CHAPTER THIRTY-THREE

BIOSIGNAL PROCESSING

GABRIELE GRATTON

Introduction

This chapter reviews issues related to the analysis of psychophysiological data. It will focus on general questions that are relevant to a variety of techniques, rather than on specific issues related to individual methods. The previous chapter of this volume also deals with statistical issues in psychophysiology, focusing on issues of inference and data interpretation.

STAGES OF DATA PROCESSING

In describing the procedures used in analyzing psychophysiological data, it is useful to distinguish among different stages of analysis. The first stage is signal enhancement and elimination of observations that are artifactual or that can be considered outliers. This stage, sometimes called "signal conditioning," involves (at least in part) techniques that are specific to each type of physiological measure.

The second stage involves data reduction (sometimes called "quantification" or "parameter extraction"). Most psychophysiological experiments involve a large number of observations per subject (i.e., dependent variables). Although this provides data sets that are rich in information content, it also increases the probability of spurious or noisy observations. In general, one of the main objectives of signal processing is maximizing the signal-to-noise ratio. A large number of observations may contain a large number of signals to study but may also contain a large amount of noise. It is therefore advantageous to select information that is relevant and to minimize redundancy between the dependent variables entered in the statistical analysis, since redundant dependent variables will not necessarily increase the signal yet may generate additional opportunities for noise. An important step in signal processing is to reduce the data set to the smallest possible number of dependent variables (each affected by the smallest possible amount of noise) while preserving the information content of the original data set as much as possible. However, a moderate degree of redundancy could be useful in distinguishing random noise from systematic noise and thus in deriving a valid and reliable estimate of the signal.

The third stage of data processing is statistical analysis, which may include hypothesis testing, model fitting, parameter estimation, and so forth. As mentioned previously, most psychophysiological experiments are multivariate in nature—that is, they involve repeated observations from the same subjects. The multivariate nature of psychophysiological measures is one of their greatest assets, for it allows the experimenters to view phenomena from different "points of view," thus affording a more complete and detailed picture of bodily phenomena. However, it often introduces special types of problems in the statistical evaluation of multivariate data sets. Two general approaches can be considered: the univariate approach, which considers one dependent variable at a time; and the multivariate approach, which considers all dependent variables (as well as their covariation) together (see Huberty & Morris 1989). Each of these approaches has advantages and disadvantages. In general, when there is a limited number of subjects in the study (relative to the number of dependent variables) and the dependent variables tend to be correlated with each other across conditions (or groups) in the same manner as within conditions (or groups), the univariate approach is more powerful. Conversely, the multivariate approach is preferable when there is a large number of subjects in the study (relative to the number of dependent variables) and when the within- and between-condition (or group) correlations among dependent variables tend to go in opposite directions. (The statistical analysis stage is treated in Chapter 32 and thus will not be considered in detail here.)
In concluding this section, it is important to remind the reader that the output of an analysis is only as good as its input (this is captured by the well-known aphorism “garbage in, garbage out”). Complex and sophisticated data analyses can in no way substitute for clear experimental hypotheses, elegant experimental designs, and accurate data collection procedures. However, they can complement these steps in providing clear and convincing data.

GENERAL ISSUES

Psychophysiological data vary along a number of dimensions, and different types of data propose different types of analytical problems and solutions to the investigator. However, there are a few overarching themes that are valid for most types of measurements. Some general issues are particularly important and worth mentioning at the beginning of this chapter. First, psychophysiological measures are typically indirect; in other words, they are measures of events occurring outside the human body that are related to events that occur inside the body. For instance, measures of the electrical activity produced by the body (such as the electrocardiogram and the electroencephalogram, ECG and EEG) are in fact measures of the difference in potential between two electrodes located on the surface of the body; positron emission tomography (PET) measures the arrival of high-energy photons to particular detectors located at some distance around the body, which in turn is related to the concentration of radioactive substances in various areas of the body; and so on. In general, psychophysiological measures are subject to some transformation with respect to the original signal generated within the body (because of physical, physiological, or psychological reasons). This transformation may partially distort the signal (and its statistical properties) and may occasionally generate artifacts that need to be recognized and eliminated from the data before statistical analysis is performed.

Second, most of the data are sampled in a discrete fashion — both temporally and spatially. The transformation of analog parameters into discrete measures may distort the signal and require specific solutions to statistical problems. One of the major problems is obtaining a representative sample of the variable of interest over both time and space. In addition, some techniques (such as measures of heart rate acceleration or deceleration) may be based on the observation of internal events (e.g., the occurrence of an R wave in the ECG) whose spacing is variable (and, in the case of heart rate changes, is a function of the variable of interest). This may make it more complicated to compare observations across trials and subjects. Investigators using these measures have developed special approaches to address this problem (see Chapter 9).

Third, most psychophysiological measures are multiply determined; that is, they are determined by the interaction of a variety of factors. Some of these factors may be extraneous to and uncorrelated with the experimental manipulations (and thus will contribute to statistical noise), while other factors may be related to the experimental manipulations in ways that differ from those intended by the researcher (and hence will contribute both to statistical noise and to systematic effects). The latter type of confounds are particularly insidious and difficult to eliminate. One of the most important characteristics of good experimental methodology is the use of appropriate control procedures to distinguish between signal and systematic noise.

Fourth, most psychophysiological measures are inherently noisy, making it difficult to distinguish between signal and noise in the raw (single-trial) observations. Several techniques have been developed to extract the signal. One of the simplest is signal averaging: the observation is repeated several times and the individual responses are averaged together. Signal averaging is based on the central limit theorem and can therefore be expected to provide reasonable estimates of the expected central tendency of the population in a large number of situations – provided that enough trials are used. Note that, for signal averaging to work appropriately, it is important that the expected value of the noise across trials be equal to zero. Although this is a reasonable assumption in many cases, there are certainly situations in which this is not true. For instance, as we will see later, frequency analysis (such as the Fourier transform) can be used to estimate the “power” or variance associated with a particular frequency. Since this value cannot be less than zero, the average of the power values for a particular frequency obtained in different trials (or in different subjects) will always be greater than zero, even if no systematic activity exists at that frequency.

A problem with signal averaging is that it eliminates potentially relevant information about the variability of the signal from trial to trial (and in some cases from subject to subject); this variability is entirely attributed to stochastic phenomena. In most cases, however, the variability of the signal is not entirely random but rather depends on the influence of intervening variables. For instance, in an experiment evaluating the size of the visual evoked potential as a function of stimulus intensity, the size of the response may also be influenced by subjects’ variables (such as allocation of attention, emotional factors, etc.) as well as by contextual conditions (stimulus sequence, noise in the environment, etc.) or still other factors. When signal averaging is used, all of these intervening variables are treated as contributors to stochastic noise and thence ignored. This may lead to loss of information or to systematic errors when contingencies exist between the experimental manipulations of interest and any of the intervening variables. Similar problems have been described with respect to other types of measures (see Estes 1956; Siegler 1987).

Another problem with signal averaging is that the signal may vary along dimensions other than its intensity: in some cases, the latency of the phenomenon or its spatial location

BIOSIGNAL PROCESSING

...
may vary from trial to trial (partly for the reasons just outlined). In this case signal averaging may produce systematic distortions, the most typical of which is a “smearing” of the response across different data points. The effect of smearing is more intense for signals with a steep gradient. In the case of electrophysiological signals that vary quickly over time, this leads to the phenomenon of “latency jitter”: the size of an evoked electrophysiological activity may be underestimated using signal averaging because the activity “peaks” at different latencies on individual trials. A similar problem exists when data from different subjects are pooled together (for a discussion of this problem and some solutions, see Möcks et al. 1988). It is interesting to note that the development of brain imaging methods, which have strong gradients in the spatial domain, has led to the recognition of a similar smearing problem when average spatial maps are computed across subjects (because of intersubject variability in brain anatomy); see Figure 1. This problem is in fact not specific to psychophysiological measures. For instance, investigators in psychology have long recognized that apparently smooth speed–accuracy functions may result from the mixture of distributions of fast responses with a low level of accuracy and slow responses with a high level of accuracy, provided that the two distributions have some degree of partial overlap in response time (Yantis, Meyer, & Smith 1991).

In order to address this problem, several investigators have proposed to measure psychophysiological phenomena on single trials (or subjects). A critical step of this procedure is the development of techniques for identifying target features on single-trial records. These techniques, although quite varied, can be grouped as “pattern recognition” algorithms. Some are based on visual inspection of individual data and are most useful when the target features are much larger than the noise. Others are computer controlled and based on statistical approaches as cross-correlation, auto-correlation, or other automatic methods. The individual records may then be aligned along the target feature and averaged together. Alternatively, some measure other than the mean can be used to describe the single-record distribution. An example of an alignment methodology in the time domain is the Woody filter (Woody 1967), which is a general-purpose statistical method; an example in the spatial domain is the alignment method proposed by Talairach and Tournoux (1988), a special-purpose approach that has become a standard practice in brain imaging.

**General Classifications**

Psychophysiological measures are usually defined as noninvasive measures of bodily functions taken with the intent of addressing issues that are relevant to psychologists (i.e., to obtain information that is relevant to the study of the human nature). This broad definition encompasses a large variety of measures that differ along a number of dimensions. Of these dimensions, some are important for the purposes of the present chapter because they result in different types of signal processing problems. I will pay particular attention to three of these dimensions.

