An Introduction to the Event-Related Potential Technique

Steven J. Luck

The MIT Press

7 ERP Localization

The ultimate goal of cognitive neuroscience is to understand how the neural circuitry of the brain gives rise to cognitive processes. This is a challenging enterprise, and one of the central difficulties is to measure how specific populations of neurons operate during the performance of various cognitive tasks. The best techniques for measuring activity in specific populations of neurons are invasive and cannot usually be applied to human subjects, but it is difficult to study many aspects of cognition in nonhuman subjects. PET and fMRI provide noninvasive means of localizing changes in blood flow that are triggered by overall changes in neural activity, but blood flow changes too slowly to permit the measurement of most cognitive processes in real time. ERPs provide the requisite temporal resolution, but they lack the relatively high spatial resolution of PET and fMRI. However, ERPs do provide some spatial information, and many investigators are now trying to use this spatial information to provide a measurement of the time course of neural activity in specific brain regions.

The goal of this chapter is to explain how this process of ERP *source localization* works in general and to provide a discussion and critique of the most common source localization techniques. Before I begin, however, I would like to provide an important caveat: I tend to be extremely skeptical about ERP localization, and this chapter reflects that skepticism. Many other researchers are also skeptical about ERP localization (see, e.g., Snyder, 1991), but I am more skeptical than most. Consequently, you should not assume that my conclusions and advice in this chapter represent those of the majority of ERP researchers. You should talk to a broad

range of ERP experts before making up your mind about source localization.

Source localization is very complex, both in terms of the underlying mathematics and the implementation of the procedures, and it would require an entire book to provide a detailed description of the major techniques. This chapter therefore focuses on providing a simple description of the two major classes of techniques—including their strengths and weaknesses—so that you can understand published research using these techniques and decide whether to pursue localization yourself.

My main advice in this chapter is that beginning and intermediate ERP researchers should not attempt to localize the sources of their ERP data. ERP localization is a very tricky business, and doing it reasonably well requires sophisticated techniques and lots of experience. Moreover, the techniques currently in wide use are not completely satisfactory. However, many smart people are working to improve the techniques, and some promising approaches are on the horizon. Thus, you may eventually want to do some source localization, and you will certainly be reading papers that report localization results. The goal of this chapter is therefore to make you an informed and critical observer of source localization efforts rather than a participant in these efforts.

The Big Picture

If you place a single dipole in a conductive sphere, you can use relatively simple equations to predict the precise distribution of observable voltage on the surface of the sphere. This is called the *forward problem*, and it is relatively easy to solve. Voltages summate linearly, which means that the forward problem is also easy to solve for multiple simultaneously active dipoles—the voltage distributions for the individual dipoles are simply added together to derive the distribution for the set of dipoles. The forward problem can also be solved for realistic head shapes.

The problem arises in solving the *inverse problem* of determining the positions and orientations of the dipoles on the basis of the observed distribution of voltage over the scalp. If only one dipole is present, and there is no noise, then it is possible to solve the inverse problem to any desired degree of spatial resolution by comparing forward solutions from a model dipole with the observed scalp distribution and then adjusting the dipole to reduce the discrepancy between the predicted and observed distributions. However, it is not possible to solve the inverse problem if you don't know the number of dipoles (or if the activity is distributed rather than dipolar) because there is no unique solution to the inverse problem in this case. In other words, for any given scalp distribution, there is an infinite number of possible sets of dipoles that could produce that scalp distribution (Helmholtz, 1853; Plonsey, 1963). Thus, even with perfectly noise-free data, there is no perfect solution to the inverse problem.

Several investigators have proposed ways around this uniqueness problem, and their solutions fall into two general categories. One approach is to use a small number of equivalent current dipoles, each of which represents the summed activity over a small cortical region (perhaps 1–2 cm^3), and assume that these dipoles vary only in strength over time; this is the *equivalent current dipole* category of source localization methods. The second category divides the brain's volume (or the cortical surface) into a fairly large number of voxels (perhaps a few thousand), and computes the set of strengths for these voxels that can both explain the observed distribution of voltage over the scalp and satisfy additional mathematical constraints; this is the *distributed source* category.

The Forward Solution

Early forward solutions assumed that the head was a sphere, which makes the computations relatively simple. That is, if we assume

that the head is a sphere and that the head's resistance is homogeneous, it is very easy to compute the voltage created at every point on the scalp by a single dipole or an arbitrarily large set of dipoles. The skull and scalp have a higher resistance than the brain, but this is easily accommodated by a model in which skull and scalp layers cover a homogeneous, spherical brain. Researchers often use a spherical approximation for computing forward solutions because it is easy to generate the model and because they can compute the forward solution very rapidly. However, this model is obviously an oversimplification, and many researchers are now using more detailed models of the head, such as *finite element models.* A finite element model divides the volume of the head into thousands or even hundreds of thousands of individual cubes. The resistance within each cube is assumed to be homogeneous (which is a reasonable approximation given that each cube is very small), but the resistances are different for each cube. In addition, the set of cubes that define the head are not constrained to form a large sphere, but instead take into account the shape of each subject's head (as determined by a structural MRI scan).

Once the researcher has divided the head into a set of cubes and determined the resistances of the individual cubes, simple equations can determine the surface voltages that a dipole in a known location would produce. Given the large number of cubes in a realistic model of the head, it takes a lot of computer power to do this sort of modeling, but as computers become faster, this sort of modeling will fall within the means of more and more researchers. The hard part is to determine the resistances of the individual cubes. The resistances can't actually be measured for a living human subject, but they can be estimated by performing an MRI scan, dividing up the head into regions of different tissue types on the basis of the MRI scan, and using normative resistance values for each tissue type (i.e., resistance values that have been determined invasively from cadavers, from nonhuman animals, and possibly from human neurosurgical patients).

Because finite element models are so computationally intensive, researchers have developed a less intensive variant, the *boundary element model*. This approach takes advantage of the fact that the brain itself has a relatively constant resistance, and most of the action is in the boundaries between the brain, the skull, and the scalp. The model therefore consists of boundary surfaces for these tissues, with the assumption that the resistance within a tissue is constant throughout the extent of that tissue. As in finite element models, the boundaries are estimated from structural MRI scans and the conductivities of each tissue type are based on normative values.

Equivalent Current Dipoles and the BESA Approach

The *brain electrical source analysis* (BESA) technique is the prototypical example of an equivalent current dipole technique. In addition, BESA has been the most commonly used method for localizing ERP sources over the past twenty years, in part because it is relatively simple and inexpensive (both in terms of computing resources and money). Thus, this section will focus on the BESA technique. The interested reader may also want to learn about the MUSIC (multiple signal characterization) technique, which provides a more sophisticated approach to computing equivalent current dipole solutions (see Mosher, Baillet, & Leahy, 1999; Mosher & Leahy, 1999; Mosher, Lewis, & Leahy, 1992).

BESA is based on the assumption that the spatiotemporal distribution of voltage can be adequately modeled by a relatively small set of dipoles (< 10), each of which has a fixed location and orientation but varies in magnitude over time (Scherg, Vajsar, & Picton, 1989; Scherg & von Cramon, 1985). Each dipole has five major parameters, three indicating its location, and two indicating its orientation. A magnitude parameter is also necessary, but this parameter varies over time and is treated differently from the location and orientation parameters.

