Linking Brainwaves to the Brain:
An ERP Primer

Alexandra P. Fonaryova Key, Guy O. Dove, and Mandy J. Maguire

Department of Psychological and Brain Sciences
University of Louisville

This article reviews literature on the characteristics and possible interpretations of the event-related potential (ERP) peaks commonly identified in research. The description of each peak includes typical latencies, cortical distributions, and possible brain sources of observed activity as well as the evoking paradigms and underlying psychological processes. The review is intended to serve as a tutorial for general readers interested in neuropsychological research and as a reference source for researchers using ERP techniques.

Over the latter portion of the past century recordings of brain electrical activity, such as the continuous electroencephalogram (EEG) and the stimulus-relevant event-related potentials (ERPs) became frequent tools of choice for investigating the brain’s role in cognitive processing in different populations. These electrophysiological recording techniques are generally noninvasive, relatively inexpensive, and do not require participants to provide a motor or verbal response. Furthermore, virtually identical procedures can be used across the entire life span (e.g., Molfese & Hess, 1978; Molfese & Molfese, 1979; Molfese & Schmidt, 1983; Nelson, Thomas, de Haan, & Wewerka, 1998). However, although the ongoing EEG reflects a wide range of neural activity related to the various sensory and cognitive functions, it also reflects the myriad of self-regulation processes ongoing in the brain at the same time (e.g., maintaining body temperature, heart rate, breathing). This intermixing of signals makes it difficult to separate cognitive and physiological contributors to the observed EEG. In contrast, the ERP approach permits inves-
tigators to link recorded signals with stimulus events more directly by focusing on the change in electrophysiological signal that occurs immediately following the stimulus event (Callaway, Gruae, & Shatton, 1975; Rockstroh, Elbert, Birbaumer, & Lutzenberger, 1982). The smaller size of ERPs relative to other physiological events can make it difficult to discern the relevant signal. To accommodate these factors, researchers employ repeated presentations of the evoking stimulus to average out potentially unrelated events.\(^1\)

ERPs have been successfully used to study both general and specific aspects of an individual’s response to events in the external as well as internal environment (e.g., Molfese, 1978a, 1978b). Neuropsychological research of cognitive functioning in various populations also demonstrated that ERP components could serve as informative markers of neurodevelopmental status in general as well as reflect development of more specific abilities (Courchesne, 1978). Additional advantages of the ERP technique over other procedures include (a) very fine temporal resolution (on the order of milliseconds or even fractions of a millisecond) that reveals even momentary changes in patterns of brain activation that otherwise could go unnoticed, and (b) relatively gross-level spatial resolution capabilities that allow for theorizing about the distribution of brain mechanisms that subserve these cognitive functions.

ERP waveforms are typically described in terms of positive and negative peaks (i.e., the most positive and negative deflections in the waves). At a general level, the labeling refers to the sequence in which the peak occurs and to its polarity. For example, \(N1\) would refer to the first negative peak in the waveform, and \(P2\) would label the second positive peak. The naming scheme for ERP components can also identify the positive and negative peaks by their latency (usually defined as the time from stimulus onset). \(N100\) in this example refers to the negative peak that occurs 100 msec following stimulus onset. \(P300\) would identify the positive peak that occurred 300 msec poststimulus onset.

In contrast to this objective peak-naming convention, functional descriptions of ERP peaks refer to their psychological interpretation. In the past, Donchin (1978) proposed a distinction between exogenous and endogenous components, suggesting that the former were sensitive mainly to physical properties of external stimuli, and the latter were affected by information processing and could be elicited even by the event absence. However, further research indicated that some components appeared to share characteristics of both groups (e.g., \(N1, P2\); Shibasaki & Miyazaki, 1992), depending on the stimulus properties. Although a variety of terms were proposed for this subgroup, such as transient (Hugdahl, 1995) or mesogenous components (Fabiani, Gratton, & Coles, 2000), in general, functional

\(^1\)Recently, Makeig et al. (2002) demonstrated that some ERP features are not independent of the background EEG and therefore proposed a single trial rather than an average analysis approach for ERP data that would provide more detailed information about cortical dynamics.
descriptions of the ERPs have shifted away from such classification toward identifying more specific cognitive processes reflected by each peak.

In addition to the latency measures and functional interpretations, ERP descriptors often include topographical scalp distributions or identify electrodes where maximum amplitudes are typically observed. Such information can be useful for interpreting ERP peaks that may occur at the same time but over different scalp areas reflecting different cognitive processes. However, the scalp distribution does not necessarily correspond to the actual brain areas generating the signal. The ERPs are generally believed to reflect postsynaptic (dendritic) potentials (Allison, Wood, & McCarthy, 1986) of a fairly extensive set of neurons activated in close temporal proximity. The orientation of the cortical columns generating the signal may affect whether the electrodes detect a signal and where on the scalp it is maximal. If the columns are perpendicular to the scalp, the likelihood of recording a strong signal is good. At the same time, columns from different brain areas may project to the same scalp area resulting in a larger signal (if the polarities are the same). Further, if the cell columns are oriented parallel to the scalp, or at some other angle to it, the signal may project to an area away from the nearest electrode above it and thus fail to be recorded or to be noted by electrodes over other scalp locations (e.g., a signal originating in the left hemisphere may be maximal over the right hemisphere). Because of this imperfect relation between the observed scalp topography and the actual brain structures involved in generating it, the scientific community has recently moved another step forward in extending ERP descriptions to include the potential brain sources of observed activity, rather than focusing on scalp distributions alone.