**FAST VERSUS SLOW MEASURES**

As mentioned before, psychophysiological measures reflect bodily changes that are related to psychological phenomena. These changes occur over time. Some of them evolve very quickly (within a few milliseconds) and others quite slowly (over seconds, minutes, or even longer times). Indeed, for each type of psychophysiological measure used, it is possible to consider a parameter that will indicate the speed at which the activity evolves. We may call this parameter the “time constant” of a particular physiological measure. This parameter depends upon three factors: (a) the rate of change of the relevant psychological phenomenon; (b) the time required by the physiological event being measured to track changes in the psychological phenomenon; and (c) the time required for the measurement.
By and large, it is possible to consider individual time-delay functions for each of these steps, although in some cases interaction terms should also be considered. Note that each step may itself comprise several substeps. Each step can be considered as having an input (coming from the previous step) and an output (influencing the next step) — but a delay will usually occur in the operation of translating an input in an output. Here is an example of a function expressing this delay:

\[ O_t - O_{t-1} = (I_t - O_{t-1})(1 - e^{-t/\tau_c}) \]  

(1)

In this function, the change in the output \( O_t - O_{t-1} \) of a particular step is a delayed function of the difference between its input \( I_t \) and its previous level \( O_{t-1} \). The delay is determined by an exponential function whose critical parameter is \( \tau_c \), the time constant of this particular step. When several steps are involved, the total time-delay function is determined by a convolution of the individual functions. That is, the total time constant of a measurement is related to the integration of the time constants of each individual step. An example of this cascade, and of its effect on the final output measure, is illustrated in Figure 2.

In this example, the initial input is a 3-sec pulse (this may simulate a stimulus of short duration). Four steps were simulated, with time constants of 0.2, 1, 8, and 3 seconds. Note that the final output is substantially delayed with respect to the input and actually peaks well after termination of the pulse (i.e., of the stimulus).

The steps contributing most to the total time constant are called “rate limiting” steps (in the example given, step 3 and, to a smaller extent, step 4). Whereas all factors determining the measurement's time constant — rate of change of the psychological event, delay of the physiological phenomena, and measurement time — are always present to some extent, there is often only one (or a few) rate limiting factor(s). The nature of the rate limiting factor and the time constant of the measurement vary between different physiological measures. I will label measures with a long time constant “slow” measures and those with a short time constant “fast” measures.

Certain measures of autonomic activity are slow, including measures of electrodermal activity (EDA, in which the psychological and/or the physiological phenomena evolve slowly over time), functional magnetic resonance imaging (fMRI, in which the physiological event under study — a change in blood oxygenation — evolves slowly), and PET (in which the response delay of the hemodynamic or metabolic systems to changes in brain activity is coupled with slow measurement time). Fast measures include those of the event-related brain potential (ERP), magnetoencephalogram (MEG), electromyogram (EMG), and event-related optical signal (EROS). For these measures, all three determining factors are fast. Some other measures, such as heart rate and pupil diameter, are intermediate between these two extremes. Measures with a short time constant have a good temporal resolution (i.e., the effects of two different psychological and/or physiological events occurring in rapid succession can be distinguished), whereas measures with a long time constant have a poor spatial resolution (i.e., events in rapid succession cannot be separated easily). This is illustrated in Figure 3, which compares the time courses of the effect of two stimuli presented in rapid succession on measures with short and long time constants.

Measures with a slow time constant provide few independent observations per unit time. In contrast, measures with a fast time constant provide a large number of independent observations over the same time unit. The processing of psychophysiological data usually requires transformation of the data into a series of numbers sampled over time (i.e., a time series). The issue then arises of the optimal temporal sampling strategy for each measure.

The sampling strategy must preserve as much as possible the information contained in the signal. The minimum sampling frequency at which information is maintained is twice the frequency of the signal (this is called the “Nyquist frequency”). At sampling rates slower than the Nyquist frequency, aliasing causes a high-frequency signal to manifest itself as a lower-frequency signal. This phenomenon is illustrated in the top portion of Figure 4. A sampling frequency that is too...
Two 200-msec stimulation pulses were presented, separated by an interval of 300 msec (this is intended to simulate two independent psychological events). The closed-circle line is obtained by assuming a cascade of three intervening steps between the stimulation and the measurement system, each introducing a delay characterized by a time constant of 200 msec. The open circle line is obtained by assuming a cascade of three intervening steps with time constants of 5 sec. Note that the short-time constant measure affords easy separation of the two pulses (i.e. psychological events), whereas effects of the two pulses with the long-time constant measure are lumped together.

slow may also completely miss the occurrence of a very fast, transient phenomenon; this is illustrated in the lower portion of Figure 4. This latter consideration may encourage investigators to sample at the highest possible frequency, yet we should also remember that – since each observation will contain both relevant information and noise – it is important to minimize the number of data points (i.e. dependent variables) entered into the statistical analysis. This may be achieved conveniently at the data reduction stage. Finally, some recording systems impose limitations on the maximum sampling rate. However, the issue of aliasing should always be considered when studying a physiological variable. This problem is particularly likely to occur when fast phenomena are studied using a relatively slow sampling rate (e.g., when EEG is measured at less than 20 Hz, ECG at less than 2 Hz, or respiration at less than 0.5 Hz).

**DIRECT VERSUS INDIRECT MEASURES**

Psychophysiological measures also differ in terms of how many steps intervene between the physiological phenomenon that is the "real" target of the study (e.g., neuronal activity, sympathetic or parasympathetic activation) and the actual physiological measure that is taken (e.g., scalp electrical activity, increased blood flow, increased heart rate or increased EDA). The indirect nature of the measures may have several consequences. First, as mentioned previously, it may introduce a lag between the psychological event of interest and observed changes. Second, each of the intervening steps may introduce distortions in the size and/or shape of the responses with respect to the original signal.

A particularly important issue related to the indirectness of psychophysiological measures is that of linearity of a measure. A measure $M$ is linear with respect to its inferential psychophysiological variable $\phi$ if the following formula is valid:

$$M = k\phi + \text{offset} \quad \text{with} \quad k \neq 0 \quad (2)$$

For all practical purposes it may be sufficient that this formula is valid (at least approximately) for only an interval of values of $M$ and $\phi$, provided that this interval contains typically observed values of the measure and the expected range of the psychophysiological factor. For the linearity assumption in this formula to be valid, two terms must remain constant across conditions: the factor $k$ and the measurement offset. The proportionality factor $k$ depends on such variables as the relationship between the psychophysiological construct of interest and the physical phenomenon actually measured, the propagation of the physiological measure to the surface, the recording and analysis procedures employed, and so forth. An example of change in the proportionality factor is a change in the way neuronal signals are transformed into a hemodynamic phenomenon. For instance, a similar increase in neuronal activity at occipital and frontal locations may result in a different increase in the scalp electrical signal measured over these two areas. The offset factor may also depend on the effect of variables other than the physiological phenomenon of interest on the observed measure. For instance, an apparent increase in the amplitude of the ERP at frontal electrodes may be the consequence of an increase in the ocular artifact. Similarly, head movements may result in apparent large changes in the MRI signal.

Another interesting example for psychophysicists is the difference between measures of EDA. Two such measures are possible, of electrical conductivity versus electrical resistance between two locations on the skin; these two measures are inversely related to each other. However, the physiological phenomenon of interest (increase in conductivity associated with the activity of the sweat glands) is directly proportional to conductivity and inversely proportional to resistance. In this case, then, conductivity is a linear measure of the physiological phenomenon whereas
resistance is not. Quigley and Berntson (1996) pointed out that a similar case can be made for measures of heart rate versus heart period (which are of course reciprocal to each other): heart period, but not heart rate, appears to vary linearly with basal autonomic activation.

Linearity is useful because it allows for linear transformations of the measure \( M \) (such as additions, subtractions, multiplications by a constant, averaging, etc.) to be considered valid also with respect to the physiological variable \( \phi \). Linear manipulations of the observed measures are commonly performed in the analysis of psychophysiological data. However, if the relationship between \( M \) and \( \phi \) is not linear (i.e., if \( k \) or the offset are not constant), then such manipulations of the dependent variable may not be valid with respect to the psychophysiological variable, which precludes drawing any inferences (other than potentially misleading ones).

Lack of linearity may occur whenever any of the intervening steps between the observed measure and the psychophysiological construct of interest involve some form of nonlinear transformations. For example, consider the steps occurring between neuronal activity and increased \( { }^{15}\text{O}_2 \) PET response. The steps include the probable release of some chemical by the active neurons, consequent local vasodilation, increased regional blood flow, and – if radioactive material is introduced into the blood flow – increased radioactivity in a particular volume of the head. In principle, the relationships between causes and effects existing at each step may be nonlinear or even nonmonotonic. In other words, it is not necessarily the case that, for a certain percent increase in neuronal activity, there is an equal percent increase in vasodilation. Therefore, linearity needs to be assessed on a case-by-case basis. For PET responses, Fox and Raichle (1984) observed an almost linear relationship between frequency of visual stimulation (between 0 and 7.8 Hz) and change in blood flow in primary visual cortex as estimated using the \( { }^{15}\text{O}_2 \) PET method (see Figure 5). The presence of linearity is critical for the application of subtraction methods during the analysis of hemodynamic responses.

Even if the linearity assumption is not met, it is still possible that the relationship between the dependent variable and the physiological phenomenon is monotonic. Whereas a monotonic relationship may limit the
use of linear transformations and of parametric statistics, it is still sufficient for the application of nonparametric statistics based on qualitative judgments about effect sizes. Finally, relationships that are nonlinear but follow a more complex function can be handled by transformations that render them linear (as, for instance, the transformation of electrodermal resistance measures into conductance measures).

So far I have discussed the linearity of the relationship between the observed measures and the physiological parameter they intend to measure. To the extent that psychophysiological measures are used to make psychological inferences, it is also important to consider the type of relationship that exists between the physiological parameter and the underlying psychological concept of interest. For example, the purpose of measuring neuronal activity in a certain brain area may be the investigation of memory or attention phenomena, and the purpose of measuring blood pressure may be the investigation of the effect of psychological stressors on the cardiovascular system. Clearly, the issue of linearity (or at least of monotonicity) of the relationship is just as important here as in the case of the relationship between the observed measure and the physiological variable.