The Essence of BESA

The BESA algorithm begins by placing a set of dipoles in an initial set of locations and orientations, with only the magnitude being unspecified. The algorithm then calculates a forward solution scalp distribution for these dipoles, computing a magnitude for each dipole at each point in time such that the sum of the dipoles yields a scalp distribution that fits, as closely as possible, the observed distribution for each point in time.

The scalp distributions from the model are then compared with the observed scalp distributions at each time point to see how well they match. The degree of match is quantified as the percentage of the variance in scalp distribution that is explained by the model; alternatively, it can be expressed as the percentage of unexplained variance (called the *residual variance*). The goal of the algorithm is to find the set of dipole locations and orientations that yields the lowest residual variance, providing the best fit between the model and the data.

This is accomplished in an iterative manner, as shown in figure 7.1. On each iteration, the forward solution is calculated, leading to a particular degree of residual variance, and then the positions and orientations of the dipoles are adjusted slightly to try to reduce the residual variance. This procedure is iterated many times using a gradient descent algorithm so that the positions and orientations will be adjusted in a way that tends to decrease the residual variance with each successive iteration. In the first several iterations, the residual variance drops rapidly, but after a large number of iterations, the residual variance stops declining much from one iteration to the next and the dipole positions and orientations become stable. There are various refinements that one can add, but this is the essence of the BESA technique.

Figure 7.2 shows an example of a BESA solution (from Di Russo et al., 2002). The goal of this source localization model was to characterize the generators of the early visual sensory components. The top of the figure shows the scalp distribution of the ERP response to a checkerboard stimulus in the upper left visual field in various
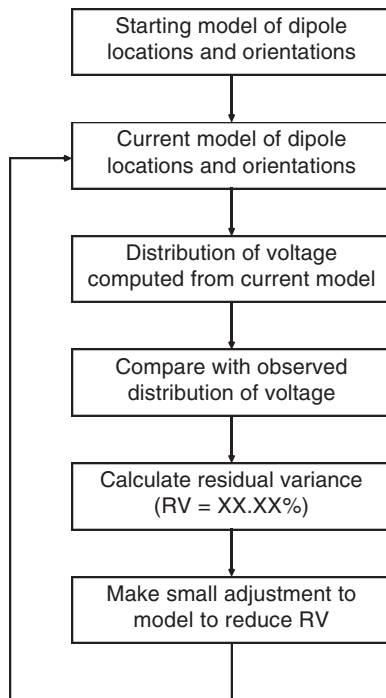
Figure 7.1 Basic procedure of the BESA technique.

time ranges, and the bottom of the figure shows the BESA model that was obtained. Each dipole is represented by a dot showing its location and a line showing its orientation; three different views of the head are shown so that the dipoles can be seen in three dimensions. Each dipole is also associated with a *source wave-form*, which shows how the estimated magnitude for that dipole varies over time.

The Starting Point
Before initiating this iterative procedure, it is necessary to decide how many dipoles to use and what their starting positions will be. These are important decisions, and they will have a very
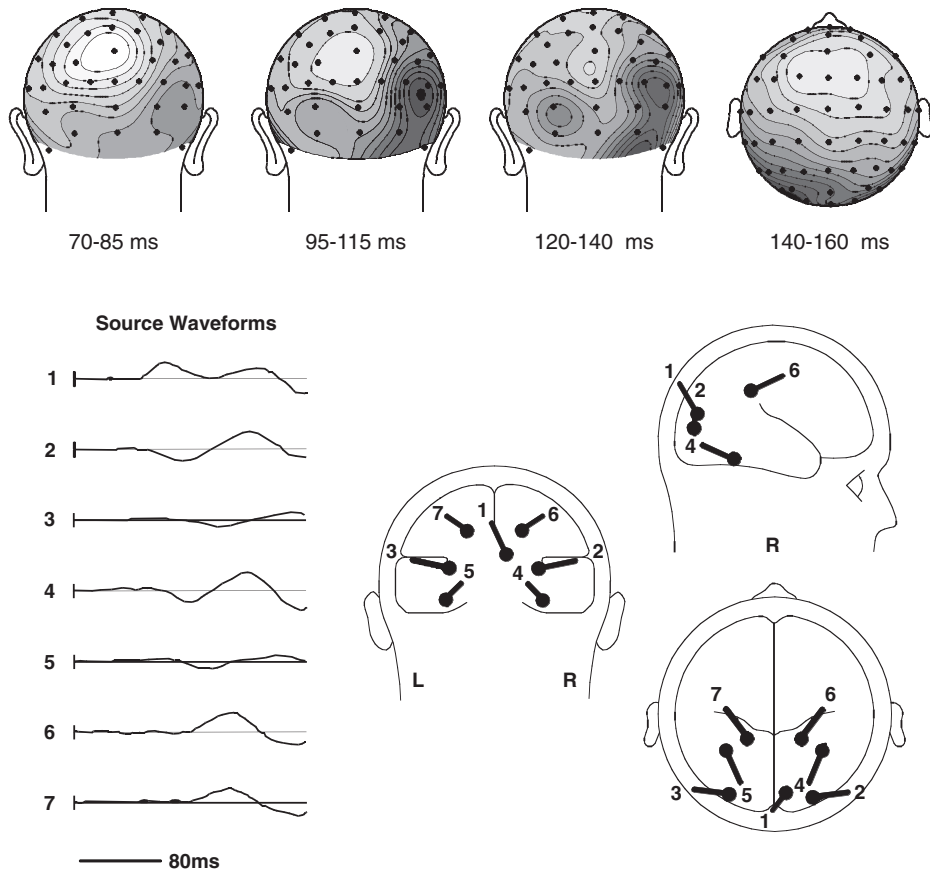
Figure 7.2   Example of an equivalent current dipole model generated using the BESA technique (from the study of Di Russo et al., 2002). The top shows the scalp distributions of voltage measured in various time ranges in response to a checkerboard in the upper left visual field. This time-space-voltage data set is modeled by a set of seven dipoles. The locations and orientations of the dipoles are shown in three different views in the lower right region, and the magnitude of each dipole over time is shown in the lower left region. (© 2002 Wiley-Liss, Inc.) Thanks to Francesco Di Russo for providing electronic versions of these images.

substantial impact on the solution the algorithm ultimately reaches. In fact, the most problematic aspect of the BESA technique is the fact that the user has a lot of control over the results (this is called *operator dependence*); consequently the results may be biased by the user's expectations.

You can use several different strategies to select the number of dipoles. One approach is to use principal components analysis (PCA) to determine how many underlying spatial distributions of activity contribute to the observed distribution of voltage over the scalp. This technique can work well in some cases. However, if the magnitudes of two dipoles are correlated with each other over time, they may be lumped together into a single component. Thus, the number of components identified by PCA can only provide a lower bound on the number of dipoles.

Another strategy is to start by using one or two dipoles to create a model of the early part of the waveform, under the assumption that the response begins in one or two sensory areas. Once a stable solution is reached for the early part of the waveform, the time window is increased and new dipoles are added while the original dipoles are held constant. This procedure is then repeated with a larger and larger time window. This procedure makes some sense, but it has some shortcomings. First, at least a dozen visual areas are activated within 60 ms of the onset of activity in visual cortex (Schmolesky et al., 1998), so representing the early portion of the waveform with one or two dipoles is clearly an oversimplification. Second, some error is likely in any dipole solution, and small errors in the initial dipoles will lead to larger errors in the next set of dipoles, and the location estimates will become increasingly inaccurate as you add more and more dipoles.

A third strategy is to use preexisting knowledge about the brain to determine the number of dipoles. For example, if you use difference waves to isolate the lateralized readiness potential, it would be reasonable to start with the assumption that two dipoles are present, one in each hemisphere.

