Available online at www.sciencedirect.com

SCIENCE DIRECT®

Biological Psychology 67 (2004) 183–218

BIOLOGICAL PSYCHOLOGY

www.elsevier.com/locate/biopsycho

Methodology

Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion

John J.B. Allen*, James A. Coan, Maria Nazarian

Department of Psychology, University of Arizona, P.O. Box 210068, Tucson, AZ 85721-0068, USA

Abstract

There exists a substantial literature examining frontal electroencephalographic asymmetries in emotion, motivation, and psychopathology. Research in this area uses a specialized set of approaches for reducing raw EEG signals to metrics that provide the basis for making inferences about the role of frontal brain activity in emotion. The present review details some of the common data processing routines used in this field of research, with a focus on statistical and methodological issues that have captured, and should capture, the attention of researchers in this field.

Keywords: Frontal electroencephalographic asymmetry; EEG; Emotion; Methods; State and trait

The field of research examining frontal electroencephalographic (EEG) asymmetries in emotion and psychopathology is now over two decades old, with over 80 published studies documenting relationships between asymmetries in frontal EEG power and emotion-related traits and states (see Coan and Allen, 2004, this issue, for review). Although data reduction and analytic techniques have varied across studies, there are many common approaches that have become quite popular for transforming raw EEG signals to metrics that provide the basis for making inferences about the role of frontal brain activity in emotion. These approaches involve many transformations of the data, and in that process involve assumptions that can impact the interpretations scientists can levy from a given pattern of results.

The aim of this paper, therefore, is to provide a general overview of some of the common steps involved in data processing in this field, highlighting the assumptions and the impact of violations of these assumptions for interpreting findings. It is important to note that none of the issues raised in this paper call to question the now well-replicated relationships between

---

* Corresponding author. Tel.: +1-520-621-4992.
E-mail address: jallen@u.arizona.edu (J.J.B. Allen).

0301-0511/$ – see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.biopsycho.2004.03.007

the metrics of EEG asymmetry and emotional constructs. The issues will, however, have implications for how the findings best be interpreted.

1. From raw signals to handy metrics

Investigators who examine frontal EEG asymmetry use a set of relatively specialized signal processing routines, which will be reviewed anon. This review is not intended so serve as a primer for basic signal processing, but rather is designed to highlight the data reduction trail typically followed in this specific research domain. For a basic primer, many sources are available, including easily accessible chapters by Gratton (2000) and by Reilly (1987), and a more in depth treatment by Glaser and Ruchkin (1976).

Fig. 1 depicts the many steps typically involved in transforming electroencephalographic signals into metrics that putatively are related to how active various brain regions may be. This process involves taking a signal collected in the time-domain (Panel A, left side), and converting it to a frequency-domain representation, usually in the form of a power spectrum (Panel A, right side). This spectrum, which collapses data across time, summarizes which frequencies are present to greater or lesser degrees in the time-domain signal. Whether data are collected from an extended resting period involving several minutes, or from discrete and relatively short emotion-related segments, this spectral analysis approach always involves examining the frequency composition of short epochs (Panel B), on the order of 1 or 2 s each, and averaging power spectra across many such epochs. In the case of resting data, this involves epoching a large data segment into many smaller epochs. In the case of EEG acquired in the context of fleeting emotional expression or experience, the data segment might still require being epoched into a few smaller epochs, and data from several such expressions or experiences would then be aggregated.

By using epochs that are only 1 or 2 s-long, one more closely approximates an assumption underlying the Fourier transform, the method used to derive power spectra from raw signals. Fourier analyses assume a periodic signal (the stationarity assumption), and moreover that any periodic signal can be decomposed into a series of sine and cosine functions of various frequencies, with the function for each frequency beginning at its own particular phase. A periodic signal is one that repeats, and does so at uniformly spaced intervals of time. Although strictly speaking EEG signals are not periodic, as the repetition of features is not precisely spaced at uniform intervals, by selecting short epochs one can analyze small segments of data that will have features that repeat in a highly similar fashion at other points in the waveform.

Epoching typically involves the construction of overlapping epochs (Panel B), as weighting functions applied in the process of "windowing" (described below) prior to frequency analysis result in the central portion of the epoch receiving the most weight, and distal portions receiving negligible weight (Panel C). By overlapping the epochs, all data points receive maximum weighting in some epoch.

Windowing is used to avoid creating artifactual frequencies in the resultant power spectra. Because Fourier transforms assume a periodic signal, it is assumed that the signal in the epoch repeats infinitely both forwards and backwards in time, and without the windowing function to reduce the ends of the epoch to near-zero values, discontinuities would

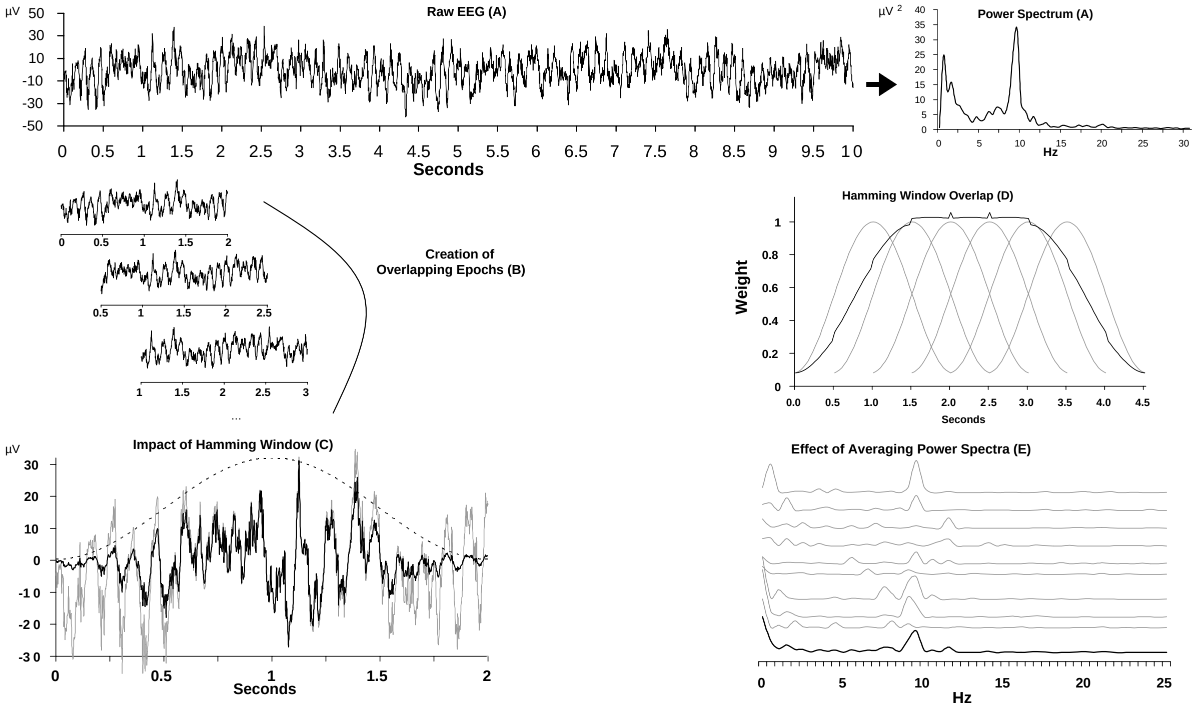


Fig. 1. Depiction of the various data reduction steps typically used in frontal EEG asymmetry research. Panel A depicts a 10-s segment of raw data from a single channel on the left, and the spectral representation of this epoch on the right. Panel B illustrates the process of epoching the longer segment into shorter overlapping 2-s epochs. Panel C depicts the impact of the Hamming window (dotted bell curve) on a single epoch, with the gray line representing the raw signal and the black line representing the signal after the application of the window. Note that a discontinuity would result if a copy of the raw (gray) signal were concatenated following this signal, but no such discontinuity would result for a similarly concatenated windowed (black) signal. Panel D displays the net weighting (black line, scaled to fit graph) of overlapping hamming windows (gray lines) for 2-s epochs. Panel E illustrates the impact of averaging power spectra. The top nine gray lines are the spectral representation of nine 2-s epochs, and the lower black line is the average spectrum. Note that alpha power (8–13 Hz) is somewhat variable from epoch to epoch, but that the average spectrum reveals a distinct alpha peak. Vertical axis in Panel E is power in μV^2.

result if one were to place a copy of the epoch immediately before or after itself. Fourier methods would introduce a variety of spurious frequencies to reconstruct a signal with such discontinuity. By windowing, the discontinuity is avoided (Panel C), but at the expense of preventing the data near the end of the epoch from being fully represented in the resultant power spectrum. The overlapping of epochs (Panel D) provides a solution to this latter problem, as data minimally weighted at the end of epoch x will be weighted more heavily in epoch $x + 1$.