**CONTINUOUS VERSUS DISCRETE MEASURES**

The distinction between continuous and discrete measures refers to the time at which information about the status of the physiological system under study is, in principle, available. Some psychophysiological measures can provide information at any time with no particular restriction. For instance, the diameter of the pupil can be monitored continuously to provide a more or less continuous measure of the level of activation of particular subcomponents of the autonomic system. Likewise, the EMG and the EEG are continuous measures of electric potentials produced by muscles and the brain respectively. However, other physiological measures (e.g., of heart rate, respiration rate) reflect modification of parameters of cyclic events that occur at some intervals within the body. With these measures, information about psychological phenomena may be available only at discrete times. For instance, the state of activation of some subcomponents of the autonomic system can be studied using changes in heart rate, but information is available only when the heart beats (or when a particular electrocardiographic wave, such as the R wave, is produced). Still other measures can only be obtained by using particular maneuvers that elicit special responses in the body. Parameters of the blink reflex, for example, can only be measured by eliciting the blink reflex itself. Therefore, using the blink reflex to monitor changes in status of a particular physiological system (such as one related to emotion) entails that information may be available at specific times only (Bradley, Cuthbert, & Lang 1991). A similar situation is obtained when the P300 component of the ERP elicited by secondary task probes is used to monitor the level of attention of subjects to primary task stimuli (Wickens et al. 1983).

I will label as "continuous" those measures that can provide information about the state of a physiological system at any time and as "discrete" those measures that provide information only at specific times. An advantage of continuous measures is that information about the time course of psychological phenomena is readily available on single-trial records. Discrete measures can also provide information about the time course of psychological events, but their effective sampling rate is limited by the interval between the successive measurement windows. Finer sampling rates require pooling information obtained across trials and thus are more complicated to obtain. An additional problem with some discrete measures is that the time at which information is available may not be completely under experimental control. For instance, certain measures related to circadian rhythms can only be obtained unobtrusively when the subject is awake. This may result in a "variable" sampling rate that may complicate the analysis. For example solutions to this problem, see Monk and Fookson (1986) and Monk (1987). Measures of heart interbeat interval changes as a function of when (during a cycle) the stimulus is presented create special problems of identification because the time at which measures are available (i.e., the next heartbeat) varies depending on the phenomenon under study. Appropriate representation of the effects requires special techniques (see e.g. Jennings et al. 1991).

**Signal Extraction and Enhancement**

**DOMAINS OF ANALYSIS**

Psychophysiological data can be considered as elements of a multidimensional matrix. We can consider the following dimensions that are specific to psychophysiological measures: (a) type of measure; (b) time of sampling (or time with respect to some anchoring event); (c) location in space (itself comprising one, two, or three dimensions). Other dimensions may also be considered, such as experimental condition, subject, group, and so on, but these dimensions are not specific to psychophysiological measures. Since some of the dimensions are grouped together (such as the three spatial dimensions), it is useful to talk about domains of analysis (the temporal domain, the spatial domain, etc.).

For some (or most) experiments, some of these dimensions may have only one cell and are therefore fixed. For instance, we may be recording only one psychophysiological measure or collecting data from only one location. It is important, however, to consider that different values on each of these dimensions all refer to the same individual
condition B with respect condition A or rather that an altogether different type of ERP activity occurs in the two conditions?

To address these issues it is usually convenient to perform some intermediate analytical (or signal processing) steps between the level of observation of the raw data and the inferential statistical analysis per se, in which the data are used to provide support for or against theoretical arguments. These preliminary steps are known as data reduction or quantification. A consequence of these preliminary analysis steps is that the data should be considered no longer as raw observations but as quantifications of parameters of some intermediate hypothetical construct that has heuristic value and is usually thought to correspond to some anatomical and/or physiological "entity" (activity in a particular brain structure in response to a particular stimulation condition, the blink response, etc.). In all cases, data reduction implies some "interpretation" of the data; that is, an analytical model is applied that (i) makes more or less explicit assumptions about the underlying structure of the phenomenon observed and (ii) involves some mathematical or logical transformation of the original (raw) data in order to reconstruct the value of some parameter of the underlying structure. The quality of the conclusions drawn from an experiment depends critically on the validity of these assumptions.

**TEMPORAL DOMAIN**

**Time Domain and Frequency Domain**

The time factor can be analyzed according to either of two basic models. According to the first model, the physiological events in the period under study evolve in a temporally ordered series that exhibits nonstationary properties (i.e., the basic structure of the activity changes over time). An example is the sequence of responses that occur in the brain after presentation of an individual stimulus. Another example is the changes in heart rate before and after the presentation of a critical stimulus. In these cases, it is profitable to use time as the main dimension of analysis (i.e., conduct a time-domain analysis). Usually, particular features of the time series (such as a maximum, a minimum, the integrated amplitude over an interval, a particular waveform) are considered as target features, and parameters of these features (such as amplitude, latency, or similarity with a particular template) are quantified and entered in the statistical analysis. In addition, time series can be averaged together (with or without alignment with respect to particular features) in order to increase the signal-to-noise ratio. This approach provides for a substantial data reduction, and data can be expressed in terms of properties of the target feature(s).

The second model, in contrast, assumes that the basic structure of the activity does not vary over time and so exhibits stationary properties in which events repeat...
in a cyclical fashion. Examples include the consistent response of the brain to a regular train of similar stimuli (when habituation is not supposed to play a role) and the oscillation in heart rate associated with respiration (sinus arrhythmia). When this model is appropriate, physiological activity can be analyzed using a set of tools that are able to extract the basic periodic structure of the activity; these methods are called frequency-domain methods. In this case, a new data series is built in which frequency replaces time as the ordering dimension, and data are expressed in terms of how much of the total variability over time is accounted for by fluctuations occurring at particular frequencies. Distinctive features of this new series can then be considered for quantification purposes (e.g., the "dominant" or most represented frequency in the data can be determined) in the same manner as for time-domain analyses. Frequency analysis may also provide information about the relative delay (or phase) at which particular rhythmic oscillations occur with respect to a reference time (such as the time of stimulation).

**Fourier Analysis and Autoregressive Methods**

Two types of frequency-domain methods are most commonly used: the Fourier transform and autoregressive methods. Both approaches express the original time series in terms of the extent that certain frequencies occur in the data. The Fourier transform is based on a "deterministic" approach, that is, it represents all of the information (variance) contained in the original data in terms of the independent contributions of equally spaced frequencies. In contrast, autoregressive methods (of which there are different varieties) are based on a "modeling" approach: they impose some constraints on the way in which individual frequencies contribute to the original time series. This modeling approach reduces the number of free parameters in the data (i.e., it reduces the dimensionality of the data). This may allow the investigator to separate signal from noise, but it also requires meeting the assumption that the constraints imposed by the procedure are valid.

The Fourier transform is a numerical method commonly used in engineering and other disciplines. A prerequisite for its application is that the sampling rate be constant across a given epoch (for more on frequency analysis techniques with irregular sampling, see Monk 1987). The basic logic of this method is that any given time series (recording epoch) can be expressed in an equivalent manner (i.e., without loss of information and without interpolation or extrapolation) in the time domain and in the frequency domain. In the frequency domain, a time series can be expressed as the sum of several time series, each characterized by the equal-amplitude oscillations of sinusoidal functions with frequencies equal to a multiple of the inverse of the length of the time series (e.g., if the recording epoch is 1 sec long than the frequencies used to describe the time series will be 1 Hz, 2 Hz, 3 Hz, etc.) plus an offset term related to the mean value across data points (usually called "zero frequency" or DC value). The number of frequencies is equal to the number of elements of the time series divided by 2. Thus, if we are sampling a 1-sec epoch at 128 Hz, the Fourier decomposition will be based on one DC value and 64 basic frequencies (1 Hz, 2 Hz, 3 Hz, ..., 64 Hz). Each basic frequency is associated with an amplitude value (related to how large the oscillation is at that particular frequency) and a "phase" value, which is related to the relative timing of the peak of the first oscillation with respect to the beginning of the time series. Derivation of the Fourier transform is numerically complex and requires a number of exponential operations that is proportional to the square of the number of elements in the time series. Application of the Fourier transform to long time series may therefore be cumbersome. Fortunately, when the number of elements in a time series is a power of 2 (such as 4, 8, 16, 32, 64, 128, ...), the numeric problem simplifies and can be carried out using a procedure called the fast Fourier transform (FFT).

The outcome of the Fourier transform is a complex number associated with each frequency in which the real and imaginary parts both contain information associated with the amplitude and phase of the basic oscillations. The real and imaginary components can be thought of as corresponding to the coordinates of a vector in a complex plane. However, it is most convenient to express the results in terms of the polar coordinates of this vector (length and angle or orientation). This transformation separates the amplitude (equivalent to the length of the vector) and phase (equivalent to the orientation of the vector) associated with each frequency. Remember that this transformation, although informative, is not linear. Therefore, averaging (or other linear operations) of data is most appropriately conducted prior to this transformation.

The Fourier transform yields a set of phase and amplitude values for each frequency. In most cases, the attention of the investigators focuses on the amplitude data (although there are exceptions, such as in the study of visual evoked potentials; Tomoda, Celesia, & Toleikis 1991). A common way of displaying the results is by plotting amplitude as a function of frequency (this function is also known as the periodogram). In some cases, instead of the amplitude of each oscillation frequency, the data are expressed in terms of the square of this amplitude. This is called the "power" of a particular frequency. **Amplitude and power are equivalent in the frequency domain to standard deviation and variance in the time domain.** Indeed, the sum of the power values for all frequencies is related to the total variance in the data. For this reason, the Fourier transform can also be considered as a method for partitioning the observed variance in a time series into subcomponents with different frequencies. Further, it is possible to apply to power frequency analysis some of the standard tools used for analyzing variance (chi-square test, analysis of variance, etc.)
Finally, it is a common practice to pool together the power for an interval of frequencies within a certain range (i.e., a frequency band). For example, in an EEG study we might be interested in quantifying the total power for frequencies ranging between 8 Hz and 12 Hz (usually called the alpha band) and, in a heart rate study on sinus arrhythmia, the power for frequencies between 0.2 Hz and 0.5 Hz (i.e., the respiration frequency range).