There are also several strategies that you can use for determining the starting positions and orientations of the dipoles, and different strategies will lead to different results. Unfortunately, most papers using the BESA technique describe the model produced by one or two different starting positions. Almost every time I have read a paper that used the BESA technique, I wished the authors had provided a detailed description of the results that would have been obtained with a wide variety of different starting locations. It would be possible for researchers to do this. Indeed, Aine, Huang, and their colleagues have developed what they call a *multi-start* approach to localization, in which they apply the localization algorithm hundreds or even thousands of times with different starting parameters (Aine et al., 2000; Huang et al., 1998). It is then possible to determine which dipole locations occur frequently in the solutions and are therefore relatively independent of the starting parameters.

Another strategy is to start with dipoles in locations that are based on preexisting knowledge about the brain. The locations could be based on general knowledge (e.g., the location of primary and secondary auditory areas), or they could be based on specific results from previous experiments (e.g., activation centers from a similar fMRI experiment). When using the latter approach, researchers sometimes say that the solution is based on *seeded dipoles*, and some studies have explicitly shown that similar results were obtained with random dipole locations and seeded dipoles (e.g., Heinze et al., 1994). This is, in some ways, a reasonable approach. However, it seems likely to lead to a confirmation bias, increasing the likelihood that the expected results will be obtained even if they are not correct.

Shortcomings of the BESA Approach

The BESA approach has several shortcomings, but the most significant problem is that there is no mathematically principled means

of quantifying the accuracy of a solution. Specifically, in the presence of noise, it is possible for a substantially incorrect solution to have the same (or lower) residual variance than the correct solution. Even with minimal noise, it is possible for a substantially incorrect solution to have a very low residual variance (especially when using more than a few dipoles). One reason for this is that each BESA dipole has five free parameters (plus a time-varying magnitude parameter). Thus, a model with only six dipoles has thirty free parameters, and a relatively large error in one of these parameters can easily be offset by small adjustments in the other parameters, resulting in low residual variance. Even if only one dipole is present, the BESA solution may be inaccurate due to noise in the data and errors in the head model. Without some means of quantifying the likelihood that a solution is correct or even nearly correct, it's hard to use a BESA solution to provide strong support for or against a hypothesis.

The second most significant shortcoming of the BESA technique is the operator dependence of the technique (as mentioned briefly in the previous section). In addition to setting the number and initial positions of the dipoles, a researcher can adjust several other parameters that control how the algorithm adjusts the positions and orientations of the dipole while searching for the configuration with the least residual variance. Moreover, at several points in the process, the researcher makes subjective decisions about adding or deleting dipoles from the solution or changing various constraints on the dipoles. I have seen ERP researchers spend weeks applying the BESA technique to a set of data, playing around with different parameter settings until the solution ''looks right.'' Of course, what ''looks right'' is often a solution that will confirm the researcher's hypothesis (or at least avoid disconfirming it).

Another significant shortcoming of the BESA technique is that it will produce an incorrect solution if the number of dipoles is incorrect. It is difficult or impossible to know the number of dipoles in advance, especially in an experiment of some cognitive

complexity, so this is a significant limitation. Moreover, BESA uses a discrete dipole to represent activity that may be distributed across a fairly large region of cortex, and this simplification may lead to substantial errors.

A Simulation Study

There have been a variety of tests of the accuracy of equivalent current dipole localizations, but they have mostly used only one or two simultaneously active dipoles (see, e.g., Cohen & Cuffin, 1991; Leahy et al., 1998). These simulations are useful for assessing the errors that might be likely in very simple sensory experiments, but they do not provide meaningful information about the errors that might occur in most cognitive neuroscience experiments.

Miltner et al. (1994) performed the most informative simulation study in the context of cognitive neuroscience. This study used BESA's spherical, three-shell head model to simulate a set of dipoles and produce corresponding ERP waveforms from thirty-two electrode sites. From these ERP waveforms, nine participants attempted to localize the dipoles using BESA. The participants consisted of ERP researchers with various levels of expertise with BESA (including three with very high levels of expertise).[1] The participants were told that the data were simulated responses from left somatosensory stimuli that were presented as the targets in an oddball task, and they were given the task of trying to localize the sources. The simulation comprised ten dipoles, each of which was active over some portion of a 900-ms interval. White noise was added to the data to simulate the various sources of noise in real ERP experiments. The simulation included two dipoles in left primary somatosensory cortex (corresponding to the P100 wave and an early portion of the N150 wave), a mirror-symmetrical pair of dipoles in left and right secondary somatosensory cortex, midline dipoles in prefrontal and central regions, and mirror-symmetrical pairs of dipoles in medial temporal and dorsolateral prefrontal regions.

This is a fairly large set of dipoles,[2] but the participants' task was made easier by at least seven factors: (1) the solution included several dipoles that were located exactly where they would be expected (e.g., the primary and secondary somatosensory areas); (2) three of the dipole pairs were exactly mirror-symmetrical (which matches a typical BESA strategy of assuming mirror symmetry at the early stages of the localization process); (3) the spherical BESA head model was used to create the simulations, eliminating errors due to an incorrect forward solution; (4) the temporal overlap between the different dipoles was modest (for most of the dipoles, there was a time range in which only it and one other dipole or mirror-symmetrical dipole pair were strongly active); (5) with the exception of the two dipoles in primary somatosensory cortex, the dipoles were located fairly far away from each other; (6) the white noise that was added was more easily filtered out than typical EEG noise and was apparently uncorrelated across sites; and (7) the simulation used discrete dipoles rather than distributed regions of activation.

Despite the fact that the simulation perfectly matched the assumptions of the BESA technique and was highly simplified, none of the participants reached a solution that included all ten dipoles in approximately correct positions. The number of dipoles in the solutions ranged from six to twelve, which means that there were several cases of missing dipoles and/or spurious dipoles. Only two of the nine participants were able to distinguish between the midline prefrontal and midline central dipoles, and the other seven participants tended to merge them into a single dipole even though the actual dipoles were approximately 5 cm apart.

Across all dipoles that appeared to be localized by the participants, the average localization error was approximately 1.4 cm, which doesn't sound that bad. However, this was a simplified simulation based on the BESA head model, and the errors with real data are likely to be greater. Moreover, there were many cases in which an individual dipole's estimated location was 2–5 cm from the actual dipole's location, and the mean errors across dipoles for

individual participants were as high as 2 cm. To be fair, however, I should note that most of the participants provided a reasonably accurate localization of one of the two primary somatosensory dipoles, the secondary somatosensory dipoles, the medial temporal lobe dipoles, and the dorsolateral prefrontal dipoles. But each of the nine participants had at least one missing dipole, one spurious dipole, or one mislocalization of more that 2 cm.

From this study, we can draw two main conclusions. First, in this highly simplified situation, dipoles were often localized with a reasonable degree of accuracy, with an average error of 1–2 cm for most of the participants (relative to a 17-cm head diameter). Thus, when reality does not deviate too far from this simplified situation, the BESA technique can potentially provide a reasonable estimate of the locations of most of the dipoles most of the time. However, some of the simplifications seem quite far from reality, so it is entirely possible that average errors will be considerably larger with most real data sets.