Most computer signal processing packages use a fast Fourier transform (FFT), which as the name implies is considerably faster and computationally less complex than the more general case discrete Fourier transform (DFT). The FFT requires that the epochs to be analyzed have 2^n data points. Data are often sampled at a rate that is a power of two, thus allowing epochs of 1 or 2 s, but in other cases of sample rates that deviate from a power of two (e.g., 250 Hz), epoch length will need to be tailored accordingly (e.g., 2.048 s). For a data segment of 1024 data points, the DFT will take about 10 times longer to arrive at the same result as the FFT.[1]

The result of the FFT is two spectra, a power spectrum and a phase spectrum. The power spectrum reflects the power in the signal at each frequency from dc to the Nyquist frequency,[2] with a spectral value every $1/T$ points, where T is the length of the epoch analyzed. The phase spectrum presents the phase of the waveform at each interval $1/T$. These two spectra can jointly be used to reconstruct the original time-domain waveform. Psychophysiologists, however, typically discard the phase spectrum and focus their analyses only on the power spectrum.

As an FFT is applied to each epoch, many power spectra result, and the average of these power spectra is ultimately taken as the basis for analysis (Panel E). The data in this resultant spectrum might entail between 20 and 200 data points (the precise number being $T \times (f/2)$, dependent on the epoch length T and the sample rate f), a substantial reduction from the raw data signal that will likely have hundreds of data points per second for several minutes. The spectra represent, therefore, a relatively economical representation of the original signal, with higher sampling frequencies and longer epochs resulting in more spectral points. Further reduction is accomplished by summarizing data within conventionally-defined frequency bands. Alpha power, either total (μV^2) or density (μV^2/Hz), is most often examined, and is typically operationalized as power between 8 and 13 Hz in adults, although lower frequencies have been examined in children (for review see Coan and Allen, 2003b), as these lower frequencies in the developing brain are assumed to be equivalent to adult alpha

---

[1] The DFT transform is a general case instantiation of the Fourier transform for discretely sampled signals, but it is computationally intensive, with the time taken to compute the spectral representation being proportional to the square of the number on points in the series. The comparable computation time using the FFT, by contrast, is proportional to $N(\log_2(N))$. For an epoch of 1024 points ($N = 1024$), the DFT will take 102.4 times longer than the FFT to compute the spectral representation of the signal.

[2] The Nyquist frequency, named after Henry Nyquist, is the fastest frequency that can be represented for a given sampling rate, and is equal to 1/2 the sampling rate. Nyquist, whose entire career was at AT&T Bell Laboratories, published a 1928 paper (Nyquist, 1928) in which he proposed a theorem that a sample rate twice as fast as the highest signal frequency will capture that signal perfectly. Stated differently, the highest frequency which can be accurately represented is one-half of the sampling rate, and this frequency has come to be known as the Nyquist frequency.

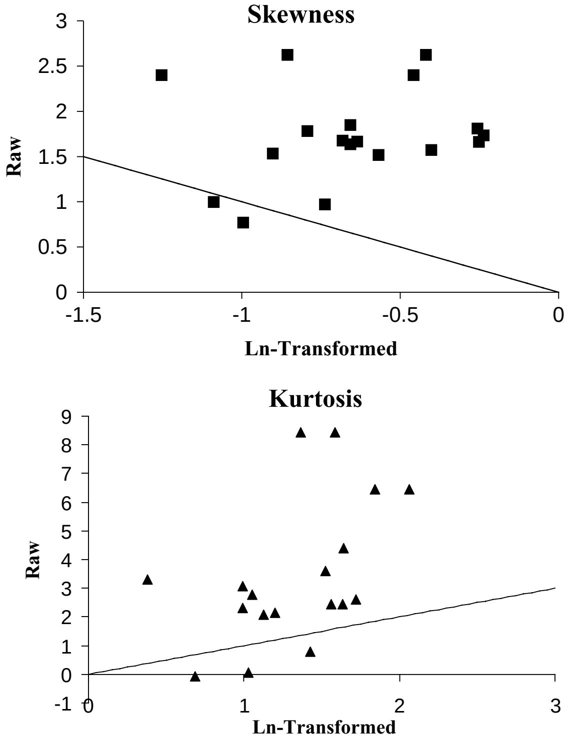


Fig. 2. Skewness statistic (top panel) and Kurtosis statistic (lower panel) for natural log transformed (*X*-axis) and raw (*Y*-axis) power values. Statistics were calculated on 34 subjects with complete resting EEG data reported in Coan and Allen (2003a) for each of 18 scalp sites (FP1, FP2, F3, F4, F7, F8, FTC1, FTC2, C3, C4, T3, T4, T5, T6, TCP1, TCP2, P3, P4) using the average reference. The solid line in each plot represents the demarcation between improvement towards normality (above the line) from greater deviation from normality (below the line) as a result of the natural-log transformation.

(e.g., Fox and Davidson, 1987). Alpha power is then taken as an index of the inverse of cortical activity (Davidson, 1988), an assumption that will be explored further below.

Alpha power at any given site then is typically natural log transformed, as untransformed power values tend to be positively skewed, as depicted in Fig. 2. The top panel of Fig. 2 depicts the Skewness Statistic for the raw power values (*Y*-axis) and the natural-log transformed values (*X*-axis) at each of 18 scalp sites. The lower panel similarly depicts the Kurtosis statistic for the same data set. The solid line in each plot represents the demarcation between improvement towards normality (above the line) and greater deviation from normality (below the line) as a result of the natural-log transformation. As can be seen from the figure, the transformation improves the skewness for 89% of the scalp sites, and improves kurtosis for 83% of the scalp sites. In absolute terms, using the 95% confidence intervals, prior to natural-log transformation, 94% of sites deviated significantly from normality in terms of skewness, and 83% deviated significantly in terms of kurtosis. Following transformation, only 33 and 39% of sites still deviated significantly from normality in terms of skewness and kurtosis, respectively. Thus although some sites still exhibit distributions of natural-log transformed scores that deviate from normality, the natural-log transformation substantially improves the distributional characteristics of the data.

2. Comparing left and right activity

Because asymmetrical activity is of interest, investigators often use a difference score ($\ln(\text{Right}) - \ln(\text{Left})$ alpha power) to conveniently summarize the relative activity at homologous right and left leads.[3] The difference score thus provides a simple unidimensional scale representing the relative activity of the right and left hemispheres, with higher scores putatively indicative of *relatively* greater left frontal activity (assuming that alpha is inversely related to activity). An additional benefit of this difference score metric is that it provides some degree of correction for overall alpha power, as large individual differences in overall alpha power could be confounded with the magnitude of the asymmetry. The correction for overall power stems from the fact that the natural log difference score metric is not a simple difference score, but a difference of natural log transformed scores. Rules of logarithmic subtraction state that the difference of two natural-log transformed scores is equivalent to the natural log transform of the ratio of these scores:

$$\ln(R) - \ln(L) = \ln\left(\frac{R}{L}\right) \tag{1}$$

Thus this difference metric is actually the natural log transform of the ratio of right to left alpha power, which provides some degree of correction for overall power expressing each subject's asymmetry in terms of a ratio. The extent of the correction is confirmed by comparing the values of the natural log difference score metric to another sometimes-utilized metric, a "normalized" difference score computed as $(R - L)/(R + L)$. This normalized difference score metric correlates over 0.99 with the natural log asymmetry metric ($\ln(\text{Right}) - \ln(\text{Left})$; Allen et al., 2004). There is in fact a nonlinear function relating these two metrics over a broad range of scores, because when either R or L gets very small, the normalized metric is bounded by the values 1 and -1 and the natural-log asymmetry metric will not have such bounds. Over the range of values encountered in asymmetry research, however, the function is almost perfectly linear, as illustrated in Fig. 3.

The difference metric is rather handy in several respects, notably that it mitigates the impact of individual differences in skull thickness that would have sizeable influences on recorded signal amplitude (Eshel et al., 1995; Leissner et al., 1970; Pfefferbaum, 1990), and the difference scores can simplify analyses, such as those involving correlations between frontal asymmetry (as a difference score) and an individual difference measure (e.g., Be-

---

[3] Similar distributional improvements as a result of natural log transformation are seen for the asymmetry scores based on these log transformed values. Comparing asymmetry scores based on the difference of natural-log transformed and untransformed values, the transformation improves the skewness of the asymmetry score for 67% of the scalp sites, and improves kurtosis for 89% of the scalp sites. In absolute terms, using the 95% confidence intervals, prior to natural-log transformation, 67% of the differences scores deviated significantly from normality in terms of skewness, and also kurtosis, but following transformation, only 22 and 33% of asymmetry scores still deviated significantly from normality in terms of skewness and kurtosis, respectively.