As mentioned before, the Fourier transform is based on the observation that any time series can be decomposed in a number of sinusoidal waves that extend with equal amplitude across the entire time series. Although mathematically valid in all cases, this way of describing the data may produce results that require further interpretation. This may occur for several reasons.

First, there may be cases in which the oscillations in the time series can be best described by functions other than a sinusoid. For instance, the pressure waves that propagate through the arteries after a heartbeat are, at first approximation, triangular in shape (sawtooth; see upper portion of Figure 7). When the shape of the oscillations departs from sinusoidal, the Fourier transform will decompose each wave into subcomponents whose individual frequencies are multiples of the basic frequency at which the wave repeats. These frequencies are labeled "harmonics" and the basic frequency is called the "fundamental." As an example, the periodogram of the sawtooth arterial pressure waves is shown in the lower part of Figure 7. Note that this periodogram indicates activity not only at the basic frequency (around 1.2 Hz) but also at some of its harmonics (such as 2.4 Hz). The pattern of harmonics is related to the particular wave shape of the basic, cyclical pattern. Given that the basic wave shape in most psychophysiological measures is not a sinusoid, psychophysiologists are often just as interested in the patterns of harmonics as they are in the fundamental frequency.

Data may also depart from the description underlying the Fourier transform when the type of activity observed in one part of the epoch differs substantially from that in another part. An extreme case is that of "phasic" features, which appear only once in a time series. Phenomena of this type, which are frequently observed in psychophysiological data, may be difficult to study using Fourier transformed data, because the variance associated with them will be distributed in a very complex manner between the DC component and various frequencies (although this variance will still be represented, in some form, in the Fourier transform). In other words, the variance partitioning that is obtained using Fourier transforms is best suited for phenomena that repeat cyclically along the time series (i.e., "stationary" waves).

Joint Time–Frequency Domain Analysis

In recent years, combinations of time-domain and frequency-domain approaches have been introduced (see e.g., Stiber & Saro 1997). The major justification for these hybrid techniques is the assumption that the brain processes information via rapid modulation of basic frequencies of activity. A frequency that has been considered as particularly important in brain activity is approximately 40 Hz—the gamma band (see Gray et al. 1989). According to this view, information processing may involve amplitude modulation of 40-Hz activity. When a new stimulus is presented, the oscillations at this
frequency increase in size. A converse phenomenon has been observed for the alpha band ("alpha blocking" or "enhancement"; see e.g. Pfurtscheller & Neuper 1992).

Another case for hybrid analysis is examining the modulation of sinus arrhythmia over time as a method for studying the time course of vagal activity. Yet another case is given by analysis of the EMG response, which often involves the occurrence of a brief series of high-frequency (typically above 30 Hz) oscillations. In all these cases, the psychological events are transient (and therefore evolve over time) yet their physiological manifestations involve modulation of a cyclical phenomenon. To represent this modulation phenomenon appropriately it is necessary to study the change over time of the amplitude of specific frequencies. This involves segmenting the time series into shorter epochs, determining the amplitude of the frequency of interest in each epoch (this analysis is performed in the frequency domain), and analyzing the time course of modulation of the amplitude of the frequency of interest (this analysis is performed in the time domain). The procedure is known as joint time–frequency analysis or JTFA (e.g. Tallon-Baudry et al. 1996). In the case of EMG, the response is often analyzed entirely in the time domain and involves: (a) recording using a relatively high-frequency bandpass (such as 10–100 Hz), (b) rectification of the wave forms (i.e., transformation of the negative value into positive values), and (c) low-pass filtering (typically below 20 Hz). This process is adequate for determination of latency measures, but it does require recordings with low noise levels.

### Amplitude-Domain Measures

All of the temporal-domain data processing steps outlined so far result in descriptions of the observed activity as a function of time, frequency, or a combination of both. However, in some cases, it is preferable to analyze the overall "amount" of activity observed without expressing it as a function of either time or frequency (Cacioppo & Dorfman 1987). This may be appropriate when the time (or frequency) variable is considered irrelevant or when variations of the measurement over time depend on factors that are not related to the physiological (or psychological) variable of interest. For instance, when the stimulation condition cannot be conveniently subdivided into specific time epochs or periodic events, it may be advantageous to consider the overall physiological response observed during an experiment instead of considering the individual responses. An example is the number of electrodermal responses during presentation of a particular videotape, which is not characterized by clearly identifiable time markers. Another case is that of measures that cannot be taken repeatedly, or whose time of acquisition is longer than the time during which the experimental manipulations are conducted. When using the deoxyglucose PET technique, for instance, measures of the amount of radioactive material accumulated in various brain areas are taken at the end of the experiment (Cherry & Phelps 1996). The temporal dimension in this case is reduced to one data point, and absolute amplitude measures are taken. Finally, in some cases the investigator wants to summarize in a unique number all the variability of observations obtained across an extended epoch (Elui 1969; see also Dorfman & Cacioppo 1990).

### SPATIAL DOMAIN

#### Spatial Dimensions

In recent years there has been a substantial growth of interest in the spatial characteristics of psychophysiological measures insofar as they correspond to the anatomy of the human body (and particularly of the brain). A particular impetus to this enterprise has been given by the introduction of functional neuroimaging methods (fMRI, EROS, PET, and single-photon emission computerized tomography or SPECT). Because brain anatomy is not conveniently expressed using frequency-domain methods (owing to the fundamental uniqueness of the various structures), other methods of analysis are generally preferred (an exception is the use of Fourier transforms in decoding the spatial information in MRI; see Cohen 1996).

Psychophysiological measures differ in the number of dimensions used and in their relationship with underlying anatomy. Some measures refer to the activity of the brain (such as EEG, ERPs, MEG, fMRI, PET and EROS), whereas others refer to other body structures (such as ECG and heart rate, EMG, pupillary diameter, measures of the blink response, measures of EDA, blood pressure, respiration). Although both types of measures (of brain activity and of bodily functions) can be taken at different locations, in general the latter are less directly dependent on brain anatomy than the former.

All measures of brain activity produce substantially different results depending on the location at which they are taken. Some of these measures provide data that can be localized to specific brain volumes (e.g., PET and fMRI). For these measures, any recorded value will have three indices, referring to the x, y, z spatial coordinates of the observed value. Other measures (such as ERPs and MEG) provide data that refer to surface observations and thus have only x and y coordinates. For these measures, algorithms have been proposed to allow for the reconstruction of depth information, thus pinpointing brain structures that are responsible for the activity observed at the surface. However, these reconstruction algorithms require a modeling effort that, for both ERPs and MEG, yields multiple solutions that are equally valid from a statistical point of view. Therefore, only limited confidence can be placed on the results.

Maps and tridimensional reconstructions obtained with various techniques can be evaluated on the basis of two major properties: (a) localization power, or the ability to attribute a particular activity to a particular location; and
(b) resolution, the ability to distinguish activity coming from two closely located points. These two properties are different and should not be confused. Some techniques may possess great spatial localization power but limited spatial resolution. This occurs because it is often the case that the larger the activity, the greater is the space over which it can be detected (even outside of the location at which it is generated). Therefore, two closely spaced sources may end up fused into a single blur.

Analysis of Surface Data: Surface Maps and 3D Inference

As previously mentioned, some psychophysiological measures (such as ERPs and MEG) can be used to generate maps of the distribution of a particular physical parameter at the surface of the head (scalp). Usually, these maps are generated using a process of interpolation: a reduced number of surface observations (usually fewer than 30 but in some cases as many as 256) are used to derive representations of the activity over the whole head (or at least a segment of it). The interpolation procedure requires the assumption that the spatial sampling used represents the surface distribution of the underlying phenomenon in a satisfactory way.

As for the time domain, one issue is determining the minimum spatial sampling required to produce representative electrical (and magnetic) maps of brain activity. In general, information theory and common sense indicate that the spatial resolution (but not necessarily the localization) is limited to twice the distance between two adjacent measurement locations (which we can define as the “spatial Nyquist frequency”). At first glance it thus may appear that a greater spatial sampling will result in maps with superior quality. Indeed, some investigators have shown that the information content increases with the number of recording locations (Srinivasan et al. 1996), a result that makes intuitive sense. However, practical and theoretical problems may limit the advantage of increasing spatial sampling of ERP and MEG data. One limiting factor is the cost of devices required to record and analyze data from a large number of channels. Another limiting factor is that measurement errors are also likely to increase with the number of recording channels, resulting in a decreased signal-to-noise ratio.

To illustrate the problem, suppose that an experimenter is interested in running an ERP experiment to determine where a potential reaches its maximum over the scalp. Let us also suppose that, in this experiment, there exists a 0.01 probability that a large, inaccurate value (i.e., an artifact) may influence the measures at any given electrode location and that, if such an artifact occurs, the maps may identify a maximum at an erroneous location. If only a small number of electrodes are used (e.g., fewer than ten), then the probability of a false maximum is relatively small ($p < 0.10$) and some confidence can be put in the results (although the spatial resolution will be very low). However, as the number of recording locations increases, the probability that at least one location will provide an artificial value also increases and so likewise the probability of incorrectly localizing the maximum. If enough electrode locations were used (say, over 100), the probability of an incorrect localization would become very high ($p > 0.63$). In this case, little confidence could be put on the results. This example underscores the need to use the appropriate number of recording locations. It also emphasizes the difference between spatial resolution and localization: procedures that enhance spatial resolution do not necessarily enhance localization, and vice versa.