Given the great variety of ERP paradigms, analyses, and proposed implications, a reader may find it challenging to make sense of the reported findings or integrate them into the more general frame of psychology. Currently, there are several reviews of ERP components available (e.g., Fabiani, Gratton, & Coles, 2000; Hugdahl, 1995); however, we are not aware of any articles that describe a wide range of ERP components and include all four characteristics: peak latency, cognitive functional significance, scalp distributions, and component brain sources. This review is intended to fill that void. The following sections describe most commonly identified components of adult ERPs: P1, N1, P2, N2, mismatch negativity (MMN), P3a, P3b, N400, and P600 in the order they appear in the brain wave.² For the purpose of consistency and clarity, the peaks are identified by their polarity and place in the sequence of components rather than by exact latency because of possible variations in the latter

²This list is not assumed to be exhaustive. Other ERP components such as the contingent negative variation (Hillyard & Picton, 1987), left anterior negativity (Friederici & Mecklinger, 1996), late positive potential (Cuthbert et al., 1995), and positive slow wave (N. K. Squires, Squires, & Hillyard, 1975) are not included in this review due to a lack of information regarding their sources or to the limited space available to cover a large amount of research.
due to developmental, environmental, or clinical effects (unless the latency is the predominant descriptor of the peak). Because peak characteristics can vary as a function of stimulus modality and reference location, our review separates data for auditory and visual paradigms and notes the references used to identify topographic maxima. Finally, different techniques used for source localization of the observed ERPs rely on different principles and therefore can produce conflicting results. Thus, findings from intracranial recordings, functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), brain electromagnetic source analysis (BESA), positron emission tomography (PET), or low-resolution brain electromagnetic tomography (LORETA) may not always agree. Consequently, the specific method for source localization is noted for each brain source listed in this article.

P1

This peak is not always easily identified, but when present it occurs approximately 50 msec after an auditory stimulus onset (also known as P50) or about 100 msec after the onset of a visual stimulus. Functionally, this component is usually interpreted as a neurophysiological indicator of preferential attention to sensory inputs (suppression of unattended information) and is thought to reflect the general level of arousal.

Auditory

The auditory P1 appears earlier in time (shorter latency) over posterior scalp electrode sites but with larger amplitudes over frontal or central regions or both. Nagamoto, Adler, Waldo, Griffith, and Freedman (1991) reported that the peak was largest over the Cz electrode (nose reference). The distribution is symmetrical over the two hemispheres, except for the anterior temporal regions where larger amplitudes are noted over the left hemisphere. Overall, peak amplitude and latency appear to decrease with age to the point where the peak disappears (Coch, Groissi, Coffey-Corina, Holcomb, & Neville, 2002).

Auditory P1 has been frequently associated with auditory inhibition (Waldo et al., 1992) and typically tested in a sensory gating paradigm where paired clicks are presented at relatively short interstimulus intervals. The amplitude of the averaged ERP to the second of the paired clicks is reduced compared to the averaged response to the first click. The magnitude of this suppression is commonly interpreted as a neurophysiological index of sensory gating. Reduced suppression is frequently reported for certain neuropsychiatric disorders, including mania and schizophrenia, where peak amplitude to paired stimuli is reported to be approximately equal (Jin et al., 1997; Patterson et al., 2000; Siegel, Waldo, Mizner, Adler, & Freedman, 1984; Waldo et al., 1991). P1 latency is often used clinically to diag-
nose neurodegenerative diseases, such as multiple sclerosis and Parkinson’s disease (Squires & Ollo, 1986).

Buchwald et al. (1992) proposed that the P1 response is associated with the ascending reticular activating system and its postsynaptic thalamic targets. Using a MEG approach, Thoma et al. (2003) and Huotilainen et al. (1998) independently localized the sources of the auditory P1 in the superior temporal gyrus. Weiss et al. (2001) coregistered auditory evoked potentials and magnetic fields. The resulting equivalent dipole model consisted of one source in the auditory cortex of each hemisphere as well as a radially oriented medial frontal source. Similar findings identifying frontal and temporal generators were reported by Potts, Dien, Hartry-Speiser, McDougal, and Tucker (1998), using the current source density approach.

Visual

The visual P1 response is different from the auditory component in terms of the evoking stimulus, neurocognitive and neurophysiological mechanism, peak latency, scalp distribution, and neural sources. The visual P1 is typically recorded in a checkerboard-reversal task or similar light-flash paradigms, but it can also be present for other visual stimuli (e.g., faces) and is largest over the occipital regions (Hugdahl, 1995). A negative component may be present at the same latency over frontal and central regions (Rossion, Campanella, et al., 1999; left earlobe reference). The amplitude of P1 generally varies with the amount of attention (Clark & Hillyard, 1996—spatial selective attention task; Mangun, Hillyard, & Luck, 1993—Posner’s cuing paradigm). Luck (1995) proposed that P1 reflects suppression of noise because the amplitude decreased for unattended locations and did not increase for attended stimuli. Mangun et al. interpreted it to reflect encoding of form and color (ventral “what” pathway). Further, the amplitude of P1 increased when speed of response was emphasized, suggesting that this peak may also reflect the level of arousal (Vogel & Luck, 2000).

Probable sources for the visual P1 were identified using PET, BESA, and LORETA methods in ventral and lateral occipital regions (Clark, Fan, & Hillyard, 1996; Gomez, Clark, Luck, Fan, & Hillyard, 1994), suggesting a striate (Strik, Fallgatter, Brandeis, Pascual-Marqui, 1998) or extrastriate (posterior fusiform gyrus) origin (Heinze et al., 1994). Rossion, Campanella, et al. (1999) submitted data from a face identification paradigm to BESA and reported similar sources as well as sources in posterior-parietal regions, indicating additional involvement of dorsal and ventral neural components.

N1

This component was originally investigated by Hillyard, Hink, Schwent, and Picton (1973) in a dichotic listening paradigm and is one of the most easily identi-
fied components, regardless of the specific analysis approach employed. There is good convergence in findings based on principal components analysis factor scores (Beauducel, Debener, Brocke, & Kayser, 2000), baseline-to-peak amplitude (Pekkala, Rinne, & Naatanen, 1995; Sandman & Patterson, 2000), and latency measures (Segalowitz & Barnes, 1993).

Generally, N1 is assumed to reflect selective attention to basic stimulus characteristics, initial selection for later pattern recognition, and intentional discrimination processing (e.g., Vogel & Luck, 2000). Latency and amplitude of the peak depend on the stimulus modality. Auditory stimuli elicit a larger N1 with shorter latency than visual stimuli (Hugdahl, 1995).

Auditory

For auditory stimuli, N1 typically occurs approximately 100 msec after stimulus onset and has maximum amplitude over frontocentral areas (Vaughan & Ritter, 1970; nose reference) or the vertex (Picton, Hillyard, Krausz, & Galambos, 1974). More recent studies differentiated it into three components with maximum amplitudes over temporal areas (latency 75 msec and 130 msec) and over vertex (latency 100 msec; Giard et al., 1994; McCallum & Curry, 1980; nose reference). Naatanen and Picton (1987) reviewed the three components of N1 and proposed that the early temporal and vertex components reflect sensory and physical properties of the stimuli (e.g., intensity, location, timing in regard to other stimuli), but the later temporal component appears to be less specific in its response and reflects transient arousal.