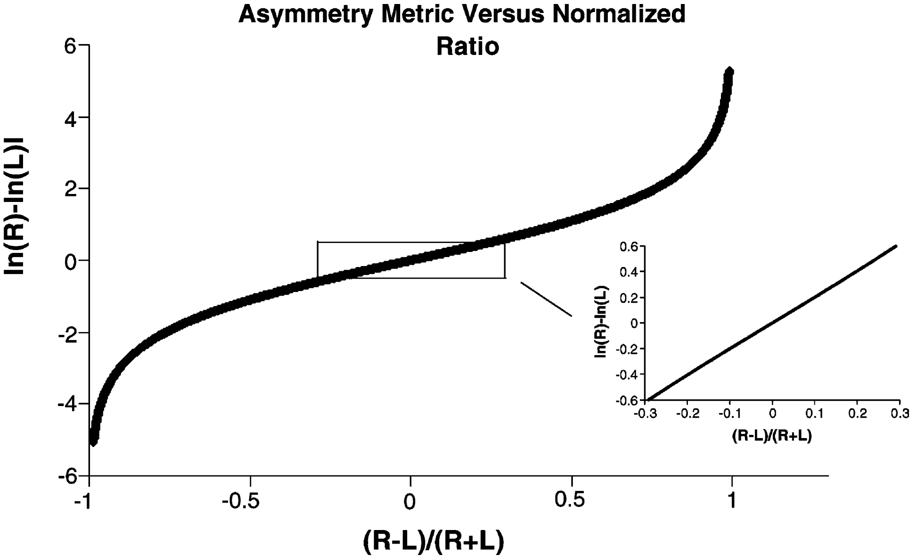


Fig. 3. The relationship of the asymmetry metric (ln(Right) − ln(Left)) and a metric normalized for overall power $((R-L)/(R+L))$, over a large range of possible alpha power values. In asymmetry research, the ln(Right)−ln(Left) metric produces scores that typically are in the range of ±0.5, the range demarcated by the two lines, where the relationship is linear. From Allen et al. (2004), reprinted with permission from Blackwell Publishing. © 2004, Society for Psychophysiological Research.

havioral Activation Scale; Coan and Allen, 2003a; Harmon-Jones and Allen, 1997; Sutton and Davidson, 1997).

Difference scores have been criticized for their potential unreliability, as errors of measurement with each of the constituent scores are compounded when the difference score is calculated. The reliability of change scores is of greatest concern, however, when the constituent scores have modest reliability. Alpha power at a given lead, however, demonstrates extremely high reliability, with coefficient alpha values typically over 0.90 based on 8 min of data. Moreover, the reliability of the difference score for frontal regions has been calculated in several studies and routinely is high, excepting frontal pole sites (e.g., coefficient alphas for frontal asymmetry (difference) scores ranging from 0.85 to 0.90 at baseline in Allen, Urry, Hitt, and Coan (2004); from 0.76 to 0.91 in Coan and Allen (2003a); a median of 0.83 in Coan et al. (2001); from 0.80 to 0.93 in Reid et al. (1998); and from 0.81 to 0.92 in Tomarken et al. (1992)).[4]

---

[4] A separate issue concerns the power of statistical tests that employ difference scores. The power of significance tests using difference scores is only indirectly influenced by the reliability of these scores. Significance tests of differences can be powerful even if the reliability of the difference scores is near zero (Overall and Woodward, 1975; Zimmerman et al., 1993). The paradox pointed out by Overall and Woodward (1975) is that difference scores with zero reliability can in fact give rise to high power to detect a significant difference. The paradox is resolved when one considers that reliability of the difference scores depends on the existence of variance in the difference score that can reliably rank-order individuals in terms of the *magnitude* of their difference scores, but that the power to detect a difference involves assessing a *mean* difference between the two scores relative to the variance in

Despite the simplicity of the difference score, the contribution of activity in each hemisphere is ultimately of interest, which will require analyses involving the examination of data from each hemisphere as a difference metric is uninformative with respect to the contribution of each constituent hemisphere (Davidson et al., 2000a). The most straightforward approach involves analyzing (ln-transformed) power at left and right sites in an analysis of variance (ANOVA) or the more general linear model (GLM), with not only region (anterior to posterior) as a factor, but hemisphere (left versus right) as well. In these models, individual differences in overall power are removed, and region specific variations in power (e.g., occipital alpha is greater than frontal alpha) are partitioned as region main effects. In such a model, with EEG power as the dependent variable, the interaction of an independent variable with hemisphere will yield the same information as a main effect of this independent variable when using asymmetry scores as the dependent variable. The follow-up tests to decompose the interaction can then examine the contribution of each hemisphere individually.

The standard ANOVA approach provides a straightforward method of examining the power from each hemisphere's lead or leads when the independent variable of interest is amenable to the ANOVA approach, such as when comparing depressed and nondepressed subjects, or when comparing two or more emotion elicitation conditions. This approach is limited, however, as the standard ANOVA with a between subjects or within subjects factor fails to allow for an examination of the power at a given lead or leads with a continuously varying independent variable such as ratings of emotional valence or intensity, or an individual difference variable such as behavioral activation. There exist a few published approaches that have included a continuous predictor in the model, the whole-head and homologous-lead residualized power approach first reported by Wheeler et al. (1993), and the hierarchical general linear model strategy (e.g., Coan and Allen, 2003a) or mixed model strategy (e.g., Kline et al., 2002).

2.1. Residualized power approach

Wheeler et al. (1993) adopted a two-stage analytic approach, examining first the correlation between the asymmetry difference score and continuous measures of self-reported

---

this difference score. Thus if one constituent score (e.g. Left activity) were for every subject a constant *k* less than the other constituent score (e.g. Right activity), then there would be no variability in the difference scores, and no reliability. On the other hand, the mean difference score would be *k*, with no variance around that mean, allowing for a powerful statistical test that the mean difference is significantly different than zero, and that a statistically significant difference has been found. The pragmatic implications are that the reliability of difference scores if are of little consequence if one wishes to test the significance of such a difference (e.g. to test that Right activity is greater than Left activity for the group as a whole), but the reliability of the difference score will be highly relevant when one is using the difference score to examine how individual differences in that difference score relate to other variables of interest (e.g. how individual differences in the asymmetry score relate to individual differences in BAS scores). In the latter case, the reliability of the difference score will impose constraints on the magnitude of the correlation that can be observed, as the maximum correlation that can be observed between two variables will be the square root of the product of the reliability of the two variables. Thus, because a sizable portion of the research examining frontal EEG asymmetry is concerned with the relationship of individual differences in frontal EEG asymmetry to other individual difference measures, the reliability of the asymmetry metric assumes great importance.

Table 1
Correlations between natural-log transformed alpha power at homologous leads collected for 8 min under resting conditions

Sites	Reference	
	AR	LM
FP1–FP2	0.997	0.998
F7–F8	0.983	0.971
F3–F4	0.990	0.992
FTC1–FTC2	0.975	0.943
C3–C4	0.977	0.981
T3–T4	0.918	0.891
TCP1–TCP2	0.944	0.948
P3–P4	0.965	0.982
T5–T6	0.907	0.932

Note. AR: average reference, LM: computer linked mastoid reference; data from 34 subjects reported in Coan et al. (2001).

affect. Upon finding significant correlations, the second stage was to investigate the contribution of each hemisphere, but unconfounded by the large individual differences in power due to irrelevant factors such as scalp thickness. Power at a given electrode (e.g., F3) was residualized, using a hierarchical regression, first entering the average power across available scalp sites, and as the second step entering power from the homologous lead (e.g., F4). The resultant residualized values were then correlated with the variable of interest (e.g., self-reported affect).

The first step of this procedure preserves individual patterns of activity across scalp sites, adjusted for overall power. The second step of this procedure was introduced by Wheeler et al. (1993) ostensibly to statistically account for volume-conducted activity from the homologous electrode. It is unclear why one would be more concerned with volume conduction from a lead over the opposite hemisphere, which in many instances is considerably further away from the site of interest than ipsilateral leads adjacent to the site. On the other hand, the activity between homologous leads is often highly correlated, with alpha power values being correlated on the order of 0.95 or even higher (see Table 1),[5] and could in part reflect the dense contralateral cortico-cortical connections between some homologous regions as well as volume conduction effects. As seen in Table 1, correlations are uniformly high, but higher yet between closely spaced homologous leads (e.g., FP1 and FP2) as compared to more widely spaced homologous leads (e.g., T5 and T6). Whether volume conducted, or the result of interconnectivity, the second step of the regression approach of Wheeler et al. then statistically controlled for shared variance between left and right homologous leads, which is likely to be substantial.

---

[5] The fact that the difference between these highly correlated sites is nonetheless predictive of state affect and individual differences merits a brief comment. The asymmetry score reflects the difference between the contribution of the activity of the left and right leads *within* subjects, whereas the correlations between sites reflect the similarities of activity at each lead *across* subjects. It is thus the case that between-person differences in alpha power at a given site are substantially larger than the within person differences between sites, but that the latter nonetheless have some degree of predictive validity.