As mentioned earlier, most maps are constructed using an interpolation process. There exist various algorithms for interpolations. A relatively simple procedure involves using a linear interpolation to derive the values in between the observed locations. One problem with this method is that the maximum of activity in the maps is bound to be at one of the locations at which the recordings were obtained; if two maps are constructed using different locations, they may differ from each other even though the actual physiological situation does not vary. Therefore, more sophisticated techniques are now commonly used. Some of them are based on fitting polynomial surfaces – or other types of surfaces obtained with some form of smoothing function – through the various observed data points (e.g. spline interpolation; Perrin et al. 1987). This interpolation may increase the localization of activity in surface maps, but at the cost of reduced spatial resolution. In addition, assumptions need to be made about the spatial frequencies contained in the data.

Some transformations of the surface maps may be useful for counteracting some limitations of the original observations. For instance, problems with surface maps of electrical potential over the head include: (a) the absolute value at each data point is dependent on the choice of reference; and (b) activity tends to be spread out across extended areas. This may make it difficult to appreciate differences between conditions. Addressing both of these issues is the transformation from maps of voltage difference to those of electric currents (“current source density” or CSD maps). According to Ohm’s law, currents flowing between two points are directly related to the differences in potential and inversely related to the resistance between the points. If the resistance is assumed to be constant, then current maps can be easily computed from voltage maps. In the case of the head, the current can be considered to flow almost exclusively along the skin, a path of lower resistance than the skull underneath. The flow of current in two dimensions can then be described as a current density, which is computed using the two-dimensional spatial derivative of the voltage map. This provides a measure of a local gradient, one that is independent of the definition of the reference level (since a local reference system is used)
and highlights local details of the maps. However, phenomena that extend across large areas become less visible, and small error variance may be amplified by this method.

Although CSD maps (and other transformed maps) of surface potentials may be useful for highlighting differences between conditions, they should not be interpreted as tridimensional reconstructions of activity inside the head. Such a reconstruction would require a model of (1) how the signal is generated inside the head and (2) how it is transferred from interior sources to the surface. Both problems require different solutions, depending on the type of measure adopted. For electrical activity (EEG and ERPs), understanding propagation of the potential to the scalp requires information about the conductive properties of various media interposed between the sources and the surface electrodes (such as gray and white matter, cerebrospinal fluid, bones, and skin). The transmission of magnetic fields is less influenced by departures from homogeneity in the media. Until recently, all source analysis algorithms were based on abstract constructions (e.g., the "three-sphere model") that modeled the head as a simple geometric shape. In the last few years, more realistic head models have been derived using MRI or computerized axial tomography (CAT) scans, and the propagation of activity to the surface has been computed using finite computation methods. Still, knowledge about the in vivo conductive properties of head tissues is imperfect and may create substantial errors in the computation.

In the last 15 years, substantial developments have been made in the modeling of the sources of electrical and magnetic activity. Three types of approaches can now be used:

1. the regional sources approach, which provides more reliable statements but has low spatial resolution (up to several centimeters);
2. linear source models, in which the influence of various cortical locations is estimated based on the assumption that activity is generated perpendicular to the cortical surface (these models possess intermediate spatial resolution, but their accuracy is difficult to evaluate); and
3. point-dipole models, which usually combine information accrued over time (thus generating a spatiotemporal dipole) – this method has greater spatial resolution and localization power, but it also has a greater level of uncertainty related to the number of sources modeled.

By and large, source modeling of electrical activity (and to some degree of magnetic activity) still requires external validation.

### Analysis of Tridimensional Spatial Data

Imaging methods provide three-dimensional descriptions of how the signal is generated inside the body. If the data refer to changes in the strength of the signal as a consequence of the functioning of a particular organ, these methods are called "functional" imaging methods. Examples of such functional techniques include fMRI and PET. Functional imaging methods vary in terms of spatial resolution and localization power, as well as along other dimensions (e.g., how direct the measures are and how much temporal resolution they possess). The great localization power of some of these techniques constitutes their major advantage, but it also creates a number of methodological problems. One of these is that the anatomy of some structures (and of the brain in particular) is quite variable across subjects. Therefore, a signal that appears at one location in one subject may appear at another location in another subject. In order to generalize conclusions across subjects, it is indispensable to "align" the data obtained in different subjects. This process requires rescaling the various anatomical features of each individual to a "standard" anatomy, so that phenomena observed in different subjects can be directly compared. A commonly used procedure to so align different brain anatomies was introduced by Talairach and Tournoux (1988). According to this procedure, structural scans of brains of different individuals (which can be obtained using MRI or CAT scans) are first aligned using a set of basic anatomical features. The brain is then subdivided into areas, and each area is independently rescaled to the dimensions of a standard provided by the Talairach and Tournoux atlas. This allows researchers to compare data across different subjects in assessing the reliability of findings and even to compute "average" brain maps across subjects. Using the standard atlas, it is then possible to associate specific functional data to particular anatomical structures of the brain (such as the Brodmann cytoarchitectonic areas of the cortex). This procedure is likely to introduce small distortions, since the large subdivisions used in the Talairach and Tournoux system may not capture all of the individual differences in anatomy. However, it is a standard method that can provide useful approximations in a number of cases.

An additional problem with three-dimensional spatial data is the huge number of dependent variables used. A three-dimensional scan of the brain based on voxels (the small volumes used for sampling a large volume) with a 1.5–3-mm side may generate more than 100,000 data points (each of which should be considered as a dependent variable). The statistical analysis of data sets with such a large number of dependent variables is very complex. Possible solutions to this problem can be found in specialized publications (see Friston 1996).

### FILTERING AND ARTIFACT PROBLEMS

One of the major issues in processing any type of data, especially psychophysiological data, is that of distinguishing the signals of interest from noise. Noise can be broadly defined as any phenomena observed in the data other than the signal(s) of interest to the investigator. In general, the ability to distinguish signals from noise can be quantified...
using a construct called the signal-to-noise ratio (S/N – the amplitude ratio between the signal and the noise). The larger the S/N, the easier it is to identify the signal. As a consequence, more reliance can be made on the observations. One of the major tasks of signal processing is therefore to increase the signal-to-noise ratio. This can be achieved by amplifying the signal, reducing the noise, or both. Procedures or devices that reduce the amount of noise present in the data are generally called filters.

Filters are based on the principle that the signal can be distinguished from noise on the basis of some characteristic feature. For instance, the signal may have a particular frequency that is different from that of the noise. Therefore, by amplifying the frequencies carrying the signal and dampening the frequencies carrying the noise, it is possible to increase greatly the S/N. This is common practice in psychophysiological recording, and it can be achieved while the data are recorded (through the use of on-line, usually analog, filters) or at any stage of data analysis (with off-line, usually digital, filters).

Filters are usually described in terms of the frequency at which they attenuate the signal. High-pass filters cut off activity with a frequency lower than a designated frequency, and low-pass filters cut off activity with a frequency higher than a designated frequency; notch filters cut off a selected frequency range and leave unaltered any activity with lower or higher frequencies. Usually, the cut-off point of a filter indicates that, at that frequency, activity is reduced by 3 dB (approximately 70%). However, other parameters are important in describing the performance of a filter. Specifically, the performance operating characteristic (POC) function describes the proportional attenuation of the signal at different frequencies. Ideally, a filter should maintain all of the activity up (for low-pass filters) or down (for the high-pass filters) to the cut-off frequency and eliminate all of the frequency above (or below) the cut-off frequency.

Most psychophysiological recordings involve on-line, analog filters. The advantage of these filters is that they can reduce high frequencies before digitization occurs, which may reduce aliasing of high frequencies. As already mentioned, aliasing occurs when digitization is performed at a rate that is less than twice that of a frequency present in the data. In such cases, activity at a high frequency will be reflected (aliased) as activity at a lower frequency. Aliasing cannot be eliminated once it occurs, so it is very important to prevent it by filtering high frequencies before digitization. Another advantage of on-line filters is that they can be used to cut off large noise oscillations, which usually occur at very low frequencies. These large oscillations may generate signals that are outside the operating range of the analog-to-digital converter, which then will “saturate” (i.e., it will provide an output corresponding to its own maximum or minimum reportable value, not to the actual value).

On-line frequency filters are, however, usually recursive: the filtering is based on activity that has already occurred and not on activity yet to occur. Hence, they introduce a phase distortion in the data. Phenomena may appear to peak at a different time from when they actually do, or their shape may be distorted. On-line low-pass filters (which cut off high frequencies) tend to displace peaks to a later time than when they really occur, and on-line high-pass filters (which cut off low frequencies) tend to displace peaks to an earlier time than when they really occur. In contrast, off-line filters can be nonrecursive and therefore need not produce phase distortion. The extent to which time displacements occur depend on the frequencies that are cut off by the filters. It is therefore preferable that data be recorded on-line using a wide bandpass filter and that filtering used for signal enhancement be mostly carried out off-line using nonrecursive filters.

Various types of off-line digital filters are available. Since they need not be analog, it is relatively simple to produce filters with good performance (i.e., sharp cut-off points). Filters with such characteristics can be built using a frequency-domain transformation of the data, but the operations required are quite complex and may render data processing exceedingly slow. It is possible to construct filters that work in the time domain and approximate the performance of these filters (for reviews see Cook & Miller 1992; Farwell et al. 1993).

Wiener (1964) introduced the concept of “optimal” filter. This is a filter whose POC function is optimized to maximally increase the signal with respect to the noise. For instance, a filter may be built to maximize the between-condition variance with respect to the within-condition variance. To build such a filter it is necessary to define the signal in some manner, so that the frequency carrying the signal can be identified a priori. In addition, an optimal filter can be practically designed only in the frequency domain – and, as just mentioned, filtering in the frequency domain may considerably reduce the speed of signal processing. These problems have limited the use of optimal filters.