The amplitude of the auditory N1 is enhanced by increased attention to the stimuli (Hillyard et al., 1973; Knight, Hillyard, Woods, & Neville, 1981; Mangun, 1995; Ritter, Simpson, & Vaughan, 1988) and by increasing the interstimulus interval (Hari, Kaila, Katila, Tuomisto, & Varpula, 1982). The latter has been attributed to contributions of additional sources from frontal cortical areas (Hari et al., 1982).

N1 appears to be most likely generated by sources in primary auditory cortex in the temporal lobe (Vaughan & Ritter, 1970). MEG, BESA, and lesion studies consistently localize auditory N1 in the superior temporal plane (e.g., Knight, Scabini, Woods, & Clayworth, 1988; Papanicolaou et al., 1990; Scherg, Vajsar, & Picton, 1989). However, several studies proposed additional sources in the frontal lobe that could be activated from the temporal lobe (e.g., Giard et al., 1994; current source density and equivalent current dipoles analysis).

Visual

The visual N1 component is usually largest over the occipital region (Hopf, Vogel, Woodman, Heinze, & Luck, 2002; reference not mentioned) or the inferior temporal regions (Bokura, Yamaguchi, & Kobayashi, 2001; average reference). N1 am-
plitude is typically larger in stimulus discrimination tasks (Mangun & Hillyard, 1990; Vogel & Luck, 2000), but it is reduced when the stimuli are presented at short intervals. The increased amplitude is attributed to enhanced processing of the attended location (Coull, 1998; Luck, 1995), including spatial properties of the stimuli (Mangun et al., 1993), and is not due to arousal because the amplitudes were larger in a task that placed no emphasis on the speed of response (Vogel & Luck, 2000). It is also not affected by inhibition as indicated by the lack of Go–No-go response differences (Bokura et al., 2001). In addition, similar to the auditory N1, a visual N1 was noted to include at least two distinct subcomponents, one occurring at 100 msec over the central midline sites and another present at 165 msec over the posterior sites (Vogel & Luck, 2000; average mastoids). The researchers attributed the more anterior visual N1 solely to response preparation processes because it could be eliminated by not requiring a motor response and decreased stimulus onset asynchrony.

Using a combination of techniques (MEG, ERP, and MRI), Hopf et al. (2002) located visual N1 sources in the inferior occipital lobe and the occipitotemporal junction. However, Bokura et al., (2001), using the LORETA approach, identified additional sources of the visual N1 in the inferior temporal lobe.

P2

The P2, like the N1 and P1, has long been considered to be an “obligatory cortical potential” because it has low interindividual variability and high replicability (Roth, Kopell, Tinklenberg, Huntsberger, & Kraemer, 1975; Sandman & Patterson, 2000; Shelley et al., 1991). The P2 component has been identified in many different cognitive tasks, including selective attention (Hackley, Woldorff, & Hillyard, 1990; Hillyard et al., 1973; Johnson, 1989), stimulus change (Naatanen, 1990), feature detection processes (Luck & Hillyard, 1994), and short-term memory (Golob & Starr, 2000; Starr & Barrett, 1987). Similar to N1, P2 has been consistently identified by principal components analysis factor scores (Beauducel et al., 2000), baseline-to-peak amplitude (Beauducel et al., 2000; Sandman, & Patterson, 2000), and latency measures (Segalowitz & Barnes, 1993). Functional interpretations of the P2 include attention modulation of nontarget stimuli (Novak, Ritter, & Vaughan, 1992) and stimulus classification (Garcia-Larrea, Lukaszewicz, & Mauguiere, 1992).

Auditory

In the auditory modality, P2 often occurs together with N1 (referred to as N1–P2 complex) and shares many characteristics of the preceding component, yet the two peaks can be dissociated experimentally and developmentally (Hugdahl, 1995;
Oades, Dittmann-Balcar, & Zerbin, 1997; see Crowley & Colrain, 2004, for a review). The maximum amplitude of the P2 can span a broader latency range (150–275 msec) compared to the N1 (Dunn, Dunn, Languis, & Andrews, 1998) and can be double-peaked (Hyde, 1997; Ponton, Don, Eggermont, Waring, & Masuda, 1996). The scalp distribution of the P2 is less localized than that of the N1 (Naatanen, 1992), but typically the highest amplitude is noted over the central region, using either the left mastoid or the linked earlobe references (Holcomb, Ackerman, & Dykman, 1986; Iragui, Kutas, Mitchiner, & Hillyard, 1993); therefore this peak is often referred to as a vertex potential or vertex positivity.

P2 is sensitive to physical parameters of the stimuli, such as pitch (Novak et al., 1992) and loudness (Hegerl & Juckel, 1993; Hillyard & Picton, 1987). Similar to the N1, the amplitude of the P2 peak gets larger as the stimulus intensity increases; however, opposite to the N1, it continues to increase for stimuli with intensity above 70 dB (Adler & Adler, 1989). Participant differences, such as reading ability, can also affect the P2 amplitude to auditory stimuli (Bernal et al., 2000).

Generators for the auditory P2 are thought to be located mainly in the primary and secondary auditory cortices (Zouridakis, Simos, & Papanicolau, 1998; MEG). Combined analyses using MEG and intracranial recordings identified possible P2 sources in the planum temporale, whereas MEG alone also located an additional source in auditory association complex (Area 22; Godey et al., 2001). When using dipole source analysis, both the N1 and P2 elicited by auditory stimuli are often represented by two dipoles: one for the primary auditory cortex and one for the secondary auditory cortex (Hegerl, Gallinat, & Mrowinski, 1994; Scherg & Berg, 1991). Using BESA and LORETA to identify dipole locations for the N1–P2 component, Mulert, Juckel, Augustin, and Hegerl (2002) identified one in the superior temporal region with a tangential orientation, whereas the second was located in the temporal lobe with a radial orientation, but sources specific to P2 have not been reliably separated from the N1 generators. Some evidence toward independent generators of the P2 comes from lesion studies reporting that damage to the temporoparietal areas did not affect properties of the P2 but resulted in reduction of the N1 (Knight et al., 1988).