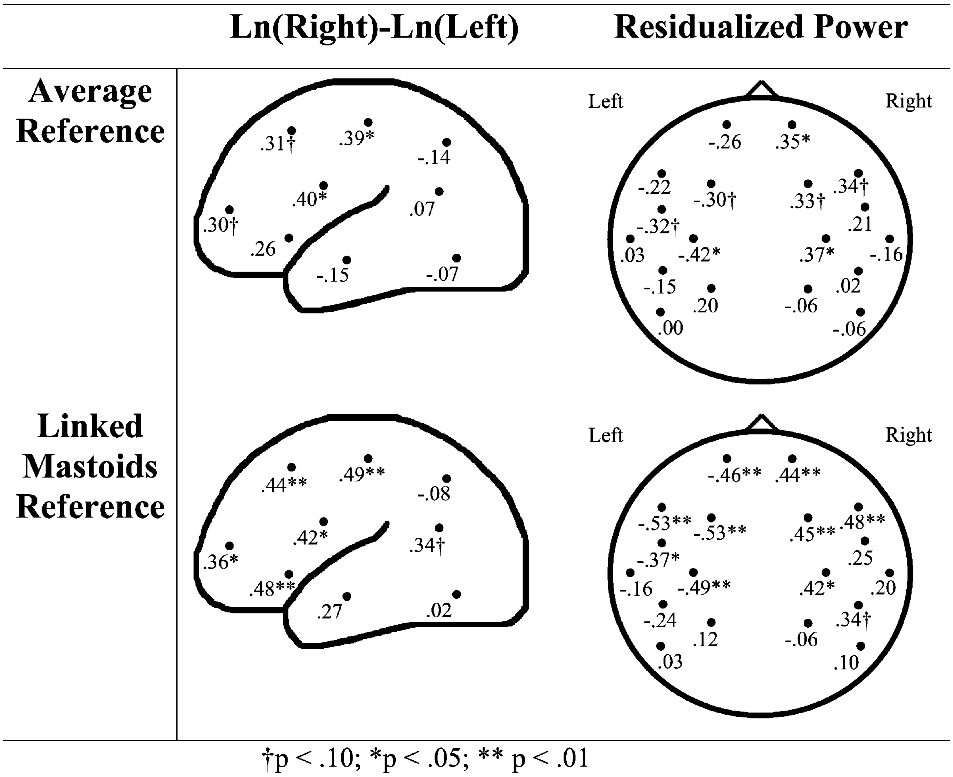


Fig. 4. Correlations between Behavioral Activation Scale (BAS) scores and EEG asymmetry (left panels) and residualized power at constituent sites (right panels), for data under an averaged reference (AR) and computer-averaged mastoids (CAM) reference. Data from subjects presented in Coan and Allen (2003a).

The results of this procedure produce what has become a fairly typical pattern: for each significant correlation between the $R - L$ difference score and a criterion, two significant correlations emerge, approximately equal in magnitude to the original correlation, but opposite in sign to one another, at the constituent leads. Correlations at the right lead maintain the sign of the $R - L$ difference score, and correlations at the left lead reverse direction. For example, Wheeler et al. (1993) found that the F4–F3 (ln-transformed) asymmetry score correlated with positive affect 0.45, and that residualized ln-transformed power at F4 correlated 0.44 and at F3 correlated −0.49 with positive affect. Similarly, Harmon-Jones and Allen (1998) found that the F4–F3 (ln-transformed) asymmetry score correlated 0.48 with trait anger, and that residualized ln-transformed power at F4 and F3 correlated 0.45 and −0.46, respectively, with trait anger. To illustrate more generally this pattern, Fig. 4 presents correlations between BAS scores and asymmetry scores (left panels) and residualized power at constituent sites (right panels). Treating these correlations themselves data points, the obtained values from residualized power at right leads correlated 0.94 with the values obtained using the difference score, and the values from residualized power at left leads correlated −0.88 with those obtained using the difference score. Additionally, the values obtained using residualized left values correlated −0.87 with those obtained using residualized right values.

Thus, the procedure that was originally devised to examine the independent contributions of each hemisphere would appear to distribute the variance relatively equally and in opposite directions across the two hemispheres, which would be expected if activity at homologous leads is extremely highly correlated, as is the case with homologous left and right lead power (Table 1). To demonstrate why such a pattern would be expected with such highly correlated data, consider the impact on residualizing left hemisphere power on right hemisphere power. The residualized score (L_{resid}) for a left hemisphere lead (L) is given by

$$L_{\text{resid}} = L - \hat{L} \tag{2}$$

where $\hat{L}$ is the predicted power at the left hemisphere lead given power at the right hemisphere lead, determined by the raw score regression (prediction) formula:

$$\hat{L} = a + b(R) \tag{3}$$

where a is the intercept and b is the unstandardized regression coefficient. In the case where L and R are nearly perfectly positively correlated (see Table 1), with the distribution of each having virtually identical means and standard deviations, the intercept a will approach 0, and the regression coefficient b will approach one,[6] reducing Eq. (3) to:

$$\hat{L} \approx 0 + 1(R) = R \tag{4}$$

Substituting the results of Eq. (4) for $\hat{L}$ in Eq. (2), it is revealed that, when L and R are nearly perfectly correlated:

$$L_{\text{resid}} = L - \hat{L} \approx L - R \tag{5}$$

Thus this residualization procedure produces residual values for left hemisphere leads that will approach the value $L - R$ as the correlation between left and right leads approaches 1.0, provided that the unstandardized regression coefficient approaches 1 and the intercept approaches 0. Similarly, by implementing Eqs. (2)–(5) for right hemisphere residualized (R_{resid}) and predicted ($\hat{R}$) scores, it will be the case that residual values for right hemisphere leads will approach the value $R-L$ as the correlation between left and right leads approaches 1.0. Therefore, this procedure will make it appear that right hemisphere leads correlate with a criterion variable in the same direction and approximate magnitude as the $R - L$ difference score, and that left hemisphere leads correlate with a criterion variable in the opposite direction but same approximate magnitude as the $R - L$ difference score.

---

[6] Empirically, it appears to be the case that the unstandardized regression weight is very close to one and the intercept is very close to zero. Resting data for 34 subjects (from Coan et al., 2001) were used to predict left hemisphere frontal activity from the homologous right hemisphere frontal activity. For the prediction of four frontal sites (FP1, F7, F3, and FTC1), each from its homologous right hemisphere site (FP2, F8, F4, and FTC2), across both LM and AR reference schemes (for a total of eight separate regressions), the median unstandardized regression coefficient was 1.028 (range: 1.012–1.071) and the median intercept was −0.039 (range: −0.098 to 0.031).

2.2. Revised residualized power approach

More recently, Davidson et al. (2000a) have proposed an improved variant on the method of Wheeler et al. (1993), one that does not include homologous lead power in the residualization calculations. This approach first residualizes the criterion variable on whole head power, and then calculates correlations between the residualized criterion variable and power at each individual site (Davidson et al., 2000a, p. 41; Davidson, 2002, personal communication). This method obviates the problem detailed above using the homologous lead to residualize power at each site, but will produce a large set of correlations (one for each scalp site) that are not tested formally in a model that can control for experiment-wise alpha slippage. Such correlations are quite informative, but ultimately must be regarded as descriptive. To adequately test the relationship between power at each site and the criterion variable, an omnibus model is required. Although the precise model will depend on the nature of the investigation, and the theory being tested, an alternative approach might be for investigators to specify a hierarchical general linear model in testing the relationship of left and right sites to criterion variables, as highlighted below. Such an approach might limit the undesirable probabilistic artifacts involved in multiple statistical tests, optimizing risk for both type 1 and type 2 errors in estimating both the impact of whole head power and effects of interest. Further, such single model approaches may economize data analytic effort and reporting.

2.3. Hierarchical general linear models

Hierarchical general linear models can simultaneously account for the multiple sources of variance contributing to the relationship between cortical asymmetry and criterion variables. Such models can include both categorical and continuous predictors, and can be constructed to test a variety of specific hypotheses of interest, including those related to overall power, hemisphere, and even reference scheme, all in a single model. In fact, interactions with reference scheme can be entered into such a model in order to determine whether relationships between asymmetry and the criterion variable are dependent upon reference scheme.

In constructing the model, some general principles may guide the investigator. First, the model should be explicitly specified, and whenever possible should be an omnibus model that can test all effects of interest at once. Second, the investigator should use theory to guide the ordering of the main effects followed by the interactions of these main effects. In most cases, main effects per se will not be of interest (e.g., they may reflect the contribution of overall power to the prediction of the criterion variable, or differences in overall power between anterior and posterior regions), but the interactions will be of interest. Interactions of hemisphere and region in predicting the criterion variable, for example, would be found if there are frontally-specific hemispheric differences in the contribution of left and right leads to the prediction of the criterion variable. A higher-order interaction with reference scheme would further indicate that the pattern of findings is reference-scheme dependent.

Of course, the use of theory and careful sensitivity to the possibility of spurious interaction effects is particularly important as the complexity of interactions increases. As interaction effects reach beyond third order, the probability of overfitting the observed data—essentially modeling meaningless residual variance—increases. While testing for most or all effects of interest in an omnibus model is highly desirable, it is not advisable if doing so requires the modeling of very complex (e.g., greater than fourth order) interactions. Thus a third principle might be to reduce the potential for spurious findings and complex interactions in the hierarchical linear model designed to test for specific contribution of hemisphere by first running a simpler but conceptually related model using the asymmetry scores. Then, following this simpler model, natural-log transformed power at constituent sites can be entered for the relevant regions where the asymmetry score identified a relationship between asymmetry and the criterion variable.

As an illustration, data from Coan and Allen (2003a) are presented, the same data that were used in Fig. 4 to illustrate the residualization approach. First, an omnibus hierarchical linear model using asymmetry scores from eight regions across the scalp under both averaged reference and computer averaged mastoids reference schemes were used to construct a model predicting BAS scores. To code reference scheme, data from each reference scheme were concatenated, and a contrast-coded variable was used to code for reference scheme (cf. Aiken and West, 1991). The model first entered the main effect of reference scheme, followed by the main effects of regions ordered according to theoretical interest and results of previous studies. Sites entered first were frontal and anterior temporal sites, followed by sites from central to parietal: F4–F3, F8–F7, FTC2–FTC1, T4–T3, C4–C3, T6–T5, TCP2–TCP1, and P4–P3. Finally, interactions of each region with reference scheme were entered to test for the reference-specific effects. In this model, only the main effect of F4–F3 was significant in predicting BAS scores ($F(1, 46) = 8.5$, $P < 0.01$), with trends for contributions from the main effects of F8–F7 ($F(1, 46) = 3.6$, $P < 0.10$) and C4–C3 ($F(1, 46) = 3.6$, $P < 0.10$). Reference scheme did not interact with any effects in this model.