In certain cases, the research interest is studying activity characterized by very specific frequencies. For instance, an investigator interested in the response to stimuli presented at regular intervals may assume that the response also should be observed at regular intervals. In this case, the frequency of the signal is well known – either the frequency of stimulation or one of its harmonics. Thus, activity at all other frequencies may be discarded as noise. An inverted notch filter (called a “bandpass” filter) may be used to selectively analyze the response and may greatly enhance the signal-to-noise ratio. Procedures of this type are used to analyze steady-state evoked responses (Tomoda et al. 1991).

So far, our discussion has focused on filters used in the temporal domain to eliminate activity at undesirable frequencies. However, filters can also be built in the spatial domain to amplify selected activity with particular spatial
properties; one example is the vector filter procedure proposed by Gratton, Coles, and Donchin (1989a). The vector filter procedure is analogous to a planned contrast executed in the space domain. Gratton et al. (1989b) and Fabiani et al. (1987) showed that this procedure may help identify ERP components.

A special type of noise that is present in most psychophysiological data is known as artifact. This term is normally used to indicate large, isolated noise activities that originate outside the system of interest. Examples of artifacts are the potentials associated with blinks and other types of eye movements during the recording of brain electrical activity, susceptibility and other artifacts due to movements during MRI and PET recordings, and missed heart beats in measures of heart rate. Artifacts may be very large—several times larger than the signal—and thus can completely obliterate the signal. Further, they may occur in a systematic fashion; that is, the frequency of artifacts may be consistently greater in one experimental condition than in another. However, artifacts are often easily recognizable because they have specific “signatures” that are readily distinguishable from regular data. For instance, blinks generate electrical potentials that exhibit characteristic spatial distribution over the scalp, and missed heartbeats in heart rate measures produce long intervals that are well outside of the normal variability.

A number of manual and automatic procedures have been developed to detect, eliminate, or compensate for the effects of artifacts. These procedures vary from one physiological measure to another and thus will not be reviewed here. It is important, however, to remember that appropriate procedures for dealing with artifacts are an essential step in the processing of any physiological signal.

Data Reduction and Quantification

QUANTIFICATION

Practically all analyses of physiological measures require a step in which a small number of numerical values are extracted from the large amount of data recorded and are then used for inferential statistics. Usually, this “quantification” itself involves two steps: (1) identifying a particular feature that is considered to represent a particular physiological event; and (2) measuring some parameter of this feature.

Feature Identification

The procedures used to identify the feature of interest are quite variable. In some cases, the feature can be identified by visual inspection of the data. This occurs when the S/N is so large (i.e., greater than 3:1) that the feature of interest can easily be distinguished from sources of noise or from other signals. High-S/N examples include blinks in electro-oculographic (EOG) recordings, EMG responses, and R waves in the ECG. In other cases, the features of interest are buried under variable and sometimes large amounts of noise or under other signals. An example of this is given by various components of the ERP, especially when measured from single trials or from averages of a small number of trials. In this case, identification of the component of interest may be quite complex and require sophisticated pattern recognition algorithms, filtering procedures, or a combination of both. Thus, identifying the P300 component of the ERP on single-trial recordings may require the use of a relatively heavy low-pass filter (band-pass 0–6 Hz or less) and pattern recognition algorithms such as cross-correlation methods etc. (see e.g. Fabiani et al. 1987; Gratton et al. 1989a). A step that is often required for identifying ERP components is the definition of a particular “time window” in which the component is expected to appear. Thus, the P300 may be defined as a positive peak in the ERP appearing in the interval of 300–700 msec; a peak appearing earlier or later may be classified as a different component. Other attributes may also be considered important for classifying data, such as the particular location at which the activity is observed.

Measurement

The second step of the quantification procedure is measuring a particular parameter of the feature of interest. Again, parameters vary along different dimensions. By and large, however, three dimensions are usually considered: (i) temporal dimension, (ii) spatial dimension, and (iii) intensity or frequency of occurrence.

Temporal Dimension. In the case of temporal dimension, activity can be considered as having a particular latency, frequency, or phase. Measures include onset, peak, center-of-gravity, and fitting of particular functions to the data. The two most commonly used measures are onset and peak latency.

Onset measures may be very informative in defining the latency of a particular physiological event (and therefore the maximum latency of the psychological phenomenon it is intended to signal). Onset can be measured as the first data point in the time series exceeding a preset value (this value can be obtained from previous work or from statistical computations of the variability in the measurements; Miller, Patterson, & Ulrich 1998). However, the exact onset time of a physiological measure can be difficult to determine when the measures contain even low levels of noise. In the case of noisy measures (such as ERP activity measured on a small number of trials), onset of a particular activity can be estimated by fitting regression lines to different segments of the data (Barrett, Shibasaki, & Neshige 1986) and then determining the point at which the regression line corresponding to the pre- and postonset periods meet. An alternative procedure is to consider the
half-amplitude latency, which is the latency required for the signal to reach half of its maximum peak (other proportions of the maximum value can be used; see Smulders, Kenemans, & Kok 1996). This alternative approach is quite reliable (Smulders et al. 1996), but determining what proportion of the maximum value to use is arbitrary and so the measure may not correspond to any particular psychological (or even physiological) phenomenon. In addition, if the maximum value is systematically different across conditions, then any determination of the latency of half-amplitude measures may be misleading.

An example is presented in Figure 8.

Peak measures are obtained (in the time or frequency domain) by selecting the time-series data point at which the measure reaches its maximum value. If several peaks are present, the peak point can be defined as the maximum value within a certain interval, or as the “nth” peak (the first peak, the second peak, …). Peak measures tend to be more reliable than onset measures. However, they may be sensitive to high-frequency noise and so it is often advisable to apply a low-pass filter before estimating the peak. The main advantage of peak latency measures is that they are easy to take, even for noisy data. A disadvantage is that, unlike onset measures, they do not necessarily identify a point in time that is of theoretical importance. Whereas we may infer from the onset of a particular activity that a particular psychological process must have been carried out for the activity to occur, the peak of the physiological response may occur some time later and have no specific meaning. An extreme example is given by the time course of the hemodynamic response (measured with fMRI) following the presentation of a train of stimuli. The moment at which the fMRI response reaches its peak may provide little information about the time course of psychological events. On the other hand, onset of fMRI activity in a particular brain area can be considered as a maximum limit for the latency of activation of that area (although, of course, activity in the area may in fact begin some time before it is detected). A similar situation exists for the latency of the peak of the skin conductance response. The peak latency of both fMRI and skin conductance signals may provide little information about the timing of psychophysiological events, yet the onset and the amplitude of either response is useful in understanding the nature of those events.

In general, temporal dimensions are considered as interval measures and hence are most commonly analyzed using parametric statistical approaches. However, temporal measures usually have a limited range. Moreover, in some cases the distribution of the measurement error departs substantially from normality, and skewness and platikurtosis may both be present (Fabiani et al. 1987). Both of these problems may lead to reduced power for statistical analysis.

Spatial Dimension. The spatial dimension is most commonly analyzed in terms of the correspondence between the location of a physiological response and a particular anatomical structure (e.g., “the blood oxygen level-dependent (BOLD) fMRI response was observed in proximity to the calcarine fissure”). With brain imaging techniques, analysis is often carried out in parallel for different data points (known in this field as “voxels”). Then, the voxels showing a response are identified and compared with anatomical maps. There is wide variation in methodology used to determine which voxels (or sets of voxels) show signs of functional activation (i.e., changes in blood flow or BOLD signal). In most cases, some type of statistic is derived for each voxel (or group of voxels). For methods based on hemodynamic or metabolic phenomena, this statistic often involves comparisons across conditions. In this case, the identified voxels are those that show differences between conditions exceeding some criterion, which is often expressed in terms of probability of alpha error.

The response is usually observed over several contiguous voxels. Indeed, given the high probability of false positives in brain imaging studies, responses are often considered meaningful only if they extend over a number of voxels. In this case, the exact location of the activity may be difficult to identify. One procedure is to consider the geometrical center of the region (group of voxels) showing significant effects. An alternative solution is to interpret the data...
as indicating a response that extends over the whole area where a significant response was observed. There are two possible problems with this latter interpretation.

First, the size of the area where a significant response is observed depends on the criterion used for considering an effect as "significant" and on the power of the measurements. Therefore, the activated area may depend on the number of trials or number of subjects used in the study. Paradoxically, if enough trials were collected (or if the S/N were sufficiently high), then larger areas of the brain could show signs of activation and the response would become less localized.

Second, for some measures (e.g. BOLD-fMRI), a large activity in one voxel may actually cause other voxels (both contiguous and noncontiguous) to show activity also. (This response pattern is due to the various procedures used to derive the measure.) For this reason, a strong activity localized in a small volume may produce effects that are difficult to distinguish from those of a weaker response localized to a larger area.

The statistical analysis of spatial information is sometimes undertaken by considering location as a categorical independent variable (a factor). For instance, the analysis of ERP scalp distribution is often carried out using electrode location as a factor in a factorial ANOVA (analysis of variance) design. Effects of experimental variables on scalp distribution would then be visible as interaction between these factors. Although quite popular, this approach has been criticized for two reasons.