Visual

In the visual domain, topographic distribution of the P2 is characterized by a positive shift at the frontal sites around 150 msec to 200 msec after stimulus onset (right mastoid reference; Heslendfeld, Kenemans, Kok, & Molenaar, 1997; Kenemans, Kok, & Smulders, 1993; van der Stelt, Kok, Smulders, Snel, & Gunning, 1998) and a large negativity, approximately 200 msec, following stimulus onset at the occipital sites (Talsma & Kok, 2001; right earlobe reference). The amplitude of a visual P2 increases with the complexity of the stimuli (Pernet et al., 2003). Using BESA dipole analysis, Talsma and Kok (2001) reported a symmetri-
cal dipole pair localized in the inferior occipital (extrastriate) area. However, the researchers noted that both topographic distribution and the exact dipole positions varied slightly for the attended and not-attended visual stimuli.

N2

The N2 component is characterized by higher interindividual variation (Michalewski, Prasher, & Starr, 1986; Pekkonen et al., 1995) and has multiple psychological interpretations, including orienting response (Loveless, 1983), stimulus discrimination (Ritter, Simson, & Vaughan, 1983; Satterfield, Schell, Nicholas, Satterfield, & Freese, 1990), and target selection (Donchin, Ritter, & McCallum, 1978), possibly reflecting task demands (Duncan et al., 1994; Johnson, 1989). Findings also show that the N2 is smaller in amplitude and shorter in latency for shorter interstimulus intervals (Polich & Bondurant, 1997).

Very few studies have investigated the “basic” N2 peak first reported by K. C. Squires, Squires, and Hillyard (1975). In their study, participants viewed two stimuli; the first was expected to give information about the upcoming second image. When that image did not match what was expected, they observed a larger N2 with frontal distribution, compared to when these expectations were met. At present, N2 is considered to be a family of responses that differ, based on the features of the eliciting stimuli, such as modality (Donchin et al., 1978) and trial presentation parameters (Ceponiene, Rinne, & Naatanen, 2002). These components share some of their functional interpretation with MMN (see MMN section) because both appear to indicate a detection of a deviation between a particular stimulus and the participant’s expectation. However, unlike the MMN studies, for the N2 to be present the participant must pay attention to the stimuli.

Auditory

Auditory stimuli elicit the highest N2 amplitudes over the central parietal region (Simson, Vaughan, & Ritter, 1977; nose reference). Based on scalp current density analysis, Bruneau and Gomot (1998) suggested that the auditory N2 has bilateral sources in the supratemporal auditory cortex.

Visual

Visual stimuli were reported to elicit the highest N2 amplitudes over the preoccipital region (Simson et al., 1977; nose reference). The N2 to visual stimuli has been shown to vary based on the task type (semantic vs. physical discrimination; Ritter et al., 1983) and stimulus type, such as written words, pictures of objects, or human faces. Using intracranial electrodes placed directly on the cortex,
Allison, Puce, Spencer, and McCarthy (1999) observed that letter strings of recognizable nouns produced an N2 component at the fourth occipital gyrus near the occipitotemporal sulci. Pictures of complex objects, such as cars and butterflies, elicited an N2 response over the inferior lingual gyrus medially and the middle occipital gyrus laterally. This effect was not present for scrambled pictures. Face-recognition tasks elicited an N2 at the fusiform gyrus and inferior temporal or occipital gyri just lateral to the occipitotemporal or inferior occipital sulci. The differential processing of human faces has led many researchers to investigate the visual processing of human faces (see the N170 section). These differing distributions indicate that the N2 peak may reflect category-specific processing (Allison et al., 1999).

**N2 AND INHIBITION**

One of the variants of N2 is associated with the Go–No-go paradigm, in which the participant is asked to respond to some stimuli (Go trials), and inhibit the response to another class of stimuli (No-go trials). The ERPs on No-go trials are characterized by a large negative peak relative to the Go trials between 100 msec and 300 msec after stimulus onset (Eimer, 1993; Jodo & Kayama, 1992; Kok, 1986; Kopp, Mattler, Goertz, & Rist, 1996; Pfefferbaum, Ford, Weller, & Kopell, 1985). Given the nature of this task, it is often thought to be associated with response inhibition (Gemba & Sasaki, 1989; Jodo & Kayama, 1992; Sasaki & Gemba, 1993). Pfefferbaum et al. (1985) showed that this response occurred both in relation to overt and covert responses, indicating that the N2 Go–No-go effect cannot be completely attributed to motor responses. Instead, it appears to be present whenever responses must be interrupted (Kopp et al., 1996).

The amplitude and polarity of the N2 inhibition response can change, depending on the complexity of the task. The amplitude of N2 was noted to increase when subjects had less time to respond (Jodo & Kayama, 1992). In some instances, the Go–No-go response has also been reported as a positive peak (Schiller, Bles, & Jansma, 2003; left mastoid reference), possibly due to large amplitude of the P300 in difficult tasks (Kiefer, Marzinsik, Weisbrod, Scherg, & Spitzer, 1998).

The N2 for both visual and auditory tasks is especially pronounced over the frontocentral electrodes when the Go response is withheld (Gemba & Sasaki, 1989; Jodo & Kayama, 1992; Mathalon, Whitfield, & Ford, 2003; Miltner et al., 2003; Pfefferbaum et al., 1985; Thorpe, Fize, & Marlot, 1996), regardless of the reference point, such as the earlobes (Jodo & Kayama, 1992), left ear (Miltner et al., 2003), and the linked mastoids (Mathalon, Whitfield, & Ford, 2003).

Mathalon et al. (2003), using both ERP and fMRI, identified the involvement of the caudal and motor anterior cingulate cortices during both correctly and incor-
rectly inhibited responses suggesting that the N2 reflects general inhibitory responses.