Thus the focus of the subsequent analysis was to examine the contribution of left and right hemisphere in the significant midfrontal region. In this hierarchical general linear model, BAS scores were the dependent variable to be predicted by (1) whole head power; (2) reference scheme and (3) natural log-transformed alpha power in the left (F3) and right (F4) hemispheres. This model, with whole head power entered first, is akin to the procedure described by Davidson et al. (2000a) to statistically partial out the effect of overall power in predicting BAS score. In this model, main effects of each site were of interest. Interactions with reference scheme were entered into the model in order to determine whether any relationships between site and BAS scores were dependent upon reference scheme.[7]

The overall model was approached statistical significance ($F(9, 54) = 1.94$, $P = 0.07$, adjusted-$R^2 = 0.12$). Results indicated a main effect of the right hemisphere at F4 ($F(1, 54) = 9.61$, $P < 0.01$, $\eta^2 = 0.15$) but not of the left hemisphere at F3 ($F(1, 54) = 2.49$, $P = 0.12$, $\eta^2 = 0.04$). No interactions with reference scheme were significant, indicating that individual differences in right frontal activity were related to differences in BAS scores, and that this main effect of the right hemisphere was not dependent upon reference scheme. Fig. 5 depicts specific left/right relationships with BAS scores for F4 and F3. To estimate regression lines for both F3 and F4 separately, two hierarchical gen-

---

[7] Had the first model included additional regions of significance, such a model could also include the interaction between hemisphere (left, right) and region (e.g. mid-frontal and lateral-frontal).

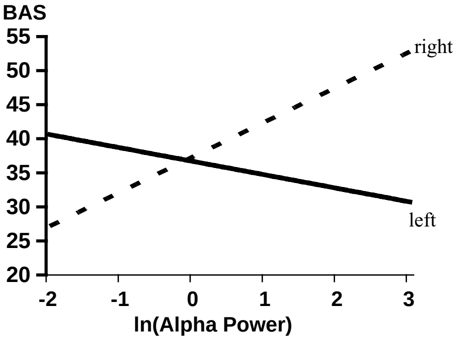


Fig. 5. Regression lines for left and right midfrontal (F3 and F4) ln-transformed alpha power predicting BAS scores. The regression equations depicted are: BAS = (5.12) × (Right) + 37.41; BAS = (−1.98) × (Left) + 36.658.

eral linear models were run, one each for F3 and F4. For these models, reference scheme, whole head alpha power, and each site were entered such that (1) the effects of reference scheme and whole head alpha power were each removed *before* the *b*-coefficient for each site was estimated, and (2) dependence upon both reference scheme and whole head power could be estimated. Notably, neither model is, by itself, statistically significant, but it is nevertheless useful to estimate such curves in the service of understanding the significant effects reported above. Apparent using this method, but not the residualization approach, is that the relationships between BAS and each hemisphere are not a mirror opposites, but in fact BAS is robustly related to right hemisphere activity, and largely not related to left hemisphere activity, a surprising result given the theoretical notions concerning the left hemisphere and approach-related motivation (Coan and Allen, 2003a,b; Davidson, 1992, 1998; Harmon-Jones and Allen, 1997).

3. Data acquisition

3.1. How much raw data should be acquired?

Sufficient data are required to ensure that reliable estimates of EEG activity are derived. Although the power spectrum derived from any single epoch via the FFT will reflect both frequencies that are common across epochs as well as those idiosyncratic to any given epoch, averaging together multiple spectra can allow those frequencies to emerge that are present in a reasonably large proportion of epochs (see Fig. 1, Panel E), while mitigating the influence of infrequent or irregular signals (Nunez, 1981), which might often be considered noise. Thus an investigator, by averaging across epochs, makes the implicit assumption that the frequencies that appear commonly across epochs are of interest, and epoch-specific variations are of little interest. In the case of estimating trait asymmetry with the goal of predicting psychological traits or psychopathology, this is clearly a reasonable assumption. On the other hand, a recent investigation found that variability from epoch to epoch was itself an important correlate of neuroticism (Minnix and Kline, 2004).

To reliably estimate EEG asymmetry at any given assessment session, investigators and reviewers often suggest that 8 min of resting EEG asymmetry are necessary to obtain adequate internal consistency reliability, as this was the number reported in the first psychometric investigation of resting EEG alpha asymmetry (Tomarken et al., 1992). Substantially fewer 1-min samples, however, also can produce acceptable estimates of internal consistency (Tomarken et al., 1992), and estimates based on even shorter time frames of 2 min have proven similarly reliable (Coan et al., 2001).

Tomarken et al. (1992) assessed the reliability of fewer than 8 min of data in a way that confounded the length of recording with the number of discrete items included in the calculation of coefficient alpha; i.e., they used the Spearman–Brown prophecy formula to estimate the reliability for shorter recording periods, estimating alpha based on six asymmetry values for 6 min of data, seven values for 7 min, and eight values for 8 min. To adequately test whether fewer minutes of recording would produce estimates of internal consistency comparable to those obtained with more minutes of recording, it would be required to keep the number of values constant despite changes in the length of recorded data, as Cronbach's alpha will be higher given more minutes (items) for analysis (Lord and Novick, 1968). In a recent study (Allen et al., 2004), reliability estimates from 2, 4, 6, and 8 min of data were compared. Specifically, the first 2, 4, and 6 min as well as all 8 min of recorded data were divided into eight blocks each. Each block contained 2-s overlapping epochs that were subjected to Fourier analysis as reviewed above. In each case, eight asymmetry values were obtained, reflecting the asymmetry score averaged across 1/8 of the total time of recording (15 s for the 2-min data, 30 s for 4-min data, 45 s for the 6-min data, and 60 s for the 8-min data). These eight values were then treated as items on an eight-item scale to assess internal consistency reliability.

Fig. 6 shows the results for frontal regions as a function of reference scheme. As can be seen in the figure, the number of minutes of recording exerts relatively little influence on the estimate of internal consistency compared to the number of blocks included in creating the estimate. Whether 2, 4, 6, or 8 min of data are utilized, very small differences are apparent when all eight data segments are used as items for the purpose of estimating internal consistency reliability. Reliability estimates begin to diverge, however, when fewer segments are utilized to estimate reliability. Thus highly internally consistent measures of asymmetry can be obtained with considerably fewer than the conventionally accepted 8 min of recorded data, provided that internal consistency is estimated with a sufficient number of constituent blocks. To highlight this point, consider a comparison of two comparable data points from Fig. 6: four 60-s blocks or eight 30-s blocks, which correspond to identical timepoints from the EEG record. In all nine cases (3 regions × 3 reference schemes), the internal consistency of the latter is higher than the former, by an average of 0.06 reliability units. It also appears to be the case that when fewer than four blocks are used to estimate the reliability, the expected rank ordering of reliabilities becomes less orderly, in some cases with longer recording blocks demonstrating lower reliability than shorter blocks. Thus, regardless of the total length of data collected, attempting to estimate reliability with insufficient blocks will lead to misleading estimates of internal-consistency reliability. If investigators have fewer than 8 min of data available, reliable estimates of asymmetry can likely be derived, but it is recommended that investigators report the internal consistency reliability of asymmetry scores based on the data