First, the ANOVA model is an additive model - so that the contribution of electrode location as a factor to the overall variance is considered as independent from other factors. However, since ERPs always reflect the difference in voltage between locations, if a phenomenon influences the voltage of an ERP activity then it is likely also to influence different electrode locations in a different manner. Therefore an interaction between electrode location and any experimental factor may be due to a change in the overall size of a particular ERP activity, to a change in its spatial distribution (or scalp distribution), or to an interaction between these factors. Some investigators (McCarthy & Wood 1985) have proposed that, before submitting ERP data to an ANOVA, the values obtained at different electrodes should be rescaled by a factor related to the absolute size of the activity (e.g., the range of values or the standard deviation across all electrode locations). This standardization process is intended to eliminate variance across conditions or subjects due to changes in component amplitude, so that any effects observed as a function of electrode location can be interpreted as being due to a change in the spatial distribution of the ERP activity. This approach appears to be justified when the ERP observed at the scalp (at all scalp locations) is determined by a single component. However, as shown in Figure 9, for overlapping components the standardization procedure does not eliminate the confounding between differences in scalp distribution and differences in component amplitude.

Second, the ERP values measured at different electrode locations are in fact observations of the same phenomenon from different vantage points. For this reason, both the signal and the noise can be correlated. The correlation of the noise observed at different electrode locations (or at different data points) generates a problem in the ANOVA model that is commonly known as "lack of sphericity" (see Jennings & Wood 1976). A solution to this problem is

Figure 9. Effect of standardizing ERP scalp distribution data in conditions with different component amplitude and component overlap. Top row: Data from a condition in which there is no component overlap (i.e., all the activity observed at all electrode locations can be attributed to one component). On the left are plots of the observed and standardized scalp distribution data from a condition in which the component is small (standardization is obtained by dividing the data by the standard deviation across electrode locations). In the middle are similar plots for a condition in which the component has medium size. On the right are plots for a condition in which the component has large size. Note that plots of the standardized scalp distributions are identical for all component sizes. Bottom row: Plot from conditions with component overlap. A component with central-maximum scalp distribution and fixed amplitude was added to the observed data presented on the top row. Note that the scalp distribution appears to change from a frontal to a central maximum as a function of component amplitude. Note also that the difference in scalp distribution is not eliminated by standardization.
to use a multivariate approach to the analysis of scalp distribution data. Several multivariate approaches have been proposed, including multivariate ANOVA (Vasey & Thayer 1987); multiple regression (e.g. the vector filter; Gratton et al. 1989a); principal component analysis; and single-value decomposition (Lamothe & Stroink 1991).

Other investigators have considered using nonparametric statistics to evaluate location information. A recent example is the use of a "bootstrapping" procedure for studying the location of the peak of a surface distribution of ERPs or optical activity (Fabiani et al. 1998). The purpose of the bootstrap method proposed by Fabiani et al. is to estimate the reliability of maxima of maps of ERP or EROS activity obtained on average data. The bootstrap method involves analysis of a large number of samples of individual trials (bootstrap replications). Each bootstrap replication is obtained by extracting at random (and with replacement) a number of trials from the original data set equal to the number of trials used to compute the regular average. For each new bootstrap replication, an average map is computed, and the location of maximum is determined. The distribution of the maxima across the different bootstrap replications is then obtained. The location of a maximum is reliable if it is obtained in a large proportion of the bootstrap replications. This bootstrap method does not require that we make assumptions about the distribution of maxima in the population, and it is therefore quite robust. The method can be modified to consider the probability that the maximum is within a certain area. One limitation of the method is that it is ad hoc, although other bootstrap procedures can be used for different applications (see Wasserman & Bockenholt 1989). Some investigators have proposed other "distribution-free" methods for the analysis of spatial distribution data (Karniski, Blair, & Snider 1994).

### Intensity

Intensity is usually quantified using one of four approaches: (1) peak amplitude, (2) integrated activity over time (also called area measure), (3) covariance with a template (Fabiani et al. 1987), and (4) frequency of response. These measures are illustrated in Figure 10.

The first two types of measures are heavily dependent on the definition of a "baseline" level (i.e., they are taken by computing the difference between the peak — or the sum of the individual values — and the baseline), whereas the covariance measure requires some hypotheses about the shape of the response to be studied (the response shape can also be obtained with statistical methods — see the next section). Area measures are less sensitive to high-frequency noise than are peak measures. However, empirical studies have shown that if a low-pass filter is applied before the peak measurement, the latter can be at least as reliable as area measures (Fabiani et al. 1987). Covariance measures can be used to separate the contribution of overlapping physiological responses — in this case, a multiple regression approach can be used — and are quite reliable. Examples of covariance measures include estimates of the amplitude of different ERP components obtained with principal component analysis (PCA — Donchin & Heffley 1978) or with stepwise discriminant analysis (SWDA — Donchin 1969; Donchin & Herning 1975; Squires & Donchin 1976). In these cases, the covariance measures are taken with respect to a "weighting" function that has been optimized on the basis of statistical properties of the data (see next section). More recently, multiple regression approaches have been used to analyze hemodynamic data (fMRI, PET, etc.) in the case of complex, factorial designs. An advantage of this approach over the "subtraction" methods traditionally used to analyze such data is that it enables identification of effects associated with interaction terms on the brain image maps. However, the assumptions underlying these methods (i.e., linearity and additivity) remain as critical for these measures as they are for subtraction methods.

All these measures are considered as providing results along interval scales, which can be analyzed using parametric statistics. The error distribution tends to be skewed (since usually the range is limited on the lower end), but platikurtosis is not a problem as it was for latency measures. Amplitude measures may therefore be more reliable than latency measures (Fabiani et al. 1987).

Measures of the frequency of response are used for signals that are quite stereotyped and easily recognizable yet
occur only on a certain proportion of the trials or only at irregular intervals. Note that, in many cases (e.g., eye blinks or electrodermal responses), response frequency will not follow a normal distribution; the distributions will have a heavy positive skew, with some subjects exhibiting a large number of responses and others very few or none. In this case, experimental variables will generally produce larger effects for the subjects exhibiting a greater base rate of responses than for those exhibiting a smaller base rate. For instance, all subjects may double their response frequency as a function of experimental manipulations; for those with a low base rate this may generate very little effect, whereas for those with a high base rate the effect may be huge. This results in a large subject × manipulation interaction with consequent reduced statistical power. In cases like this, a logarithmic transformation of the response frequency may help increase the statistical power of the analysis. Alternatively, inferential analysis can be conducted using a log-linear approach (Kennedy 1983). For a discussion of other approaches (standardization, rescaling, etc.), see Ben-Shakhar (1985) and Stemmler (1987).

**STATISTICAL APPROACHES TO DATA REDUCTION**

As mentioned previously, psychophysiological data are essentially multivariate. One of the major steps in data analysis is the reduction of the sometimes bewildering number of dependent variables to a smaller number that is appropriate for inferential analysis. Statisticians have developed a number of procedures to deal with the problem of data reduction on the basis of the statistical properties of the data. One of the simplest and most popular approaches is PCA (see Donchin & Heffley 1978). Principal component analysis is a statistical procedure for grouping dependent variables in a smaller subset of underlying (or “latent”) variables that have the following properties: (a) they explain as much as possible of the variance and covariance of the original set of dependent variables; and (b) they are orthogonal to each other (in other words, they are uncorrelated). The components can be rotated (using procedures such as VARIMAX) to ease interpretability. The PCA components are interpreted in terms of their correlations (or covariances) with the original variables (also called component loadings). Donchin and his collaborators wrote a series of papers outlying the application of PCA to the analysis of ERP data. In this approach, ERP data are viewed as a data matrix in which data points are viewed as variates (i.e., dependent variables in a multivariate approach) while different subjects, conditions, and electrode locations are viewed as individual observations. A standard PCA is then run on the covariance matrix obtained from this data matrix. In most cases, a VARIMAX rotation is run on the results of the PCA. The component loadings are used to describe different components of the ERPs. These loadings are then used as sets of weights to compute component scores for each subject, condition, and location. The component scores are used as estimates of the amplitude of each component for each of the original waveforms and are submitted to inferential analysis (e.g., ANOVA). This approach has been used in a number of studies (Karis, Fabiani, & Donchin 1984; McCallum & Curry 1984; Ruchkin, Sutton, & Stega 1980; Squires, Squires, & Hillyard 1975), mostly because it provides separate estimates of the amplitude of individual components even when they overlap in time.

This PCA/VARIMAX/ANOVA approach has been criticized (see Möcks 1986; Möcks & Verleger 1985; Wood & McCarthy 1984) on the grounds that variance is misattributed to different components. Indeed, the subdivision of variance into different components as obtained with PCA is entirely arbitrary, since there is an infinite number of ways to subdivide the variance on the basis of the same number of components. Therefore, it is unjustified to consider the components obtained with PCA in terms of physiological entities per se. Further problems with the use of PCA in the analysis of ERPs are the difficulty of incorporating latency shifts within the PCA model (Möcks 1986) and possible correlations between components. However, PCA remains an interesting attempt at a principled way for data reduction, one in which as much as possible of the original variance is retained while the number of independent variables is minimized. As already mentioned, this is one of the major goals of signal processing.

More recently, similar approaches have been proposed to address the issue of data reduction. For instance, Maier and colleagues (1987) used a “spatial” PCA (in which the dependent variables used are spatial locations rather than data points) as a preliminary step for source analysis of ERPs. Spatial PCA has also been used to determine the number of active components at a certain moment in time as a preliminary step for dipole analysis (see the BESA of Scherg & von Cramon 1986). A sequence of spatial and temporal PCAs was employed by Spencer, Dien, and Donchin (1997), whereas Möcks (1988) proposed a multivariate approach that also combines spatial and temporal information. Singular-value decomposition has also been proposed as an alternative model to PCA (see the EMSE of Greenblatt 1996). As with PCA, a problem with most of these approaches is that the result is not unique, since (an infinite number of) alternative descriptions of the data are possible. Further, it is often difficult to interpret the result of the various data reduction procedures used and to compare results obtained in different experiments.