N170

The N170 peak is another member of the N2 family and ranges in latency between 156 msec and 189 msec (Bentin, Allison, Puce, Perez, & McCarthy, 1996; George, Evans, Fiori, Davidoff, & Renault, 1996; Jemel, Pisani, Calabria, Crommelinck, & Bruyer, 2003; Rossion, Campanella, et al., 1999; Rossion, Delvenne, et al., 1999; Taylor, McCarthy, Saliba, & Degiovanni, 1999). It is associated primarily with visual processing of human faces. The topographic distribution of the N170 component for both familiar and unfamiliar faces is largest over the occipitotemporal regions (Allison et al., 1999; Bentin et al., 1996; George et al., 1996; Jemel et al., 2003). These results are consistent across studies and reference points, such as the mastoids (Allison et al., 1999) and the nose (Jemel et al., 2003). N170 amplitude is significantly larger in response to faces than other natural or human-made objects (Bentin & Deouell, 2000; Eimer, 2000), and patients suffering from prosopagnosia do not show an N170 response to faces (Bentin & Deouell, 2000). However, Tanaka and Curran (2001) recently proposed that the N170 is not specific to human faces, but reflects expert object recognition. In their study dog experts showed an increased N170 to pictures of dogs, but not birds, whereas bird experts showed the opposite effect.

Intracranial recordings of evoked potentials (Allison et al., 1999; Bentin et al., 1996) and fMRI studies (Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997) all point to the fusiform gyrus as the possible neuroanatomical substrate of N170. However, source localization of the N170 using BESA identified the potential source in the lateral occipitotemporal region outside the fusiform gyrus (Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002).

MMN

First described by Naatanen, Gaillard, and Mantysalo (1978), the MMN is a negative deflection that has a typical latency of 100 msec to 250 msec. The amplitude is largest at frontal and central electrode sites (Fabiani et al., 2000; Liebenthal et al., 2003) and has been replicable with different reference points, including the tip of the nose (e.g., Liebenthal et al., 2003; Pekkonen et al., 1995), the earlobe, and noncephalic locations (Aarts, Kraus, McGee, & Nicol, 1991). MMN is elicited using an oddball paradigm where an occasional deviant stimulus is presented in a stream of more frequent standard stimuli (but see Naatanen, Pakarinen, Rinne, & Takkegata, 2004, for a five-deviants paradigm). Because MMN paradigms typically do not require attention to the stimuli, they have been widely used in developmental research (Csepe,
1995; Csepe, Dieckmann, Hoke, & Ross, 1992; Kraus, Koch, McGee, Nicol, & Cunningham, 1999) and sleep studies (Alho, Saino, Sajaniemi, Reinikainen, & Naatanen, 1990; Campbell, Bell, & Bastien, 1991). Although MMN is associated with considerable test–retest reliability (Pekkonen et al., 1995), it may be affected by many paradigm characteristics. Many researchers reported increased MMN with shorter ISIs (e.g., Naatanen, 1992), whereas others noted that ultra-short ISIs (e.g., 350 msec) may lead to reduced amplitudes for some types of stimuli (e.g., affect intensity but not frequency discrimination; Schröger, 1996). Some reports indicate a substantially reduced MMN response in subjects not attending to the stimuli (Paavilainen, Alho, Reinikainen, Sams, & Naatanen, 1991). MMN characteristics may also depend on the number of trials because too many deviant trials may allow a participant to habituate to the particular stimulus, thus diminishing the MMN amplitude. McGee et al. (2001) mapped the habituation of adults, children, and guinea pigs for complex and simple stimuli and found that as the number of exposures increased, the size of the MMN response decreased in a nonlinear fashion. Further, the exact time for habituation varied as a function of the complexity of the stimuli.

Auditory

In the auditory modality, the MMN can be evoked by any perceivable physical deviance from the standard stimulus, such as changes in tone duration, frequency, intensity, and interstimulus interval (Rosburg, 2003). It is thought to be an index of the early, preattentive sensory memory, most likely only echoic memory (Naatanen, 1992). Most often MMN is used as a measure of a participant’s ability to discriminate linguistic stimuli (e.g., speech sounds with different voice onset time or place of articulation; Naatanen, 1992). ERPs elicited by the standard stimuli are subtracted from the ERPs for the deviants. The resulting difference wave is typically used in the statistical analyses. The subtracted component generally displays onset latency as short as 50 msec and a peak latency of 100 msec to 200 msec (Naatanen, 1992).

Rosburg (2003), using MEG, reported that dipole locations for the MMN were located within the temporal lobe (see also Alho, 1995), but exact placement varied, based on the stimulus properties. Dipoles for frequency and duration deviants differed significantly from each other in the anterior–posterior direction and were located significantly inferior in comparison to the intensity deviants. Liebenthal et al. (2003) recorded fMRI and ERP data simultaneously and noted increased blood-oxygenation-level-dependent signal in the right superior temporal gyrus and the right superior temporal plane.

Visual

The MMN for visual stimuli has been difficult to obtain (Fabiani et al., 2000), although there is some evidence that it can be captured with optical imaging techniques. Source localization techniques suggest the involvement of the pri-
mary visual cortex or adjacent areas or both (Gratton, 1997; Gratton, Gabiani, Goodman-Wood, & DeSoto, 1998).

P3

At this time, the P3 is the most extensively researched ERP component. It was first identified by Sutton, Tueting, Zubin, and John (1965) in a cuing paradigm as a pronounced positivity over parietal areas (one third of the distance from Cz to the external auditory meatus; bilateral earlobe reference) that occurred in response to an unexpected stimulus type approximately 300 msec after stimulus onset. This effect was present for auditory (clicks) and visual (light flashes) stimuli. Currently, the most typical paradigm for eliciting the P3 component, also known as P3b, is the oddball paradigm where a target stimulus is presented infrequently among more common distracter stimuli. However, Polich, Eichen, and Collins (1994) noted that P3 could also be elicited in a single stimulus paradigm where a rare stimulus is presented randomly in time. Unlike the MMN paradigms, for a P3 to be elicited, the participant must pay attention and respond (overtly or covertly) to the stimuli. In addition, the ratio of target to distracter stimuli must be low (the fewer targets, the larger the peak). P3 amplitude is affected by attention (Overtoom et al., 1998; Strandburg et al., 1996), stimulus probability, stimulus relevance, and the amount of processing resources available, such as in single versus dual tasks (Donchin, Miller, & Farwell, 1986), the quality of selection (Johnstone, Barry, Anderson, & Coyle, 1996), and attention allocation (Jonkman et al., 2000). Polich (1990) indicated that length of the interstimulus interval could also affect the amplitude independently of stimulus probability, with shorter intervals resulting in a larger P3. P3 latency was reported to vary with stimulus complexity (McCarthy & Donchin, 1981), effectiveness of selection (Robaey, Breton, Dugas, & Renault, 1992; Taylor, Sunohara, Khan, & Malone, 1997), and sustained attention (Strandburg et al., 1996).