The second main conclusion is that any single dipole in a given multiple-dipole BESA model has a significant chance of being substantially incorrect, even under optimal conditions. Dipoles may be mislocalized by several centimeters or completely missed; multiple dipoles may be merged together, even if they are fairly far apart; and spurious dipoles may be present in the model that correspond to no real brain activity. Thus, even if the average error is only 1–2 cm for most dipoles, this simulation suggests that BESA solutions for moderately complex data sets may typically contain at least one missing dipole, one spurious dipole, or one 2–5 cm localization error. And the accuracy of the technique is presumably even worse for real data sets that deviate from the simplifications of this simulation.

Although the BESA technique has been widely used over the past twenty years, most ERP researchers now appreciate its limitations. There is a clear trend away from this technique and toward more sophisticated equivalent current dipole approaches and distributed source approaches.

Distributed Source Approaches

General Approach

Instead of using a small number of equivalent current dipoles to represent the pattern of neural activity, it is possible to divide the brain up into a small number of voxels and find a pattern of activation values that will produce the observed pattern of voltage on the surface of the scalp. For example, you could divide the surface of the brain into a hundred little cubes. Each cube would contain three dipoles, one pointing upward, one pointing forward, and one pointing laterally (a single dipole of an arbitrary orientation can be simulated by varying the relative strengths of these three dipoles). You could then find a pattern of dipole strengths that would yield the observed distribution of voltage on the surface of the head. This would provide you with an estimate of the distribution of electrical activity throughout the brain.

The problem with this approach is that even this relatively coarse parcellation of the brain requires that you compute 300 different dipole strengths. That is, your model has 300 free parameters to be estimated. Generally speaking, you need at least as many independent data points as you have free parameters, and even if you have voltage measurements from 300 electrodes, they are contaminated by noise and are not independent of each other. Consequently, there are many different sets of strengths of the 300 dipoles that could produce the observed ERP scalp distribution. This is the problem of *nonuniqueness*. And it would get even worse if we wanted to divide the brain into even smaller voxels.

Cortically Constrained Models

Researchers have developed several strategies to avoid the nonuniqueness problem. One strategy is to reduce the number of dipoles by assuming that scalp ERPs are generated entirely by

currents generated in the cerebral cortex, flowing perpendicular to the cortical surface (which is a reasonable assumption in most cases). Instead of using a set of voxels that fills the entire volume of the brain, with three dipoles per voxel, this approach uses structural MRI scans to divide the cortical surface into hundreds or thousands of small triangles, each with a single dipole oriented perpendicular to the cortical surface. This *cortically constrained* approach dramatically reduces the number of free parameters in the model (although some error may be introduced by inaccuracies in the cortical surface reconstruction). The result is a model of the distribution of electrical activity over the cortical surface.

Figure 7.3 illustrates this approach, with a slice through a cartoon brain that shows the cortical surface and recording electrodes for the left hemisphere. The cortical surface has been divided into a number of small patches, and each patch is treated as a dipolar current source pointing perpendicular to the cortical surface. This reduces the number of dipole locations and orientations compared to dividing the entire volume of the brain into voxels, each of which contains three dipoles. However, the number of dipoles needed in a real experiment is still very large (usually in the hundreds or thousands), and there is still no unique pattern of dipole strengths that can account for the observed distribution of voltage on the scalp. That is, the use of a cortically constrained model reduces the number of internal patterns of activity that could explain the observed distribution of voltage over the scalp, but it does not bring the number all the way to one (i.e., to a unique solution).

The non-uniqueness problem in cortically constrained models can be appreciated by considering sources 15 and 16 in figure 7.3. These sources are almost perfectly parallel to each other, but they are inverted in orientation with respect to each other (i.e., the outer surface of the cortex points downward for source 15 and upward for source 16). This is a common occurrence given the extensive foldings of the human cerebral cortex. The non-uniqueness problem occurs because any increase in the magnitude of source 15
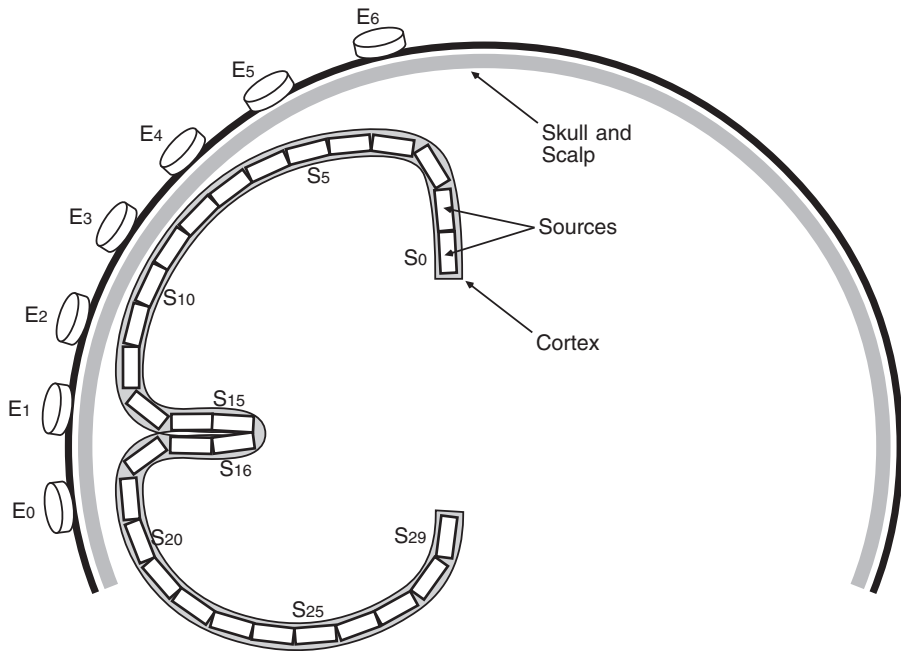
Figure 7.3 Example of the electrical sources and measurement electrodes used by cortically constrained distributed source localization methods. The figure shows a coronal section through the brain. The cortex is divided into a large number of patches that are assumed to be the electrical sources (labeled S_0–S_{29} here). Each source is modeled as a dipole that is centered in the corresponding cortical patch and oriented perpendicular to the patch. The voltage corresponding to each source propagates through the brain, skull, and scalp to reach the recording electrodes (labeled E_0–E_6), and the voltages from the different sources simply sum together.

can be cancelled by an increase in the magnitude of source 16, with no change in the distribution of voltage on the surface.

The Minimum Norm Solution

To get around the nonuniqueness problem (for both whole-brain and cortically constrained models), Hämäläinen and Ilmoniemi (1984) proposed adding an additional constraint to the system. This constraint is based on the fact that the cancellation problem—as

exemplified by sources 15 and 16 in figure 7.3—allows the magnitudes of nearby sources to become huge without distorting the distribution of voltage on the scalp. These huge magnitudes are a biologically unrealistic consequence of the modeling procedure, and it therefore makes sense to eliminate solutions that have huge magnitudes. Thus, Hämäläinen and Ilmoniemi (1984) proposed selecting the one solution that both produces the observed scalp distribution and has the minimum overall source magnitudes. This is called the *minimum norm* solution to the problem of finding a unique distribution of source magnitudes.

One shortcoming of the minimum norm solution is that it is biased toward sources that are near the surface, because a larger magnitude is necessary for a deep source to contribute as much voltage at the scalp as a superficial source. However, this problem can be solved by using a *depth-weighted minimum norm* solution that weights the magnitudes of each source according to its depth when finding the solution with the minimum overall source magnitudes.

