical evidence against simplistic explanations of behavior that invoke a single mechanism.

### *Covert Monitoring*

A final ERP approach involves using ERPs as a means of "covertly monitoring" processing in situations in which behavioral output is uninformative, inapplicable, or unavailable. There are three general situations in which this approach is applied: (1) assessing processing in individuals who cannot or will not make a behavioral response (e.g., infants, coma patients); (2) assessing processing under conditions in which requiring a behavioral response might invalidate the task (e.g., monitoring the processing of unattended stimuli); and (3) assessing processes that might not be evident in behavior (e.g., the processing of subliminal stimuli). In this section, we will provide examples of all three of these situations.

Behavioral methods used with infants almost always take advantage of the fact that infants tend to orient toward some types of stimuli (e.g., complex, dynamic, or novel stimuli) more than other types of stimuli (Brennan et al., 1966). And if they exhibit greater looking times toward one category of stimuli than another, then this is evidence that they were able to distinguish between these categories (Spelke, 1985). The categories can be simple sensory categories (e.g., the presence versus absence of a fine pattern) or complex conceptual categories (e.g., animal versus nonanimal). However, it is always possible that infants are able to make a particular discrimination even if they fail to exhibit any behavioral orienting on the basis of this discrimination. Moreover, these techniques are difficult to use prior to about 4 months of age owing to poor motor control. Event-related potentials can be useful in these situations to determine whether the brain has made a given discrimination.

For many years, ERPs have been used in this way to determine whether newborn infants might be suffering from hearing loss. Specifically, a rapid sequence of clicks is presented, and the amplitude and latency of the early brainstem evoked responses are used to determine whether the sensory response is abnormal (Stapells, 1989). The auditory MMN component has also been widely used to assess the ability of infants to make more complex perceptual discriminations, such as distinctions between phonemes (see Chapter 6, this volume). Other components have been used to assess higher-level aspects of visual processing in infancy, such as face perception, and even higher-level cognitive discriminations (see Chapter 17, this volume). It is generally easier to assess lower-level sensory processes than higher-level cognitive processes with ERPs, because the sensory processes can typically be assessed without any kind of task. Higher-level processes are typically task-dependent, and it is difficult to teach infants a task that will elicit these processes reliably. One can sometimes take advantage of spontaneous differences in processing between, for example, rare and frequent stimulus categories, but these spontaneous differences may habituate before enough trials have been acquired to obtain reliable averaged ERP waveforms.

Event-related potentials can also be used in individuals who are unable to make behavioral responses due to a medical condition. In amyotrophic lateral sclerosis, for example, ERPs have been used to create brain–computer interfaces that allow patients to communicate with their families and caregivers

(Silvoni et al., 2009). Another recent example comes from coma research, where ERPs have been used to predict which patients are likely to recover (Fischer et al., 2004). There are also cases in which an individual might refuse to make a valid behavioral response, such as a suspect in a in crime, and ERPs have been used to assess whether people have knowledge of an event that they are not admitting (e.g., Farwell & Donchin, 1991).

Another type of covert monitoring approach is used when the requirement to make a behavioral response might interfere with the processing of a task. The most obvious example of this arises in attention research, in which ERPs have been widely used to compare the processing of attended and unattended stimuli (see Chapter 11, this volume). Requiring a behavioral response for an unattended stimulus presumably creates an incentive to attend to the stimulus, which is problematic for the study of attention. However, because ERPs can be recorded just as easily for unattended stimuli as for attended stimuli, they can be used to assess the processing of stimuli for which there is absolutely no incentive to attend.

This approach has also been used extensively in language research (see Chapter 15, this volume). In studies of sentence comprehension, it is difficult to assess the processing of each individual word by means of behavioral measures, because this would require interrupting the sentence for a response. Eye movement measures have often been used for this purpose in studies of reading, because the eye movements are a naturally occurring part of the reading process. However, the eye movements are still discrete events that occur some time after the eyes have landed on a given word, and they are applicable primarily in the context of written language comprehension rather than spoken language comprehension.