The issue of modality effects on the P3 is not very clear. Some findings suggest that P3 characteristics are not identical across various modalities (e.g., Johnson, 1989, 1993). Katayama and Polich (1999) used a three-stimulus oddball paradigm and reported larger P3 amplitudes for visual stimuli, whereas the auditory stimuli were associated with shorter latencies. Similar findings have been noted by others using the traditional oddball design (e.g., Picton, Stuss, Champagne, & Nelson, 1984; Simson et al., 1977). Nevertheless, the general consensus in the field is that stimulus modality has no significant effect on the P3 amplitude and latency (Picton et al., 1984; Simson et al., 1977) or scalp topography (Polich, Eichen, & Collins, 1996).

The functional interpretation of the classic P3 is diverse—some view it as an indicator of memory updating (Donchin & Coles, 1988), whereas others believe that it reflects a combination of processes that vary by task and situation, including
more elaborate active stimulus discrimination and response preparation (Verleger, 1988). P3 latency is assumed to reflect the duration of stimulus evaluation (Donchin & Coles, 1988). The P3 component has also attracted attention in clinical studies. Because P3 amplitude varies with the amount of attention paid to the stimuli, this component is widely studied in populations with attention deficits (e.g., ADHD) where it is interpreted to reflect information regarding various attentional functions. Further, P3 latency was reported to be related to cognitive abilities, with shorter latencies associated with better performance (Emmerson, Dustman, Shearer, & Turner, 1990; Polich & Martin, 1992).

Sources of the P3 are not clearly identified, but intracranial recordings indicate that at least some are expected to be in the medial temporal lobe (Neshige & Luders, 1992; O’Donnel et al., 1993), including the hippocampal region (Paller, McCarthy, Roessler, Allison, & Wood, 1992), parahippocampal gyrus, amygdala, or thalamus (Katayama, Tsukiyama, & Tsubokawa, 1985). Tarkka, Stokic, Basil, and Papanicolaou (1995) investigated the possible sources and reported that selecting only one region (e.g., hippocampus or thalamus) resulted in poor BESA model fit, but combining the different locations produced a better model. Their findings are consistent with earlier observations using MEG analyses that located sources in the floor of the Sylvian fissure (superior temporal gyrus) as well as deeper sources in the thalamus or hippocampus or both (Papanicolaou, Baumann, & Rogers, 1992; Rogers et al., 1991). Lesion and BESA data suggest that at least some of the P3 generators are located deep within the temporoparietal area or in the temporoparietal junction (Hegerl & Frodl-Bauch, 1997; Knight, Scabini, Woods, & Clayworth, 1989).

P3a

A variant of P3, known as P3a, appears to have a different scalp distribution with frontal maximum and slightly shorter latency for stimuli in visual (Courchesne, Hillyard, & Galambos, 1975; right mastoid reference) versus auditory (Knight, 1984) and somatosensory (Yamaguchi & Knight, 1991) modalities. This frontal P3a occurs when a participant is not required to actively respond to the targets (N. K. Squires et al., 1975) or when a novel stimulus is added to the standard two-stimulus oddball paradigm (Coull, 1998).

Frontal P3a is assumed to reflect involuntary attention as well as inhibition. In Go–No-go paradigms, P3a was larger in amplitude in No-go than Go conditions (maximal at parietal sites for Go; Bokura et al., 2001; Fallgatter & Strick, 1999; Kopp et al., 1996). Regarding its neural substrate, Bokura and colleagues used the LORETA approach and identified sources of P3a in the medial parietal lobe (317 msec) and in the left superior prefrontal cortex (651 msec) for Go trials; for the No-Go trials the sources originated in the left lateral orbitofrontal cortex (365
msec; similar to Casey et al., 1997; Weisbrod, Kiefer, Marzinzik, & Spitzer, 2000). Underscoring the prefrontal cortex connection, P3a can be reduced by lesions to frontal cortex (Knight, 1991). Using BESA, Hegerl & Frodl-Bauch (1997) located auditory P3a near the superior temporal plane in both hemispheres. Similar to the P3b results, these findings were highly reliable as evidenced by almost identical replication across two separate data sets from 54 adults collected 3 weeks apart.

**N400**

This negative component occurs approximately 400 msec after stimulus onset and is usually associated with visual and auditory sentence comprehension tasks. This phenomenon was first identified by Kutas and Hillyard (1980a, 1980b) in a paradigm where words of a sentence were visually presented one after another at fixed intervals. The last word of the sentence was syntactically appropriate and either congruous (“He took a sip from the water fountain”) or incongruous (“He took a sip from the transmitter”) with the rest of the sentence. The incongruous words elicited a larger amplitude N400 response than the congruous words. Further, the amplitude of the N400 was correlated with the degree of incongruency of the sentence and the final word. Kutas and Hillyard (1983) reported that the N400 effect was elicited for semantic, but not syntactic, deviations from expected endings. The N400 is also elicited in semantic word pairs (Rugg, 1985), semantic priming tasks (Bentin, McCarthy, & Wood, 1985; Ruz, Madrid, Lupiáñez, & Tudela, 2003), and matching semantic material to visual displays (Huddy, Schweinberger, Jentzsch, & Burton, 2003).

The amount of attention necessary to produce the N400 and the precise cognitive processes involved remain unclear (Osterhout & Holcomb, 1995). Holcomb (1988) reported that the N400 is more robust when attention is required but can occur even when participants are not attending to the stimuli. However, Bentin, Kutas, and Hillyard (1995) observed that in a dichotic listening task, the N400 was absent for material presented in the unattended ear. The amount of effortful semantic processing required is also unclear. Kutas and Hillyard (1993) identified an N400 effect in tasks that did not require any semantic processing, whereas Chwilla, Brown, and Hagoort (1995) found no N400 when the attention was not directed to the meaning of the stimuli (see also Ruz et al., 2003). One consistent finding is that N400 can be elicited by anomalies in language presented in various modalities, including auditory presentation (Connolly & Phillips, 1994; Connolly, Phillips, Steward, & Brake, 1992; Holcomb, Coffey, & Neville, 1992; McCallum, Farmer, & Pocock, 1984) and American Sign Language (Neville, 1985). However, N400 did not occur when participants were presented with anomalies in music, which is believed to involve a structure similar to language (Besson, Faïta, & Requin, 1994; Besson & Macar, 1986). More recently N400 response was also
noted in response to incongruent solutions for mathematical multiplication problems (Niedeggen, Rosler, & Jost, 1999).