Other researchers have proposed other types of minimum norm solutions that reflect different constraints. The most widely used of these alternatives is the low-resolution electromagnetic tomography (LORETA) technique, which assumes that the voltage will change gradually (across the volume of the brain or across the cortical surface) and selects the distribution of source magnitudes that is maximally smooth (Pascual-Marqui, 2002; Pascual-Marqui et al., 2002; Pascual-Marqui, Michel, & Lehmann, 1994). The smoothness constraint may be reasonable in many cases, but sharp borders exist between adjacent neuroanatomical areas, and these borders would sometimes be expected to lead to sudden changes in cortical current flow. Indeed, if an experimental manipulation is designed to activate one area (e.g., V3) and not an adjacent area (e.g., V4), then the goals of the experiment would be incompatible with an assumption of smoothness. On the other hand, gradual changes in activity are probably the norm within a neuroanatomical area, so the smoothness constraint may be appropriate in many cases. Note

also that, because of its smoothness constraint, the LORETA technique is appropriate only for finding the center of an area of activation and not for assessing the extent of the activated area. In contrast, nothing about the original and depth-weighted minimum norm solutions will prevent sharp borders from being imaged.

It is possible to combine empirical constraints with these mathematical constraints. For example, Dale and Sereno (1993) describe a framework for using data from functional neuroimaging to provide an additional source of constraints that can be combined with the minimum norm solution (see also George et al., 1995; Phillips, Rugg, & Friston, 2002; Schmidt, George, & Wood, 1999). Like the LORETA approach, this approach has the advantage of using biological information to constrain which solution is chosen. However, it is easy to conceive of situations in which an fMRI effect would not be accompanied by an ERP effect or vice versa (see Luck, 1999), so the addition of this sort of neuroimaging-based constraint may lead to a worse solution rather than a better one. It may also produce a confirmation bias: when you use fMRI data to constrain your ERP localization solution, you're increasing the likelihood of finding a match between the ERP data and the fMRI data even if the ERP is not generated at the locus of the fMRI BOLD signal. A simulation study suggested that the most problematic situation arises when an ERP source is present without a corresponding fMRI source (Liu, Belliveau, & Dale, 1998). This study also indicated that the resulting distortions are reasonably small if the source localization algorithm assumes a less-than-perfect correspondence between the ERP and fMRI data. It remains to be seen whether the use of fMRI data to probabilistically constrain ERP source localization leads to substantial errors when applied to real data.

Unlike equivalent source dipole approaches, minimum norm-based techniques will always find a unique solution to the inverse problem, and they do it largely automatically. However, a unique and automatic solution is not necessarily the correct solution. The correctness of the solution will depend on the correctness of the

assumptions. As Ilmoniemi (1995) discussed, an approach of this type will provide an optimal solution if its assumptions are valid, but different sets of assumptions may be valid for different data sets. For example, if one applies the LORETA technique to the three-dimensional volume of the brain without first creating a model of the cortical surface, the smoothness assumption will almost certainly be violated when areas that are distant from each other along the cortical surface abut each other due to the folding pattern of the cortex. But if the smoothness constraint is applied along the reconstructed 2-D cortical surface, then this assumes that subcortical regions do not contribute to the data, which may be incorrect.

The Added Value of Magnetic Recordings

As described in chapter 1, the EEG is accompanied by a magnetic signal, the MEG, and event-related electrical potentials (ERPs) are accompanied by event-related magnetic fields (ERMFs). Because the skull is transparent to magnetism, it does not blur the MEG signal, and this leads to improved spatial resolution for MEG recordings. Another benefit of MEG is that, because magnetism passes unimpeded through the head, MEG/ERMF localization does not require a model of the conductances of the head; it simply requires a model of the overall shape of the brain. Thus, it can be advantageous to apply localization techniques to ERMFs rather than ERPs.

ERMF localization faces the same non-uniqueness problem as ERP localization, but combining ERP and ERMF data provides a new set of constraints that can aid the localization process. The main reason for this is that the voltage field and the magnetic field run in different directions, and they therefore provide complementary information. As figure 7.4A illustrates, the magnetic field runs in circles around the current dipole. When the current dipole is oriented in parallel to the skull, the magnetic field exits the skull on one side of the dipole and reenters the skull on the other side (figure 7.4B). The strength of the magnetic field varies as a function
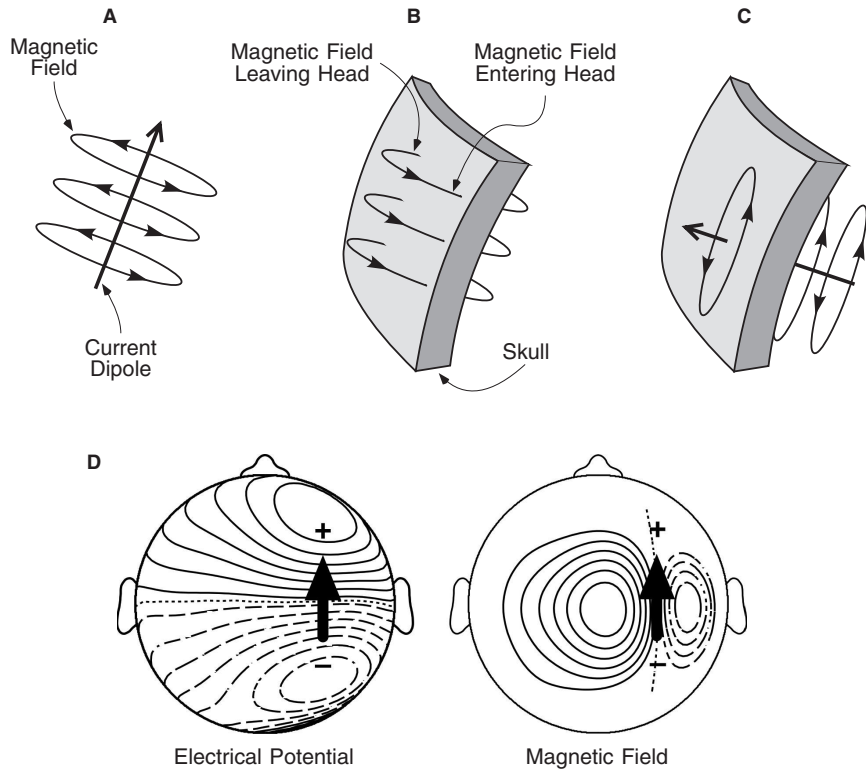
Figure 7.4   Relationship between an electrical dipole and its associated magnetic field. An electrical dipole has a magnetic field running around it (A), and when the dipole is roughly parallel to the surface of the head, the magnetic field leaves and reenters the head (B). If the dipole is oriented radially with respect to the head, the magnetic field does not vary across the surface of the head (C). When a dipole runs parallel to the surface of the head (represented by the arrow in D), there is a broad region of positive voltage at the positive end (solid lines) and a broad region of negative voltage at the negative end (dashed lines), separated by a line of zero voltage (represented by the dotted line). The magnetic field, in contrast, consists of magnetic flux leaving the head on one side of the dipole (solid lines) and reentering the head on the other side (dashed lines), separated by a line of no net flux (dotted line). Thanks to Max Hopf for providing the electrical and magnetic distributions shown in (D).

of distance from the dipole, just like the strength of the electrical field, but the voltage distribution is broader due to the blurring of the scalp (figure 7.4D). In addition, the magnetic and electrical distributions are oriented at 90 degrees with respect to each other. As figure 7.4D shows, the positive and negative electrical potentials appear at the positive and negative ends of the dipole, and the line of zero voltage runs perpendicularly through the center of the dipole. The efflux and influx of the magnetic field, in contrast, occur on the left and right sides of the dipole, and the zero flux line runs in parallel with the orientation of the dipole.