The third variety of covert monitoring involves asking questions about processes that might not be evident in behavior. That is, the brain may engage in a given process and reach a specific result without that result reaching awareness or triggering a behavioral response. The most obvious case of this arises in research on perception without awareness. By using ERPs, it is possible to determine how much information has been extracted from a stimulus that fails to reach awareness. For example, research has shown that a specific type of masking (*object substitution masking*) does not eliminate the orienting of attention to a target stimulus, as indexed by the N2pc component (Woodman & Luck, 2003a), but it does impair the processes needed to generate an N400 difference between words that match versus mismatch a semantic context (Reiss & Hoffman, 2006). This pattern of results indicates that this variety of masking operates after early perceptual processing but prior to semantic analysis. Similarly, stimuli that are associated with a given response will activate the preparation of that response, as indexed by the LRP, even if the subject is unaware of the stimulus and does not actually execute the response (Dehaene et al., 1998).

## Conclusions

The ERP technique provides a unique and highly informative perspective on brain processing, but like all techniques it suffers from challenges, difficulties, and limitations. The goal of this chapter was to chronicle both the positive and negative sides of ERPs, exploring issues that are often unaddressed in the literature while providing a detailed set of strategies that allow the technique to be optimally employed. We hope that these recommendations allow the reader to understand and avoid the down sides of ERP research while also adopting our view that the positives of ERP research outweigh the negatives.

## Acknowledgment

We thank Bill Gehring and Greg Hajcak for helpful editorial comments and suggestions.

## Notes

1 The negative peaks of the waveform are sometimes referred to as *troughs*; however, there is nothing special about whether the activity is positive or negative in polarity. Therefore, we will refer to both the positive and negative deflections in the waveform as *peaks*.

2 The necessity for summation across large groups of neurons to observe a scalp ERP has implications that are often neglected by researchers. First, the magnitude of an ERP will depend on both the size of the individual postsynaptic potentials and the number of neurons that are active. Second, many neurons that are simultaneously active within a given cortical region may actually be doing very different things, and an ERP component may therefore reflect a mixture of different neural responses. Bill Gehring suggested to us that recording the ERP waveform is analogous to measuring the number of cars crossing the San Francisco Bay Bridge at a given time of day: These cars may have nothing in common except that many of their drivers are heading home for dinner. Thus, ERPs may be useful for answering broad questions about neural activity (analogous to asking when most people end their workday in San Francisco) and not as useful for answering narrow questions (analogous to asking where individual cars are going or what their occupants are doing).

3 Analogous effects can be seen for neural firing rates; for example, the duration of a change in the firing rate of a typical neuron in visual cortex following a brief stimulus is typically at least 100 ms. This is presumably a result of PSPs that last at least 100 ms.

4 The claim that brain processes involve individual brain areas requires us to be a bit more specific about what we mean by the term *process*, because much brain activity involves the interaction of multiple brain areas. We are using the term *process* to mean an elementary computation that might plausibly occur within a single brain area (e.g., spatial filtering based on lateral inhibition within an area) rather than a multistep computation that likely involves the coordinated operation of multiple brain areas (e.g., retrieval of an item from memory).

5 It should be noted that these two problems are not this extreme in all cases. For example, if one measures the amplitude or latency of the P3 peak when this component is much larger than all of the other components, then these measures will not be greatly distorted by the overlapping components. However, other shortcomings of peak measures still apply in this situation, and small differences between groups or conditions could easily reflect differences in the overlapping components.

## References


Bates, A. T., Kiehl, K. A., Laurens, K. R., & Liddle, P. F. (2002). Error-related negativity and correct-response negativity in schizophrenia. *Clinical Neurophysiology*, *113*, 1454–1463.

Bertrand, O., Perrin, F., & Pernier, J. (1991). Evidence for a tonotopic organization of the auditory cortex with auditory evoked potentials. *Acta Otolaryngologica*, *491*, 116–123.