For both visual and auditory stimuli, the N400 is larger over the parietal and temporal regions in the right hemisphere (Atchley & Kwasny, 2003—linked mastoids; Holcomb et al., 1992—left mastoid reference). N400 latency varies with the modality of the task, with visual stimuli resulting in an earlier peak relative to the auditory presentation (475 msec vs. 525 msec), but only over the temporal, anterior temporal, and frontal sites (Holcomb et al., 1992). Further, the shortest latency in the visual modality was noted over the parietal and temporal sites, whereas in the auditory modality it was recorded over parietal and occipital areas (Holcomb et al., 1992). Hemisphere asymmetry for latency measures was noted only in the visual modality where N400 occurred earlier over the left hemisphere (Holcomb et al., 1992).

The N400 is likely to arise from multiple generators that are segregated both functionally (Nobre & McCarthy, 1994) and spatially (Halgren et al., 1994; McCarthy, Nobre, Bentin, & Spencer, 1995). Results of intracortical recordings point to the parahippocampal anterior fusiform gyrus (McCarthy et al., 1995; Nobre, Allison, & McCarthy, 1994) or medial temporal structures near the hippocampus and amygdala (Halgren et al., 1994; Nobre & McCarthy, 1995; Smith, Stapleton, & Halgren, 1986), whereas others suggest locations in the lateral temporal region (Simos, Basile, & Papanicolaou, 1997; MEG).

P600

This component has two functionally different interpretations, one associated with memory processes and another related to language. Although both peaks have roughly similar topographies, they appear to have different brain sources.

Some researchers proposed that the P600 component, especially the one associated with language, is a delayed variant of the P3 because both peaks have relatively similar scalp distributions and are both sensitive to probability manipulations (e.g., Coulson, King, & Kutas, 1998; Gunter, Stowe, & Mulder, 1997). However, Osterhout, McKinnon, Bersick, and Corey (1996) reported evidence that the P3 and P600 have sufficiently different scalp topography, are differentially sensitive to manipulations of stimuli and tasks, and have additive effects when they are co-elicited (see also Osterhout & Hagoort, 1999).

P600 AND MEMORY

This version of P600 is typically observed in recognition–recall memory paradigms and is often referred to as an old–new effect. Typically, the peak onsets at
400 msec and continues for approximately 400 msec to 600 msec (Allan, Wilding, & Rugg, 1998). Maximum amplitudes are noted over the left temporoparietal regions in studies using linked mastoid references (Donaldson & Rugg, 1999; Rugg, Cox, Doyle, & Wells, 1995) or average reference (Curran, 1999, 2000; Curran & Cleary, 2003).

The P600 old–new effect often co-occurs in time with a frontal N400 effect present over the left frontocentral areas, starting at 300 msec to 500 msec poststimulus and continuing to 1,200 msec and beyond (Allan et al., 1998; Curran, 1999, 2000; Curran & Cleary, 2003; Wilding & Rugg, 1996). Jordan, Kotchoubey, Groezinger, & Westphal (1995) noted that during the learning phase of a free recall task larger N400 and P600 amplitudes were elicited by items that were later forgotten. However, the two components have different functional interpretations. P600 is assumed to reflect recognition of the stimuli (Allan et al., 1998; Rugg, 1995; Rugg & Doyle, 1992), whereas frontal N400 is associated with stimulus familiarity (Allan et al., 1998; Curran & Cleary, 2003; Duzel, Vargha-Khadem, Heinze, & Mishkin, 2001; Friedman & Johnson, 2000; Guillem, Bieu, & Debruille, 2001; Mecklinger, 2000; Nessler, Mecklinger, & Penney, 2001; Rugg, Cox, Doyle, & Wells, 1995; Rugg, Schloerscheidt, & Mark, 1998; Wilding & Rugg, 1996, 1997a, 1997b).

Numerous studies of recognition memory reported a larger P600 in response to “old” stimuli previously presented to the participant compared to “new” stimuli that were not experienced before (e.g., Rugg & Doyle, 1992; Smith, 1993), whereas the opposite is true for frontal N400 (Johnson, Kreiter, Russo, & Zhu, 1998). The P600 old–new effect also occurs for items that are incorrectly judged as new (Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997). In addition, it is often larger for correctly recognized words than falsely recognized lures (Curran, 2000), and it can be affected by depth of processing (Paller & Kutas, 1992; Paller, Kutas, & McIsaac, 1995; Rugg, Allan, & Birch, 2000; Rugg et al., 1998) and the amount of retrieved episodic information (see Friedman & Johnson, 2000, for a review). Further, the amplitude of the P600 peak increases with better memory performance (Curran & Cleary, 2003; Olichney et al., 2000). A number of experiments have demonstrated that P600 old–new effects could also occur in the absence of intentional retrieval (Curran, 1999; Paller & Kutas, 1992; Paller, Kutas, & McIsaac, 1995). However, some have reported that intentional retrieval resulted in enhanced P600 old–new effects (Badgaiyan & Posner, 1997; Paller & Gross, 1998).

Although most of the P600 studies involve visually presented stimuli, some work has employed auditory stimuli. For example, Curran (unpublished manuscript, cited in Curran & Cleary, 2003, p. 201) noted no difference in the size of the P600 when the words were studied in one modality but tested in another. Similarly, Wilding and Rugg (1996, 1997b) reported the old–new P600 effect after training subjects on auditory stimuli and testing them when the same stimuli were pre-
sented visually. These findings suggest that the component is not modality specific.

Various techniques consistently identified several brain sources for the P600 old–new effect. Using intracranial ERP recordings during continuous recognition tasks, Guillem, N’Kaoua, Rougier, and Claverie (1995) noted P600 responses in prefrontal regions and anterior temporal lobe structures. Further, Guillem, Rougier, and Claverie (1999) reported a large amplitude P600 response in the anterior cingulate gyrus. Similar findings were obtained in studies employing PET and ERP methods. PET data indicated that regional cerebral blood flow in the left posterior hippocampus, left frontal and temporal cortex, and left anterior cingulate were greater during the recognition of deeply processed (sentence generation vs. alphabetic judgment) words (Rugg et al., 1998). Henson, Rugg, Shallice, Josephs, and Dolan (1999) used event-related fMRI imaging and found that during the study period words subsequently recalled versus judged as familiar were associated with increased activity in a posterior left prefrontal region. However, during the memory task, recalled words were associated with enhanced responses in anterior left prefrontal, left parietal, and posterior cingulate regions relative to familiar judgments.