MEG fields differ from EEG fields in another key way as well. If the current dipole is perfectly perpendicular to the skull, as in figure 7.4C, the magnetic field does not exit and reenter the head, and it is essentially invisible. As the dipole tilts from perpendicular toward parallel, a recordable magnetic field begins to appear again. In contrast, a large and focused voltage will be present directly over a perpendicular dipole. A dipole near the center of the head will act much like a perpendicular dipole, generating a reasonably large voltage on the surface of the scalp that is accompanied by a magnetic field that does not exit and reenter the head and is therefore effectively invisible. Thus, magnetic signals are largest for superficial dipoles that run parallel to the surface of the skull, and fall off rapidly as the dipoles become deeper and/or perpendicularly oriented, but voltages do not fall off rapidly in this manner.

The different effects of dipole depth and orientation on electrical and magnetic signals provide an additional set of constraints on source localization solutions. In essence, there are many internal source configurations that can explain a given electrical distribution, and there are also many internal source configurations that can explain a given magnetic distribution. But there will be far fewer configurations than can explain both the electrical distribution and the magnetic distribution. Consequently, the combination of magnetic and electrical data is substantially superior to either type of data alone. The main drawbacks of combining magnetic

and electrical data compared to using magnetic data alone are that (a) a more complex head model is needed for the electrical data, and (b) some effort is required to ensure that the electrical and magnetic data are in exactly the same spatial reference frame.

Can We Really Localize ERPs?

Each of the source localization techniques described in this chapter has shortcomings. Of course, any scientific technique has limitations and shortcomings, but the shortcomings of source localization techniques are fundamentally different from the shortcomings of other techniques for localization of function. This section will explore these differences and consider a new approach that seems more promising.

Source Localization as Model Fitting

To understand the essence of ERP source localization, it is useful to compare it with a "true" neuroimaging technique, such as PET. In the most common PET approach, radioactively labeled water molecules travel through the bloodstream, where their diffusion properties are straightforward. Consequently, the number of radioactive molecules in a given volume of the brain can be directly related to the flow of blood through that part of the brain. When a labeled water molecule decays, it gives off a positron, which travels a known distance (or distribution of distances) before colliding with an electron. This collision leads to a pair of annihilation photons that travel in opposite directions along the same line. When these high-intensity photons are picked up simultaneously by two detectors within the ring of detectors around the subject, there is a high likelihood that they were generated somewhere along the line between the two detectors, and the decaying isotope is known to have been within a certain distance from this line. Thus, by combining the known physics of radiation with various probability distributions, one can directly compute the maximum likelihood

location of the radioactively labeled water molecules and the margin of error of this location. The story is analogous, although more complicated, for fMRI.

Because the ERP localization problem is underdetermined, mainstream ERP localization techniques employ a different approach. That is, they do not simply compute the maximum likelihood location of an ERP source, along with a margin of error, on the basis of the physics of electricity and magnetism. Instead, ERP localization techniques generate *models* of the underlying distribution of electrical activity, and these models are evaluated in terms of their ability to satisfy various constraints. The most fundamental constraint, of course, is that a given model must recreate the observed distribution of voltage over the surface of the head. However, a correct model may not fit the data exactly, because noise in the data distorts the observed distribution. Consequently, any internal configuration that is, say, 95 percent consistent with the observed scalp distribution might be considered acceptable. Unfortunately, there will be infinitely many internal configurations that can explain an observed scalp distribution, especially when only a 95 percent fit is required.

Additional constraints are then added to select one internal configuration from the many that can explain the observed scalp distribution. Each source localization technique embodies a different set of these additional constraints, and the constraints can be either mathematical (as in the use of the minimum norm) or empirical (as in the use of fMRI data to constrain ERP localizations). The most straightforward empirical constraint is the use of structural MRI scans to constrain the source locations to the cortical surface. However, this alone does not lead to a unique solution (and it may not always be the case that all scalp ERP activity arises from the cortex). Researchers therefore add other constraints, but there is usually no way of assessing whether these constraints are correct and sufficient.

The bottom line is that ERP localization leads to a *model* of the internal configuration of electrical activity, not a *measurement*

of the internal distribution of electrical activity. In contrast, PET and fMRI provide measurements and not merely models. PET, for example, provides a measurement of the internal distribution of radioactively labeled blood. This measurement is derived from more basic measurements, but that is true of most sophisticated scientific measurements. And although the PET measurements are not error-free, this is true of any measurement, and one can specify the margin of error. It is more difficult to describe exactly what the BOLD signal reflects in fMRI, but the location of this signal is measured with a known margin of error. In contrast, one cannot use surface electrodes to measure the distribution of internal electrical activity with a known margin of error.

People occasionally ask me how accurately ERPs can be localized, hoping for a quantification of accuracy that they can compare with the accuracy of PET and fMRI. My response is that the accuracy of ERP localization is simply undefined. That is, in the absence of any constraints beyond the observed scalp distribution, radically different distributions of internal electrical activity would produce the observed scalp distribution, and the margin of error is essentially the diameter of the head.

Once constraints are added, some source localization approaches could, in principle, quantify the margin of error. For example, it would be possible to state that the estimated center of a region of activation is within X millimeters of the actual center of the activated region. Or it would be possible to state that the amount of estimated current flow within each patch of cortex is within Y percent of the actual current flow. I've never seen anyone do this in the context of a serious experiment, but it would be an extremely useful addition to the source localization techniques. However, the margin of error that could be specified in this manner would be meaningful only if the constraints of the model were fully adequate and the only sources of error arose from noise in the ERP data (and perhaps errors in specifying the head model). If the constraints were insufficient, or if they reflected incorrect assumptions about the underlying neural activity, then the margin of error

would be meaningless. Thus, the source localization techniques that are currently in widespread use do not, in practice, provide a meaningful estimate of the margin of error.

Probabilistic Approaches

The commonly used source localization techniques attempt to find a single pattern of internal electrical activity that best explains the observed scalp distribution (along with satisfying other implicit or explicit constraints). When I read a paper that reports source localization models, I always wonder what other distributions would fit the data as well as, or almost as well as, the reported solution. Are all reasonable solutions similar to the reported solution? Or are there other solutions that are quite different from the reported solution but fit the data and the constraints almost as well? After all, the presence of noise in the data implies that the correct solution will not actually fit the data perfectly, so a solution that explains only 97 percent of the variance may be closer to the correct solution than a solution that accounts for 100 percent of the variance.

Figure 7.5 shows a simulation presented by Koles (1998) that illustrates this problem. Koles created a distributed source by using several nearby dipoles arranged along a curved surface, and fit an equivalent current dipole to the scalp distribution produced by this distributed source. He placed model dipoles systematically at a variety of locations, and measured the residual error for each of these locations. As the right side of figure 7.5 shows, the error was lowest near the location of the best equivalent current dipole, but there wasn't much difference in error over a fairly wide range of locations. This is exactly the sort of information that one needs to know when trying to evaluate a source localization model.