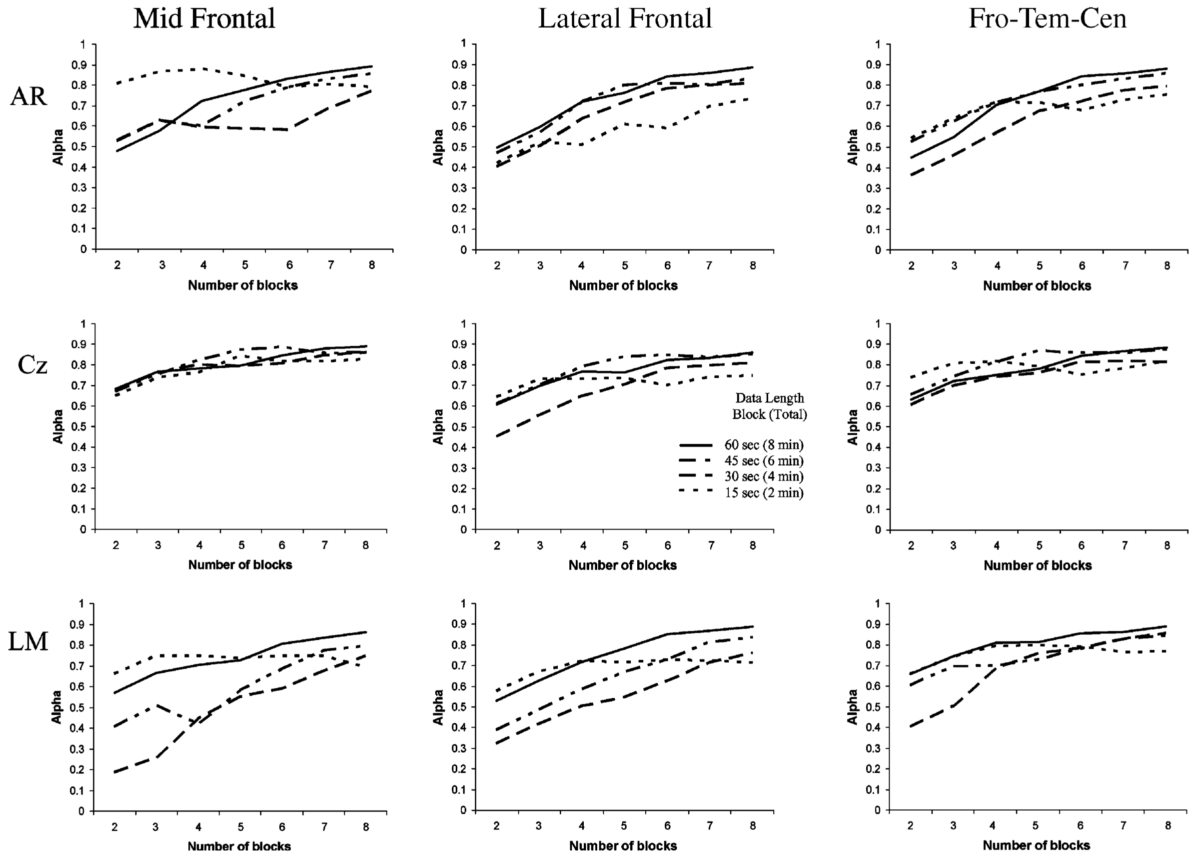


Fig. 6. Cronbach's alpha internal consistency estimates for resting alpha asymmetry as a function of region, reference scheme, length of data recording, and number of blocks (items) used to calculate alpha. The number of subjects ranges from 19 to 28, reflecting that some subjects did not have enough artifact-free 2-s epochs to compute power spectra for the for shorter recording intervals, or that a recording site was bad for a given subject. Midfrontal: F4–F3, lateral frontal: F8–F7, and Fro-Tem-Cen: fronto-tempo-central. From Allen et al. (2004), reprinted with permission from Blackwell Publishing. © 2004, Society for Psychophysiological Research.

available, detailing how many epochs were treated as items in the calculation of Cronbach's alpha.

3.2. What reference montage is preferred?

The choice of reference has been referred to as "perhaps the most divisive issue among current EEG researchers" (Davidson et al., 2000a, p. 33). Although rational arguments have been levied in favor of one or another reference scheme (e.g., Hagemann et al., 2001; Reid et al., 1998), it remains an empirical question which reference scheme has the greatest predictive validity with respect to motivation, emotion, and psychopathology. Investigators would ideally like measures of spectral power at a given site to reflect the activity at that site, and not at the reference lead. For this purpose, investigators often search for a relatively inactive reference, and have used linked ears or mastoids, averaged ears or mastoids, or an average reference comprised of the average of activity at all recorded EEG sites. The average reference, given a sufficiently large array of electrodes in a spherical arrangement around the head, will nicely approximate an inactive reference, as activity generated from dipoles will be revealed as positivity at one site and negativity at a site 180° opposite this site, with the sum across sites thus approaching zero with a sufficiently representative sample of the sphere. Smaller montages, and those that do not provide coverage approximating the sphere, however, will have more residual activity in the average reference.

Especially troubling is the Cz reference, which has been utilized more often in the EEG asymmetry literature than other reference montages (see Coan and Allen, 2003b for review). The Cz reference has been criticized as potentially under- or over-estimating activity at the target site (Hagemann et al., 2001). Moreover, empirical comparisons of data from different reference schemes have found Cz to be the least related to other reference schemes (e.g., Hagemann et al., 2001; Reid et al., 1998). The fact that many studies have successfully identified predicted relationships using the Cz reference suggests at least two non-mutually exclusive possibilities: (1) significant results using the Cz reference reflect, in part, not only the relationship of constructs with frontal asymmetry, but also with sources of variance unique to the Cz reference (e.g., overall alpha power); and/or, (2) asymmetry scores using the Cz reference may have more irrelevant variance (error or systematic) with respect to asymmetry, and may therefore result—across studies—in inconsistencies in the pattern of empirical relationships with motivation, emotion, and psychopathology. Distinguishing between these possibilities will be facilitated if investigators report results from multiple reference montages. Moreover, various reference schemes can be conceptualized as contributing unique sources of error variance to any given analysis, providing the researcher with semi-independent measures of EEG activity, with findings that are statistically independent of reference scheme being considered the most generalizable, being less likely to reflect only the reference-specific "method" variance (cf. Campbell and Fiske, 1959).

3.3. Impedances in asymmetry research

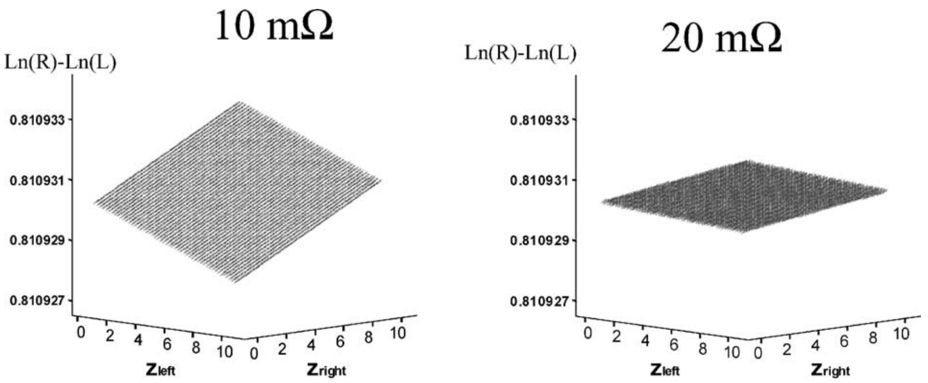
It has been customary in EEG asymmetry research to strive to obtain low and symmetrical impedances during subject preparation. Intuitively, this seems desirable, as one would wish to have a strong noise-free signal by lowering impedance, and would like to guarantee that any observed alpha power asymmetries reflect an underlying asymmetry in activity, rather than an asymmetry in impendence to recording the underlying signal. As pointed out by Ferree et al. (2001), however, contemporary high impedance amplifiers mitigate the impact of scalp impedances on the recorded signal, as the loss in the observed signal due to scalp impedances is directly related to the average impedance of the measurement and the reference electrode, and inversely related to the amplifier input impedance. Because amplifier input impedances are typically on the order of tens (or even hundreds) of megaohms, and scalp impedances on the order of a few kilohms, small changes in scalp impedance do not appreciably impact the observed signal, as the magnitude of the amplifier input impedance is *at least* 1000 times greater than the scalp impedance.

The observed voltage for a given electrode E with a given reference electrode R is the measured voltage difference between these electrodes, or $V_E - V_R$. This difference, $V_E - V_R$, is influenced by electrode impedance Z_E and reference impedance Z_R and input impedance Z_{in} as follows (Ferree et al., 2001, p. 538):

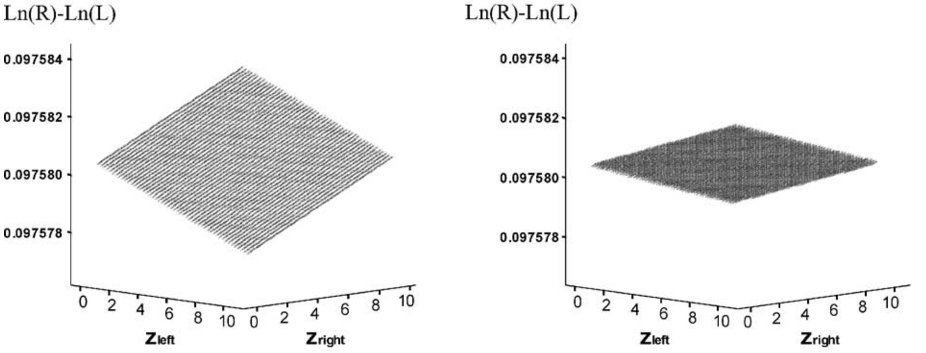
$$V_E - V_R = V_D \left(2 - \frac{Z_E + Z_R}{Z_{in}}\right) + V_C \left(\frac{Z_E - Z_R}{Z_{in}}\right) + O\left(\frac{1}{Z_{in}}\right)^2 \tag{6}$$

where V_D is the actual differential-mode signal $(V_{E\_true} - V_{R\_true})/2$, and V_C is the common-mode signal $(V_{E\_true} - V_{R\_true})/2$. The latter term V_C results primarily 60 cycle (US) or 50 cycle (Europe) ambient noise, and the extent to which it emerges is a function of impedance mismatch. The former term V_D is primarily the signal of interest, resulting from voltage potential differences between the two sites, but attenuated by the ratio of the scalp impedances to the amplifier input impedance. Since the scalp impedances are a tiny fraction of the size of the input impedance, even appreciable differences in scalp site impedance will not measurably attenuate the voltage potential difference observed between the two sites. The final term in the equation is a residual term to account for other sources in the differential amplifier circuit that influence the observed voltage potential difference, the sum of which are negligible (Ferree, 2002, personal communication). The mathematical notation O is standard for “order” in Taylor series, and simplifies the expression without appreciably altering the result obtained with the simplified equation involving only the first two terms. In the full equation there are a series of higher order terms involving powers of $1/Z_{in}$, which the term $O(1/Z_{in})^2$ denotes. With a high input impedance Z_{in}, the impact of $(1/Z_{in})^2$ will be negligible, and the impact of higher powers approaches zero.