Brennan, W., Ames, E. W., & Moore, R. W. (1966). Age differences in infants' attention to patterns of different complexities. *Science*, *151*, 354–356.

Butler, P. D., Martinez, A., Foxe, J. J., Kim, D., Zemon, V., Silipo, G., Mahoney, J., Shpaner, M., Jalbrzikowski, M., & Javitt, D. C. (2007). Subcortical visual dysfunction in schizophrenia drives secondary cortical impairments. *Brain*, *130*, 417–430.

Davis, P. A. (1939). Effects of acoustic stimuli on the waking human brain. *Journal of Neurophysiology*, *2*, 494–499.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P. F., & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395*, 597–600.

Di Russo, F., Martinez, A., & Hillyard, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cerebral Cortex*, *13*, 486–499.

Di Russo, F., Martinez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, *15*, 95–111.

Donchin, E., & Heffley, E. F., III. (1978). Multivariate analysis of event-related potential data: A tutorial review. In D. Otto (Ed.), *Multidisciplinary perspectives in event-related brain potential research* (pp. 555–572). Washington, DC: U.S. Government Printing Office.

Donchin, E., Ritter, W., & McCallum, W. C. (1978). Cognitive psychophysiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, & S. H. Koslow (Eds.), *Event-related brain potentials in man* (pp. 349–441). New York: Academic Press.

Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology*, *28*, 531–547.

Fischer, C., Luaute, J., Adeleine, P., & Morlet, D. (2004). Predictive value of sensory and cognitive evoked potentials for awakening from coma. *Neurology*, *63*, 669–673.

Galambos, R., & Sheatz, G. C. (1962). An electroencephalographic study of classical conditioning. *American Journal of Physiology*, *203*, 173–184.

Gehring, W. J., Gratton, G., Coles, M., & Donchin, E. (1992). Probability effects on stimulus evaluation and response processes. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 198–216.

Girelli, M., & Luck, S. J. (1997). Are the same attentional mechanisms used to detect visual search targets defined by color, orientation, and motion? *Journal of Cognitive Neuroscience*, *9*, 238–253.

Hickey, C., Di Lollo, V., & McDonald, J. J. (2009). Electrophysiological indices of target and distractor processing in visual search. *Journal of Cognitive Neuroscience*, *21*, 760–775.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society: Biological Sciences*, *353*, 1257–1270.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*, 65–96.

James, W. (1890). *The principles of psychology*. New York: Holt.

Javitt, D. C. (2000). Intracortical mechanisms of mismatch negativity dysfunction in schizophrenia. *Audiology and Neurotology*, *5*, 207–215.

Javitt, D. C., Spencer, K. M., Thaker, G. K., Winterer, G., & Hajos, M. (2008). Neurophysiological biomarkers for drug development in schizophrenia. *Nature Reviews Drug Discovery*, *7*, 68–83.

Jeon, Y. W., & Polich, J. (2003). Meta-analysis of P300 and schizophrenia: Patients, paradigms, and practical implications. *Psychophysiology*, *40*, 684–701.

Kiesel, A., Miller, J., Jolicoeur, P., & Brisson, B. (2008). Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, *45*, 250–274.

Kutas, M. (1997). Views on how the electrical activity that the brain generates reflects the functions of different language structures. *Psychophysiology*, *34*, 383–398.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, *197*, 792–795.

Laming, D. (1979). Choice reaction performance following an error. *Acta Psychologica*, *43*, 199–224.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Luck, S. J. (in press). Event-related potentials. In D. L. Long (Ed.), *APA handbook of research methods in psychology*. Washington, DC: American Psychological Association.

Luck, S. J., & Ford, M. A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences, USA*, *95*, 825–830.