In my view, it is misguided to attempt to find a unique solution given the uncertainties inherent in ERP localization. A better approach would be to report the entire range of solutions that fit the data and constraints to some criterion level (e.g., a fit of 95 percent or better). Moreover, it would be useful to report the range of
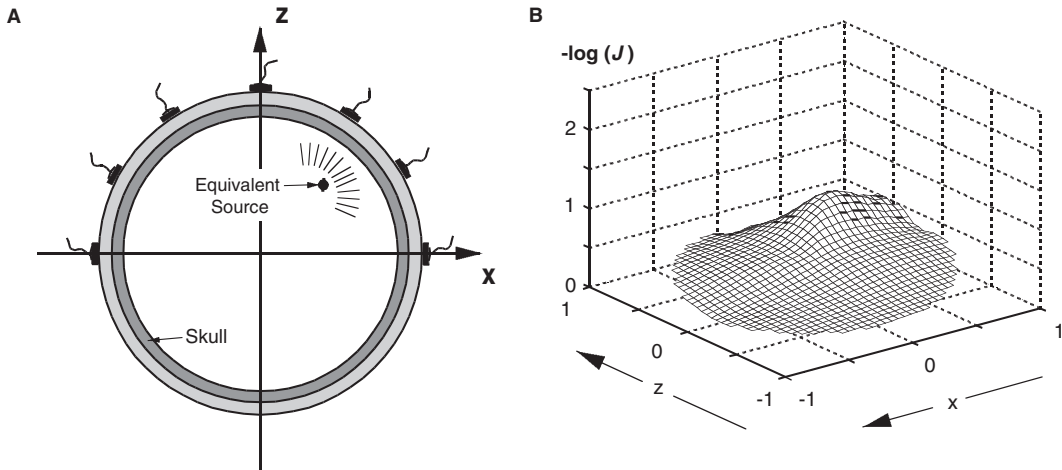
Figure 7.5 Example of a distributed electrical source (A) and the amount of error that would occur in estimating this source by a single dipole over a range of positions (B). Although a single point of minimum error exists, the amount of error varies little over a fairly wide range. (Adapted from Koles, 1998. © 1998 Elsevier Science Ireland Ltd.) Thanks to Zoltan Koles for providing electronic versions of these images.

solutions obtained as one adds and removes various constraints. If, for example, one were to find high levels of estimated activity in a particular region in almost any solution, no matter what constraints were used, then this would give us considerable confidence that this region really contributed to the observed ERPs.

A few investigators have explored this general sort of approach. For example, as described in the section on equivalent source dipole approaches, Huang, Aine, and their colleagues (1998) have developed a multi-start approach in which the dipole localization procedure is run hundreds or thousands of times with different starting positions. This makes it possible to see which dipole locations are found frequently, independent of the starting positions or even the number of dipoles in the model. This approach potentially solves the most significant shortcomings of equivalent source dipole approaches. In particular, the solutions are largely operator-independent, and it is possible to assess the likelihood that a given

dipole location occurred because of the starting positions of the dipoles or because of incorrect assumptions about the number of dipoles.

Although the multi-start approach addresses these shortcomings, it still falls short of providing a quantitative description of the probability that a particular brain area contributed to the observed ERP data. That is, a dipole may have been found in a given region in some percentage of the solutions, but the localization approach does not guarantee that the space of adequate solutions is sampled completely and evenly. However, Schmidt, George, and Wood (1999) have developed a distributed source localization technique based on Bayesian inference that provides a more sophisticated means of assessing probabilities. This technique is similar to the multi-start technique insofar as it generates thousands of potential solutions. However, its basis in Bayes's theorem allows it to provide a more complete and quantitative description of the space of possible solutions. This is the most promising localization technique that I have seen. Unfortunately, it has not yet been widely applied to real experiments, and other groups of researchers have not yet thoroughly explored its limitations. Nonetheless, the more general principle embodied by this approach and the multi-start approach—which systematically explore the space of likely solutions—seems like the best direction for the development of source localization techniques.

Recommendations

I will end this chapter by providing some recommendations about whether, when, and how you should use source localization techniques. My basic conclusion is that ERP localization is extremely difficult, and it should be attempted only by experts and only when the solution space can be reduced by well-justified constraints, such as structural MRI data and the combination of electrical and magnetic data. In addition, the most commonly used techniques are useful primarily for obtaining converging evidence rather than providing a conclusive, stand-alone test of a hypothe-

sis, although ongoing developments may someday allow source localization data to provide definitive results.

Source Localization and Scientific Inference

To assess the value of source localization techniques, it is useful to put them into the context of general principles of scientific inference. Perhaps the most commonly cited principle of scientific inference is Popper's (1959) idea of falsification. A commonly used, although less commonly cited, extension of this idea is Platt's (1964) notion of *strong inference*, in which the best experiments are those that differentiate between competing hypotheses, supporting one and falsifying the other.

How do source localization techniques fare when judged by these standards? Not well. I don't think anyone really believes that a single source localization model can conclusively falsify a hypothesis or definitively decide between two competing hypotheses. There are simply too many uncertainties involved in source localization. On the other hand, it is rare that a single experiment using any method is 100 percent conclusive, so this standard may be unrealistic.

A more flexible approach is to apply Bayes's Theorem to scientific inference. In this context, we can summarize Bayes's Theorem by stating that a new result increases the probability that a hypothesis is true to the extent that (a) there is a high probability of that result being true if the hypothesis is true, and (b) there is a low probability of that result being true if the hypothesis is false. In other words, a finding that is consistent with a hypothesis does not give us much more faith in the hypothesis if the finding is likely even if the hypothesis is wrong.

In this context, a given source localization model will have value to the extent that it not only supports a specific hypothesis but is also unlikely to have been obtained if the hypothesis is false. It's the second part of this equation that is especially problematic for source localization models, at least as researchers typically use them. As discussed in the previous section, source localization

models provide an estimate of the internal distribution of electrical activity, but they do not typically quantify the probability that the estimate is incorrect (which is related to the probability that the finding would be obtained even if the hypothesis is false). However, this problem can be overcome, at least in principle. For example, the probabilistic approaches described in the previous section are designed to provide information about the range of possible solutions, making it possible to assess the probability that activity would appear in a given location in the models even if the corresponding brain location were not truly active. Thus, although most source localization methods are not well suited for this kind of scientific inference, this does not appear to be an intrinsic limitation of the entire source localization enterprise.

Another commonly cited principle of scientific inference is the idea of *converging evidence*, which was first developed in the context of perception research (Garner, Hake, & Eriksen, 1956) but is now widely used in cognitive neuroscience. The basic idea is that many interesting questions about the mind cannot be answered by means of any single method, but a clear answer can be obtained when many methods with different strengths and weaknesses converge on the same conclusion. This is a common use of source localization models. That is, the researchers understand that the models are not conclusive evidence that the ERPs are generated in specific brain regions, but they believe that the models are valuable insofar as they converge with data from other sources. Box 7.1 provides an example of this from my own research.