Fig. 7 depicts the impact of mismatched impedances under conditions likely to be encountered in a psychophysiological laboratory. The data depicted in Fig. 7 show the observed asymmetry score (ln(Right) − ln(Left)) as a function of amplifier input impedance, and impedance at left (z_{left}) and right (z_{right}) leads. Data in the left panel depict the impact of mismatched left and right lead impedances with an input impedance of 10 mΩ (that of the Neuroscan Synamps system, Neuroscan a Compumedics Company, El Paso, TX), and data in the right panel depict the same with an input impedance of 20 mΩ (that of the Grass Model 12 Neurodata system, Grass Telefactor an Astro-Med Inc. Product Group, West Warwick, RI). Data were obtained by solving Eq. (6) independently for left and right leads, for impedances ranging from 0.2 to 10 kΩ, assuming a reference electrode impedance of



Right signal 5 μV larger than left



Right signal 0.5 μV larger than left

Fig. 7. Asymmetry score (difference of natural log scores) as a function of amplifier input impedance, and impedance at left (z_{left}) and right (z_{right}) leads. For both plots, reference electrode impedance is set to 1 kΩ. The computed voltage at both left and right leads was squared to produce power units, and the difference of the natural log transformed power values ($\ln(\mathrm{Right}) - \ln(\mathrm{Left})$) was plotted on the vertical axis. Top two panels depict the observed asymmetry score when the right lead's true signal was 5 μV larger than the left, and the lower two panels depict the asymmetry score when the right lead's true signal was 0.5 μV larger than the left. Note that the top two panels are on the same scale, and the bottom two panels share a different scale.

1 kΩ.[8] The resultant values for left and right leads were then used to compute the asymmetry score ($\ln(\mathrm{Right}) - \ln(\mathrm{Left})$) for all combinations of left and right impedances.

Two aspects of Fig. 7 warrant comment. First, the overall impact of impedance mismatch ranging from 0 to 10 kΩ between left and right leads is negligible, and apparent only in the sixth decimal place of the asymmetry score. Differences between hemispheres and between groups of subjects, on the other hand, are readily apparent in the first decimal place (cf. Henriques and Davidson, 1990, 1991; Reid et al., 1998). Differences in left and right lead

---

[8] Results remain essentially unchanged if reference impedance is higher than 1 kΩ, as the reference impedance appears similarly in the equation for calculating observed left lead and observed right lead voltage.

impedances are thus unlikely to spuriously create or mask veritable differences in left and right alpha power. Second, the impact of the mismatch is further attenuated as a function of the higher input impedance amplifier. Asymmetry scores vary by 7.7×10^{-6} with the 10 MΩ input impedance, and by 3.9×10^{-6} with the 20 MΩ input impedance, when true voltage differences are 0.5 μV. Variation would be even less with higher input impedances (such as the 200 MΩ input impedance of the Net Amps, Electrical Geodesics Inc., Eugene, OR).

3.4. Dealing with ocular and muscular artifacts

EEG recordings may contain not only brain electrical activity, but non-cerebral contributions to the observed signal, including artifactual contributions of the scalp muscles and potentials generated by eye movements and blinks (Gratton, 1998; Picton et al., 2000). Although careful screening and rejection of data segments contaminated by the artifacts is likely to remove many of the artifacts, it may be desirable to obtain an estimate of the extent to which such artifacts may be influencing the results of an investigation.

3.4.1. Electrooculographic (EOG) influences

The eyes, being ion-filled imperfect spheres, carry a positive charge at the relatively leptokurtotic cornea, and a negative charge at the relatively platykurtotic retina. Being mobile, these charged spheres create electrical fields that are observed as signal in the case of EOG recordings, or artifact in the case of EEG recordings. Moreover, the conductive eyelid acts as a variable resistor as it slides across the cornea, momentarily distributing the ocular potential across the scalp. Thus ocular movements and blinks can be observed in scalp-recorded EEG, with the magnitude of the EOG signal decreasing as a function of the distance from the eyes. Although a majority of the signal of ocular origin is in the delta and theta range (Gasser et al., 1985; Hagemann and Naumann, 2001), slower than the 8–13 Hz alpha range of interest in EEG asymmetry research, some activity in the alpha band will inevitably be present, some potentially of neural origin (cf. Iacono and Lykken, 1981). The concern that activity of ocular origin may contaminate scalp-recorded EEG has prompted investigators utilizing EEG asymmetry to often reject epochs containing blinks or other ocular artifact. Moreover, the concern that the EOG signal may contain alpha-band activity of neural origin has discouraged investigators from employing a simple correction procedure that subtracts a portion of the time-domain EOG signal from the time-domain EEG signal, for doing so might also subtract alpha activity of neural origin.

Hagemann and Naumann (2001) carefully examined the contribution of ocular signals to scalp-recorded EEG asymmetry scores. Reviewing the literature, Hagemann and Naumann (2001) suggested that ocular artifacts are not likely to artifactually create or mitigate alpha-band asymmetries from homologous scalp leads, because: (1) power in the alpha band that is observed in EOG recordings is predominantly neural in origin, thus making it unlikely that ocular movements and blinks will appreciably alter scalp-recorded alpha activity, and (2) eye-movements and blinks are propagated relatively symmetrically. The symmetric propagation of vertical eye movements and blinks is apparent in the raw signal, but even lateral eye movements will be reflected similarly in the power spectra of left and right sites due to such movements creating similar magnitude but phase reversed deflections; the FFT will produce similar power spectra but different phase spectra, with only the power spectra of interest in EEG asymmetry research.

Assessing the contribution empirically, Hagemann and Naumann (2001) found that alpha asymmetry scores $(\ln(\text{Right}) - \ln(\text{Left}))$ derived from 8 min of resting EEG were highly similar when computed with versus without epochs containing ocular artifacts. The correlations between asymmetry scores from a dataset that included all epochs free of non-ocular artifacts and the same dataset with ocular-contaminated epochs were greater than 0.82 for all regions except the frontal pole, which was substantially lower. Effects of ocular-contaminated epochs, however, were larger for single sites than for corresponding asymmetry scores, further supporting the notion that ocular artifacts propagate symmetrically across most of the scalp.

Hagemann and Naumann (2001) concluded that the control of ocular artifacts may thus be unnecessary for correlational analyses involving alpha asymmetry scores, but that analyses involving mean levels may be influenced by ocular artifacts. Although the data in support of their conclusion is relatively strong—as the correlations are high between asymmetry scores from data with versus without artifacts—it is worth noting two issues that were not considered or assessed fully, and that remain to be investigated empirically. First, no relationships between alpha asymmetry and a criterion variable (e.g., BAS scores, cf. Coan and Allen, 2003a; Harmon-Jones and Allen, 1997; Sutton and Davidson, 1997) were investigated. Although the high correlation between asymmetry scores obtained from data with and without ocular artifacts would suggest that each would demonstrate similar correlations to a criterion variable, it is possible that the variance that is not shared by the two sets of asymmetry scores is differentially related to the criterion variable. Because the correlation between the two sets of asymmetry scores is attenuated the most at frontal leads by the inclusion or exclusion of epochs with ocular artifacts, and because it is precisely these regions that are of greatest interest with respect to the criterion variable, the possibility is amplified that the two sets of scores may relate differentially to a criterion variable. The second issue to consider derives from this possibility. Despite the careful analysis of Hagemann and Naumann (2001), one cannot differentiate between two possibilities: (1) that the true brain activity is invariant across epochs with and without ocular artifacts, but the presence of the ocular activity influences the observed EEG recording, or (2) the asymmetry scores differ because the true EEG activity differs as a function of whether blinks or eye movements are occurring. Given that eye-blinks show predictable relationships to cognitive processing and attention (Stern et al., 1984), this latter possibility must be considered in earnest.

3.4.2. Facial electromyographic (EMG) activity

Scalp-recorded EEG alpha activity may artifactually reflect the contribution of EMG activity (Cacioppo et al., 1990; Friedman and Thayer, 1991). Although the vast majority of the power in the EMG signal is faster than the alpha band, EMG activity has broad frequency characteristics with some small proportion of activity evident in the alpha band. This problem is potentially exaggerated by the fact that facial EMG asymmetries—sometimes similar in direction to reported cortical EEG asymmetries—have been observed (Borod et al., 1997), although the consistent and robust finding to emerge from this literature is an asymmetry characterized by greater left side activity in facial expressions in general, across all specific emotions and elicitation procedures.

Friedman and Thayer (1991) examined the potential magnitude of the EMG contribution to EEG recordings with the use of a redundancy analysis, which can be used to account for overlap between cortically derived alpha power and alpha power due to facial muscle activation. In their analysis, facial EMG accounted for 7% of the variance in cortical EEG activity, while cortical EEG activity accounted for only 3% of the variance in facial EMG, suggesting that facial EMG is likely to be responsible for a small but potentially important portion of the variance in scalp-recorded EEG. This study did not, however, specifically address the extent to which asymmetries in facial EMG activity were contributing to asymmetries in scalp EEG.