Luck, S. J., Fuller, R. L., Braun, E. L., Robinson, B., Summerfelt, A., & Gold, J. M. (2006). The speed of visual attention in schizophrenia: Electrophysiological and behavioral evidence. *Schizophrenia Research*, *85*, 174–195.

Luck, S. J., Girelli, M., McDermott, M. T., & Ford, M. A. (1997). Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cognitive Psychology*, *33*, 64–87.

Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1000–1014.

Luck, S. J., & Hillyard, S. A. (1999). The operation of selective attention at multiple stages of processing: Evidence from human and monkey electrophysiology. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 687–700). Cambridge, MA: MIT Press.

Luck, S. J., Kappenman, E. S., Fuller, R. L., Robinson, B., Summerfelt, A., & Gold, J. M. (2009). Impaired response selection in schizophrenia: Evidence from the P3 wave and the lateralized readiness potential. *Psychophysiology*, *46*, 776–786.

Luck, S. J., Mathalon, D. H., O'Donnell, B. F., Hämäläinen, M. S., Spencer, K. M., Javitt, D. C., & Ulhaaus, P. F. (in press). A roadmap for the development and validation of ERP biomarkers. *Biological Psychiatry*.

Luck, S. J., & Vecera, S. P. (2002). Attention. In S. Yantis (Ed.), *Stevens' handbook of experimental psychology: Vol. 1: Sensation and perception* (3rd ed., pp. 235–286). New York: Wiley.

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, *382*, 616–618.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*, 432–440.

Mazaheri, A., & Jensen, O. (2008). Asymmetric amplitude modulations of brain oscillations generate slow evoked responses. *Journal of Neuroscience*, *28*, 7781–7787.

Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, *35*, 99–115.

Miller, J., Ulrich, R., & Schwarz, W. (2009). Why jackknifing yields good latency estimates. *Psychophysiology*, *46*, 300–312.

Nuechterlein, K. H. (1977). Reaction time and attention in schizophrenia: A critical evaluation of the data and theories. *Schizophrenia Bulletin*, *3*, 373–428.

Picton, T. W., Alain, C., Woods, D. L., John, M. S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., & Trujillo, N. J. (1999). Intracerebral sources of human auditory-evoked potentials. *Audiology & Neurotology*, *4*, 64–79.

Poldrack, R. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*, 59–63.

Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, *71*, 264–272.

Reiss, J. E., & Hoffman, J. E. (2006). Object substitution masking interferes with semantic processing: Evidence from event-related potentials. *Psychological Science*, *17*, 1015–1020.

Shulman, R. G. (1996). Interview with Robert G. Shulman. *Journal of Cognitive Neuroscience*, *8*, 474–480.

Silvoni, S., Volpato, C., Cavinato, M., Marchetti, M., Priftis, K., Merico, A., Tonin, P., Koutsikos, K., Beverina, F., & Piccione, F. (2009). P300-based brain–computer interface Communication: Evaluation and follow-up in amyotrophic lateral sclerosis. *Frontiers in Neuroscience*, *3*, 60.

Spelke, E. (1985). Preferential-looking methods as tools for the study of cognition in infancy. In G. Gottlieb & N. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life: A methodological overview* (pp. 323–363). Norwood, NJ: Ablex.

Spencer, K. M., Dien, J., & Donchin, E. (2001). Spatiotemporal analysis of the late ERP responses to deviant stimuli. *Psychophysiology*, *38*, 343–358.

Stapells, D. R. (1989). Auditory brainstem response assessment of infants and children. *Seminars in Hearing*, *10*, 229–251.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1656–1674.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2005). Pushing around the locus of selection: Evidence for the flexible-selection hypothesis. *Journal of Cognitive Neuroscience*, *17*, 1907–1922.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences USA*, *90*, 8722–8726.

Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, *400*, 867–869.

Woodman, G. F., & Luck, S. J. (2003a). Dissociations among attention, perception, and awareness during object-substitution masking. *Psychological Science*, *14*, 605–611.

Woodman, G. F., & Luck, S. J. (2003b). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 121–138.