Until source localization techniques routinely provide meaningful, quantitative information about the probability that a given model is correct, the main role of source localization models will be to provide converging evidence. However, not all cases of converging evidence are created equal: If a model uses weak methods to create a given source localization model, then this model will provide only weak converging evidence. And the value of such models is questionable, especially given the time and expense often involved in the modeling process. At present, source local-

Box 7.1 Converging Evidence

The following is an example of how I have used source localization to provide converging evidence for a specific hypothesis. My initial visual search experiments examining the N2pc component suggested that this component reflects the focusing of attention onto a target and filtering out irrelevant information from the distractor objects (Luck & Hillyard, 1994a, 1994b). This seemed similar to the types of attention effects that Moran and Desimone (1985) observed in single-unit recordings from area V4 and from inferotemporal cortex, but I had no way of localizing the N2pc to these areas. Then Leonardo Chelazzi conducted a series of follow-up studies in Desimone's lab using visual search tasks that were more similar to the tasks that I had used to study the N2pc component (Chelazzi et al., 1993, 1998, 2001). The onset of the attention effects in these single-unit studies was remarkably similar to the onset time of the N2pc component, and this suggested that the N2pc component might reflect the same neural activity as the single-unit attention effects. To test this hypothesis, I conducted a series of N2pc experiments that paralleled Chelazzi's single-unit experiments, and I found that the N2pc component responded to several experimental manipulations in the same way as the single-unit attention effects. To provide converging evidence, I collaborated with Max Hopf and Hajo Heinze on a combined ERP/ERMF study of the N2pc component using the cortically constrained minimum norm approach (Hopf et al., 2000). The resulting source localization model was consistent with a source in the general area of the human homologues of monkey V4 and IT (with an additional source in posterior parietal cortex). In this manner, the source localization data provided converging evidence for a link between the N2pc component and a specific type of neural activity (see Luck, 1999 for an extended discussion of this general approach, which combines traditional hypothesis testing with source localization).

ization models provide reasonably strong converging evidence only if they are the result of state-of-the-art methods and only if they are developed thoughtfully and carefully.

Specific Recommendations

My first specific recommendation is to avoid techniques that involve substantial input from the operator (which is true of many, but not all, equivalent current dipole approaches). These

techniques are so prone to experimenter bias that they can provide only the weakest sort of converging evidence. In fact, I would argue that these models are often worse than no models at all, because they provide the illusion of strong evidence when in fact the evidence is weak. The one exception to this recommendation is that these approaches may be adequate when a combination of three criteria are met: (1) the data are very clean; (2) you can be sure that only one or perhaps two dipoles are present; and (3) you have good reason to believe that the electrical activity is relatively focused rather than being distributed over a large region. Localization is fairly easy under such conditions, and validation studies have shown that localization errors average approximately 1 cm or less when these criteria are met (see, e.g., Cuffin et al., 1991; Leahy et al., 1998).

My second specific recommendation is to obtain structural MRI scans from each subject so that you can create a reasonably accurate head model and use one of the cortically constrained approaches (which typically involve distributed source solutions rather than equivalent current dipole solutions). For most experiments in cognitive neuroscience, it is very likely that the activity is generated exclusively in the cortex with a perpendicular orientation, and this provides a powerful constraint that reduces the solution space considerably. The most common versions are the depth-weighted minimum norm and LORETA techniques, described previously in this chapter. LORETA is well suited for situations in which (a) you want to determine the center of each activated region, (b) you do not care about the spatial extent of the activated regions, and (c) the activated regions are well separated from each other. If these conditions are met, LORETA appears to work quite well (see, e.g., the impressive LORETA/fMRI correspondence obtained by Vitacco et al., 2002). If these conditions are not met, I would recommend using the depth-weighted minimum norm approach.

My third recommendation is to use difference waves to isolate a single component (or small set of components; see chapter 2 for

more discussion). The more components are active, the more of a mess you will have to sort out. Equivalent current dipole approaches become particularly problematic when more than a few sources are present, but this can also be a problem for distributed source approaches.

My fourth recommendation is to record ERMFs in addition to ERPs. ERMFs have two major advantages over ERPs. First, they are not blurred and distorted by the high resistance of the skull, leading to greater resolution and a smaller space of possible solutions (see, e.g., the simulation results of Leahy et al., 1998). Second, because biological tissues are transparent to magnetism, it is not necessary to create a model of the conductivities of the brain, skull, and scalp, and this eliminates one possible source of error. ERMF recordings do have a couple disadvantages, though. First, they are very expensive, both in terms of the initial capital investment and the maintenance costs (particularly the coolant). Second, ERMF recordings will not be able to detect sources that are deep or perpendicular to the surface of the head (note, however, that this becomes an advantage rather than a disadvantage when ERMFs are combined with ERPs). In my experience, source localization is just too uncertain when based on ERPs alone.

The bottom line is that source localization is extremely difficult, and any serious attempt at localization will require sophisticated methods and considerable costs (both in terms of time and money). If you simply record ERPs from a large number of channels and try to fit a half dozen dipoles to the data with no additional constraints, it's not clear what you will have learned. At best, you will gain some weak converging evidence. At worst, you will be misled into believing in a solution that is simply incorrect.

I would like to end by noting that the clear strength of the ERP technique is its temporal resolution, not its ability to localize brain function. It is therefore sensible to use this technique primarily to answer questions that require temporal resolution, leaving questions about localization of function to other techniques. When a question requires a combination of temporal and spatial resolution,

a combination of ERPs, ERMFs, structural MRI scans, and fMRI data may provide reasonably strong evidence, but the commonly used methods for localizing ERPs/ERMFs do not make it clear how strong the evidence is. As new techniques are developed—particularly those based on probabilistic approaches—we may eventually get to the point where we can have a reasonably high (and known) level of certainty.

Suggestions for Further Reading

The following is a list of journal articles and book chapters that provide useful information about ERP and MEG source localization.

Dale, A. M., & Sereno, M. I. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience*, *5*, 162–176.

George, J. S., Aine, C. J., Mosher, J. C., Schmidt, D. M., Ranken, D. M., Schlitt, H. A., Wood, C. C., Lewine, J. D., Sanders, J. A., & Belliveau, J. W. (1995). Mapping function in the human brain with magnetoencephalography, anatomical magnetic resonance imaging, and functional magnetic resonance imaging. *Journal of Clinical Neurophysiology*, *12*, 406–431.

Hämäläinen, M. S., Hari, R., Ilmonieni, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Review of Modern Physics*, *65*, 413–497.

Koles, Z. J. (1998). Trends in EEG source localization. *Electroencephalography & Clinical Neurophysiology*, *106*, 127–137.

Luck, S. J. (1999). Direct and indirect integration of event-related potentials, functional magnetic resonance images, and single-unit recordings. *Human Brain Mapping*, *8*, 15–120.

Mangun, G. R., Hopfinger, J. B., & Jha, A. P. (2000). Integrating electrophysiology and neuroimaging in the study of human brain

function. In P. Williamson, A. M. Siegel, D. W. Roberts, V. M. Thandi & M. S. Gazzaniga (Eds.), *Advances in Neurology* (Vol. 84, pp. 33–50). Philadelphia: Lippincott, Williams, & Wilkins.

Miltner, W., Braun, C., Johnson, R., Jr., Simpson, G. V., & Ruchkin, D. S. (1994). A test of brain electrical source analysis (BESA): A simulation study. *Electroencephalography & Clinical Neurophysiology*, *91*, 295–310.

Pascual-Marqui, R. D., Esslen, M., Kochi, K., & Lehmann, D. (2002). Functional imaging with low-resolution brain electromagnetic tomography (LORETA): a review. *Methods & Findings in Experimental & Clinical Pharmacology*, *24 Suppl C*, 91–95.

Scherg, M., Vajsar, J., & Picton, T. (1989). A source analysis of the human auditory evoked potentials. *Journal of Cognitive Neuroscience*, *1*, 336–355.

Schmidt, D. M., George, J. S., & Wood, C. C. (1999). Bayesian inference applied to the electromagnetic inverse problem. *Human Brain Mapping*, *7*, 195–212.

Snyder, A. (1991). Dipole source localization in the study of EP generators: A critique. *Electroencephalography & Clinical Neurophysiology*, *80*, 321–325.