**OTHER QUANTIFICATION ISSUES**

**Absolute versus Relative Effects: The Problem of Baseline and Reference**

In a number of cases, psychophysiological measures are intended to reflect a change in a physiological variable from...
a level that exists prior to the introduction of an experimental manipulation (such as the presentation of a stimulus or the administration of a drug) to a level that exists after the manipulation. The level of the physiological variable before the manipulation begins is called the resting or baseline level.

For some variables, this level is characterized by the absence of any measurable activity (for instance, there may be little EMG activity in a muscle at rest). Most often, however, the targeted physiological system is already active prior to the experiment and remains active during the experiment (this is clearly true for the heart and the brain). In this case, the baseline level may change as a function of subject, location, and time of measurement. For this reason, most physiological measures are expressed as changes with respect to a baseline level. Determination of an appropriate baseline level is not always simple, because it may not be possible to induce a pure “rest” condition. In most cases, even when the experimental condition does not require a measurable overt response, the subject may be active predicting and evaluating situations from both a cognitive and an emotional point of view. In addition, the subject may be paying attention to internal or external stimuli that are not experimentally controlled.

All of these psychological activities may influence the measurements obtained during the baseline period. If they do so in a manner that is systematically different from that observed at the time of the experimental measurement, or in different ways in different experimental conditions, then there is the possibility the systematic effects may influence the measurement. For instance, in most experiments using ERPs, measurements are taken as differences with respect to a prestimulus baseline level. However, if subjects can anticipate certain properties of the stimulus during the prestimulus period, it is possible that an anticipation component of the ERP (e.g., the contingent negative variation or CNV; Walter et al. 1964) may be elicited during this interval. The differential measures taken at a later time will then be affected by the presence of a CNV during the baseline period. Likewise, if the dependent variable is heart rate then the expectation of an external stimulus may induce heart rate deceleration (Lacey et al. 1963). Note, however, that if the research interest is in comparing conditions in which anticipation is expected to be (on average) the same then this is not necessarily a critical problem, since a similar value is subtracted from both conditions. For this and like reasons, it is generally easier to compare activity observed in different experimental conditions than to draw conclusions from activity observed in a single condition alone. In other words, it is much easier to consider relative than absolute effects.

Whereas the baseline problem is common to most physiological measures, some measures (such as ERPs) also entail the problem of the reference. The term “reference” is used to indicate a comparison value that is valid for all points in space; this contrasts with the term “baseline,” which is used to indicate a comparison value that is valid for all points in time. In some cases, physiological measures are obtained by considering the difference in electric potential between two locations. Electric potentials do not possess absolute values and are defined as differences between two energy levels. Therefore, measures of electric potentials should be considered with respect to the locations that are involved. It is, however, possible to compare the difference in potential that exists between point A and point B with that between point C and point B. If the first difference is more positive than the second then we can conclude that point A has a more positive voltage level than point C. In this case, point B will be used as a common reference point. If measures were taken only at a single point in time, which point (A, B, or C) is used as a common reference would make little difference. However, voltage measures are generally taken at different times, and points of maximum difference (or peak values) are identified. The point at which A is most different from B may not be the same point at which A is most different from C. Therefore, the time of the occurrence of the peak on A may differ depending on whether B or C is used as a reference. For this reason, appropriate selection of a reference location is a critical aspect of recording electrical activity.

Historically, there have been three approaches to the problem of reference selection. The first approach is to consider measurements obtained between selected pairs of electrode locations. Such bipolar measurements are common practice in the study of EMG, ECG, and EOG. Bipolar derivations can be used to study local phenomena: the two electrodes are placed very close to each other, so that activity generated far away is likely to influence them equally (and therefore to cancel out) whereas activity generated locally is more likely to generate differences. This approach is adequate when the research interest is studying the time course of activity of relatively simple electric field configuration – such as for EMG, EOG, and (to some extent) ECG. For EEG studies, the bipolar approach fails to account for the complexities of the fields generated by the brain and is therefore used only in certain clinical applications or special cases (e.g., recording auditory brainstem averaged evoked potentials, BAEPs).

A second approach, which uses a common reference system, is to consider measurements from a number of locations as differences with respect to the same location (or to a common value). There are a number of choices for the location of the reference value, including the ears, mastoids, nose, forehead, or locations outside the head (extracephalic reference). The latter may pick up electrical activity from the heart (ECG) and therefore require special compensation procedures (Fortgens & de Bruin 1983). Finally, some researchers have advocated the use of an average reference, which is obtained by subtracting or adding a value to all of the observed locations so that the algebraic
sum of all of the potential values observed at different locations is equal to zero. Although arguments in favor of one or another of these systems have been proposed (see e.g. Skrandies & Lehmann 1982), there is evidently no clear advantage to using one reference system rather than another: the choice is arbitrary. However, since the selection of the reference may alter the shape of waveforms (including both the amplitude and latency of the peaks), it is very important that the reference selection be made explicit and that the same reference system be used when comparing across different data sets. The common reference system is the one most used in ERP research. Its greatest advantages are ease of computation and the possibility it offers of comparing values obtained at different locations. Its major drawback is the arbitrariness of the choice of the reference.

A third approach is to transform the data so that they are expressed in terms of absolute dimensions (such as current flow) rather than as relative dimensions (such as difference in potential). As we have seen (see the section entitled “Analysis of Surface Data”), this can be achieved by computing the local gradient in potential rather than the difference with a common reference (with the assumption of constant resistance between electrode locations).

**Subtraction, Comparisons, and Linearity of the Measures**

In most cases, psychophysicists are interested in comparing physiological responses in two (or more) conditions. The comparison process usually involves a study of the statistical reliability of differences between measures. This, in turn, involves computing several differences across subjects, time points, and locations. Data obtained in different conditions can be interpreted either along ordinal scales (which require using less powerful nonparametric statistics) or along interval scales (which may be analyzed using parametric statistics). The basic assumption of interval scales is that differences between intervals at any level of the scale are directly comparable – for example, the difference between the values 2 and 4 in a given scale has the same significance as the difference between the values 22 and 24. This allows the investigator to subtract common terms and to compare data directly. We previously considered the similar property of linearity (see the section entitled “Direct versus Indirect Measures”).

However, most psychophysiological measures depart to some extent from linearity. One of the reasons is that they often have a limited range because physiological systems feature feedback mechanisms that counteract extreme values in order to maintain the body’s homeostasis. For example, heart rate is unlikely to drop below 40 beats per minute even under strong vagal activation. There are also other limitations due to the number of units (e.g. neurons) that can respond to particular stimuli. Even so, for most psychophysiological measures there exists an interval of values for which linearity can occur. For practical purposes, it is advantageous for measurements to be taken – if at all possible – in conditions where the psychophysiological measure is within the range in which it exhibits linearity.

Range and scale differences may also vary as a function of subject or of location on the body. For example, certain subjects may exhibit higher variability in a psychophysiological variable than others. Similarly, measures of a psychophysiological variable may have a greater range at one location than another. This will make it more difficult to compare data across subjects or locations. To counteract this problem, standardized measures (in which the data are transformed so as to have equal means and standard deviations) are often used in these types of comparisons.

For a number of psychophysiological variables, it has been found that variability is correlated with the average level of the variable itself for a particular subject or condition. This relationship sometimes takes the form of “ceiling” or “floor” effects, which in psychophysiology often have been considered as examples of the law of initial values (Wildar 1957). Such effects produce departures from linearity, and psychologists have often used special transformations (e.g. logit or probit) to correct such departures. Sometimes the effects are actually proportional to the original value, in which case a logarithmic transformation of the data (or the division by the average value) may equalize variances (and differences) across conditions – and thereby produce a scale that has linear properties. For a discussion of several procedures for comparing conditions with different base rates, see Wainer (1991).

**Concluding Remarks**

In this chapter we examined several issues related to the analysis of psychophysiological data. In most cases, data analysis procedures can be viewed as an effort to extract meaningful information from data that are often noisy. In these conditions, a major part of the data analyst’s work is to increase the signal-to-noise ratio. Although there exist procedures that increase this ratio under all conditions, it is often the case that a priori hypotheses about what type of signal to expect may help in the design of appropriate analysis methodologies. These hypotheses need not be that specific (e.g., suggesting that activity at a particular latency at a particular location is determined by factor A or factor B). Even simple hypotheses – that some activity should be observed somewhere during a latency interval – help in designing appropriate analytical methods. In general, however, the power of the analytical procedure increases with the specificity of the hypotheses that are entertained.

On the other hand, it is often the case that psychophysiological phenomena are quite complex, and unpredicted effects are obtained in a number of studies. In some cases, discovery of these new effects is what affords the greatest scientific advancements. This generates a dilemma for the
data analyst: How can maximum power be obtained while maintaining a wide focus in the analysis? In most cases the answer lies in alternating studies conducted with an "exploratory" attitude (enabling investigators to detect unexpected findings) with other, "hypothesis-driven" studies (which provide more rigorous tests of specific hypotheses). This approach emphasizes the need for data analysis to be integrated with experimental design in a bidirectional fashion.

NOTES
I thank Monica Fabiani, John Cacioppo, Lou Tassinary, and Gary Berntson for their comments on previous versions of this manuscript.
1. Note that this formula may not apply to all cases and is only provided as an example of the typical influence of delaying factors on the time course of physiological measures. In fact, not only the time course but also the intensity of the output with respect to the input may vary from case to case.
2. There are, however, notable exception to this rule, especially in the study of reflexes. For instance, Hackley and Johnson (1996) used properties of the blink response (a peripheral measure) to determine whether a particular phenomenon (prepulse inhibition) is under the control of cortical or subcortical structures (clearly an issue of function localization within the brain). This indicates that, if appropriate experimental designs are used, peripheral measures can provide some information about the localization of function within the brain.
3. This phenomenon can be minimized by using special type of filters (e.g., elliptical and/or Bessel filters).

REFERENCES