P600: SYNTACTIC POSITIVE SHIFT (SPS)

Kutas and Hillyard (1983) first reported that syntactic anomalies elicited a small early negativity and a small later positivity rather than a standard N400 response. A decade later, two independent research teams identified a specific component, variously referred to as P600 (Hagoort, Brown, & Groothuysen, 1993) or the syntactic positive shift (SPS; Osterhout & Holcomb, 1992). This component typically consists of a slow positive shift, lasting up to 300 msec, which begins approximately 500 msec after word onset and is widely distributed across the scalp with posterior maxima (Brown, Hagoort, & Kutas, 2000). Most researchers working on the P600–SPS reference scalp electrodes to either a single or linked mastoids.

The P600–SPS is typically elicited by various syntactic or morphosyntactic violations (for a review, see Osterhout, McLaughlin, & Bersick, 1997), including violations of agreement (Coulson et al., 1998; Hagoort et al., 1993), phrase structure (Hagoort et al., 1993; Neville, Nicol, Barss, Foster, & Garrett, 1991), subadjacency (Neville et al., 1991), and subcategorization frame (Hagoort & Brown, 2000; Osterhout & Holcomb, 1992). It has also been elicited by syntactically ambiguous sentences (Frisch, Schlesewsky, Saddy, & Alpermann, 2002). The P600–SPS was reported in studies using English (Neville et al., 1991), Dutch (Hagoort et al., 1993), German (Rosler, Friederici, Puts, & Hahne, 1993), and Italian languages (Angrilli et al., 2002). Although it is usually elicited by visually presented written
stimuli, it can also be elicited using naturally produced speech (Friederici, Pfeifer, & Hahne, 1993; Hagoort & Brown, 2000).

The P600–SPS is commonly thought to reflect additional syntactic processing in response to a parsing failure (Friederici & Mecklinger, 1996; Hagoort et al., 1993; Hagoort, Wassenaar, & Brown, 2003; Osterhout, 1994). It is elicited not only by syntactically incorrect sentences that require repair (Coulson et al., 1998; Hagoort et al., 1993; Neville et al., 1991; Osterhout & Mobley, 1995) but also by syntactic anomalies such as garden-path sentences that require reanalysis (Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Osterhout & Holcomb, 1992). The P600–SPS has also been recently shown to occur in response to syntactic ambiguity (Frisch et al., 2002). Münte et al. (1998) challenged the syntactic specificity of the P600–SPS. Examining ERP responses to morphosyntactic, semantic, and orthographic violations, they found that each elicited similar late positivities.

Investigation of the neuroanatomical sources of the P600–SPS, using rapid-presentation event-related fMRI methods, has identified greater activation in the superior parietal cortex and the precuneus and posterior cingulate on the medial surface in response to morphosyntactic violations compared to normal sentences (Kuperberg et al., 2003). Aphasic patients with lesions in basal ganglia failed to display a P600 effect in response to auditory stimuli containing syntactic violations, but they had a clear P3b in response to an oddball paradigm (Frisch, Kotz, & von Cramon, 2003). Another study involving patients with left subcortical lesions restricted to the basal ganglia found a modulated P600–SPS response with a reduced amplitude compared to that of normal individuals (Friederici, von Cramon, & Kotz, 1999). These results suggest that the basal ganglia play a crucial role in the modulation of the syntactic P600, but not in the modulation of the P3b.

CONCLUSIONS

The purpose of this review was to provide a comprehensive summary of the peak characteristics, paradigms, and typical interpretations of the results for the commonly identified ERP components.

From the review, it is evident that a notion of individual peaks reflecting single cognitive processes is a long-outmoded view. In the early years of electrophysiological research, equipment limitations made it very difficult or impossible to record or analyze more than a single peak or to record from more than a few electrode sites. This may have led investigators to conclude that the measured component was the sole indicator of the cognitive process in question. In the interim, decades of research and advances in technology have increasingly demonstrated that each of the ERP components can be elicited by multiple stimuli and paradigms that tap different cognitive processes. This view is consistent with the
common understanding of brain organization—the same structures may participate in different processes to varying degrees at different times.

Further, it is clear that peak characteristics can be affected by the procedures used to record ERPs. Differences in number of trials or length of the intertrial intervals, variations in stimulus intensity or modality can contribute to inconsistent outcomes. Therefore, to increase the chance of successful replication, investigators must routinely report (and review) such details.

In addition, it is our intention to caution researchers about potential problems of interpretation, directly linking the scalp distribution of an ERP component with brain structures located below the specific electrodes. As noted in this review and elsewhere (e.g., Coles & Rugg, 1995), brain sources of the components are often not located immediately below the electrode that recorded the maximum amplitude. In some cases, the sources are not even in the same hemisphere. Development of carbon electrodes as well as brain source analysis software now allows researchers to coregister ERPs with fMRI data, to map ERP components onto brain structures, and to model potential sources of the observed activity across procedures. Therefore, a change in the language used to report electrophysiological results is needed. We propose that investigators guard against using brain structure terminology, such as “frontal regions produced a larger peak” and instead indicate the electrode locations, for example, “electrodes over frontal regions recorded larger amplitudes.” For a researcher to make a claim regarding the source of activity, the method used to determine the source (e.g., MEG, BESA, etc.) must be described.

Finally, combining ERP measures with other behavioral indicators (e.g., response time, number of correct responses, scores on standardized assessments) invariably provides more detailed information concerning the cognitive processes under study. This also provides a means to map the ERP findings onto the extensive behavioral literature that already exists. Such an approach may lead to increased understanding of brain–behavior relations and to development of innovative neurocognitive assessment techniques that may be increasingly sensitive to otherwise less noticeable changes.

ACKNOWLEDGMENTS

This work was supported, in part, from Grant R01HD17860 from the National Institutes of Health; and Grant R215K000023 from the U.S. Department of Education.

We thank Dennis L. Molfese for comments on earlier drafts of this manuscript.

REFERENCES