Coan et al. (2001) assessed the influence of EMG on scalp recorded alpha during a directed facial action task using two strategies. The first, a strategy used also by Davidson and colleagues (cf. Davidson, 1988; Davidson et al., 2000b), involves assessing EMG frequencies at scalp sites of interest. This approach extracts EMG frequencies (70–80 Hz in Davidson et al., 2000b or 70–90 Hz in Coan et al., 2001) from the power spectrum at each site involved in the EEG analysis. Because Coan et al. (2001) were analyzing EEG asymmetry scores, EMG asymmetry scores ($\ln(\text{Right}) - \ln(\text{Left})$) were computed on this EMG frequency band—for all the same regions as were included in the EEG analysis. These EMG range asymmetries were used as *changing covariates* in a multivariate repeated measures analysis of covariance (MANCOVA), which assumes that the EMG covariate changes within groups with the dependent variable across levels of the independent variable—in this case the particular facial expression. This changing covariate approach then correlates the change in the covariate with the change in the dependent variable and subsequently analyzes the residual variance in a standard MANOVA. Using this strategy, Coan et al. (2001) found that statistically adjusting for the EMG variance in this way did not change any of the significant relationships between facial pose and EEG asymmetry.

The second strategy used by Coan et al. (2001) involved an examination of alpha frequencies derived from bipolar EMG leads. This analysis was motivated by noting that the previous method assumes that all frequencies of the EMG are equally likely to show asymmetry effects that differ by the manipulation. But it is conceivable that EMG activity in the 70–90 Hz band may relate differently to a criterion variable than EMG activity in the 8–13 Hz band, the band of particular interest. Thus whereas the first analysis strategy examined EMG frequencies in EEG leads, the second examined alpha frequencies in EMG leads. This second approach derived power spectra from bipolar EMG activity in the frontalis and the temporalis muscle regions. Alpha power asymmetry scores derived from these EMG leads were thus included as changing covariates the analyses. Because unlike the first approach, where each region had its own covariate, this approach produced solely a frontalis alpha asymmetry score and a temporalis asymmetry score, one way repeated measures MANCOVAs were conducted separately for each region (e.g., one for F4–F3, one for F8–F7, etc.) since each region could not have its own changing covariate. Analyzing each region separately, thus increasing the number of analyses conducted, actually provided a more stringent test of whether the relationships between the manipulation and EEG asymmetry were influenced by myogenic contributions, because the chances increased of finding that the covariates rendered a previously significant effect nonsignificant. Statistically controlling for the EMG variance in this way, however, did not change any of the significant relationships between facial pose and EEG asymmetry.

4. Interpretive issues

4.1. Alpha and activity

A guiding assumption underlying the interpretation of findings involving frontal EEG alpha asymmetry is that greater alpha power is indicative of less cortical activity in broad underlying regions (cf. Davidson, 1988). Although there is good evidence to support this assumption, one might question whether this relationship is ubiquitous.

It has been well documented that sensory input shows modality-specific blocking of alpha activity at cortical regions involved in processing such input. With visual stimuli, blocking of alpha over the occipital region occurs about 0.3 s after the presentation of the visual stimulus (Berger, 1932; Jasper and Cruickshank, 1937; Knott, 1938), but this latency has been found to vary with intensity and duration of the stimulus (Cruickshank, 1937; Durup and Fessard, 1936a,b) and to diminish somewhat with a motor response related to the stimulus (Knott, 1938, 1939; Travis et al., 1937). Recovery time from blocking is generally about 1 s, but it too varies with stimulus intensity and duration (Cruickshank, 1937; Jasper and Cruickshank, 1937; Motokawa and Tosiada, 1941).

Similar but less dramatic effects are observed with other sensory modalities. Auditory stimuli, for example, block occipital alpha less effectively than visual stimuli and with a somewhat longer latency (Berger, 1930; Gibbs et al., 1935; Travis et al., 1937). Other sensory stimuli, such as tactile, cutaneous, pain (Berger, 1931, 1932; Jasper and Cruickshank, 1937; Livanov, 1940; Travis and Barber, 1938) and gustatory (Kitamura, 1939) have been found to block alpha, at least in their respective cortical areas.

Thus, sensory stimulation that should require active cortical processing leads to modality-specific alpha blocking, a principle that might lead to the inference that diminished alpha recorded over any cortical region signifies greater cortical activity. A test of this hypothesis in regions other than primary sensory regions is hampered by the lack of clearly defined stimuli to specifically engage those cortical regions in active processing, although several studies have provided data quite consistent with the notion that greater alpha power is indicative of less cortical activity in the underlying regions thought to subserve task performance (e.g., Davidson et al., 1990).

A consideration of the genesis of the alpha rhythm might prove illuminative for the assumption that diminished alpha recorded over any cortical region signifies greater cortical activity. A series of studies by Andersen and colleagues (Andersen et al., 1967a,b) suggest that thalamic rhythmicity drives cortical ensembles, the latter which comprise a large portion of scalp-recorded EEG activity. Andersen et al. (1967b) examined spindles in animals anesthetized with barbiturates, making the inference that such spontaneous rhythmic spindle activity is homologous with the human alpha rhythm. Several findings highlight the basis of their conclusion that thalamic rhythmicity drives cortical rhythmicity, including: (1) destruction or cooling of cortical regions leaves thalamic spindle activity unchanged (Andersen et al., 1967b); (2) damage or removal of the thalamus abolished cortical spindle activity (Andersen et al., 1967b); (3) unilateral destruction of thalamic tissue resulted in the disappearance of ipsilateral cortical spindle activity (Andersen et al., 1967b); (4) synchronous cortical spindles were not observed in relatively closely spaced cortical regions (those separated by 2 mm or more), but were observed over a much larger distance between a group of thalamic cells and the cortical area to which they projected (Andersen et al., 1967a). Thus Andersen and colleagues concluded that spontaneous cortical rhythmicity was "generated exclusively by thalamic neurons" (Andersen et al., 1967b, p. 258). Although cortical systems provide inputs to the thalamus that can disrupt the rhythmicity, the thalamus, and particularly the reticularis nucleus (Steriade et al., 1985), appears to be responsible for synchronizing cortical EEG activity.

Recent confirmation of a relationship between thalamic activity and scalp-recorded alpha activity in humans derives from a positron emission tomography (PET) study (Larson et al., 1998). During approximately 30 min, EEG and [^{18}F]-2-fluoro-2-deoxy-D-glucose (FDG) PET recordings were obtained. Global alpha (8–13 Hz) power was then correlated with glucose metabolism, and cortical alpha power was strongly and inversely related to glucose metabolism in the thalamus (Larson et al., 1998). This finding is consistent with the notion that thalamic activity in response to sensory or cortical input will disrupt alpha rhythmicity.

Thus scalp recorded EEG alpha activity—in a very coarse sense both spatially and temporally—is inversely related to thalamic activity. Global alpha power across electrodes and across 30 min relates to thalamic metabolism. Ultimately, however, investigators would wish to know whether EEG alpha at a given scalp lead is related to cortical activity in the tissue beneath that lead, a question addressed by Cook et al. (1998). Using $H_2{}^{15}O$ PET imaging allowed them to examine activity in 2 min segments, with a total of eight such segments per subject. EEG power was calculated for 4 Hz wide bins, starting at 0 Hz and extending to 40 Hz, at 1 Hz intervals (e.g., 0–4, 1–5, 2–6, etc.). Cerebral perfusion under each electrode was calculated, and then correlated with each of the EEG spectral bins, resulting in a plot of correlations between EEG power and cerebral perfusion as a function of frequency. Frequency range played a major role in the relationship of EEG power with perfusion, such that lower frequencies (those bins with a center frequency below 8 Hz) had a positive relationship to perfusion, middle frequencies (bins with center frequencies from 8 to 12 Hz) had a negative relationship, and upper frequency ranges (center frequency varied depending on specific operationalization) had a positive relationship (Fig. 8). Apparent from the figure is that relative power shows a closer correspondence to underlying cortical activity than does absolute power, which may reflect that the latter is confounded by variations in scalp thickness much more than the former. Additionally, although alpha power is inversely related to underlying cortical activity *no matter which of the montages was used*, the relationship of beta power (13–30 Hz) to underlying activity varied substantially as a function of recording montage, exhibiting either a positive or negative relationship depending on the particular recording montage used.

Thus there is reasonable support for the assumption that greater alpha at a scalp lead reflects less cortical activity in a broad region(s) contributing to electrical activity recorded at that lead. Recent data of Cook et al. (1998) suggest, however, that a tighter correspondence between cortical activity and scalp-recorded EEG is possible with a reattribution technique these authors have called *cordance*, although the correspondence in the alpha band is not vastly improved using this cordance measure (see the lower panel of Fig. 8, taken from Cook et al., 1998). Whether asymmetry in reattributed power demonstrates relationships with emotion and individual differences, however, remains an empirical question, but one worth investigating given the tighter coupling of EEG to brain function that appears possible using this technique (see also Leuchter et al., 1994, 1999, 2002).

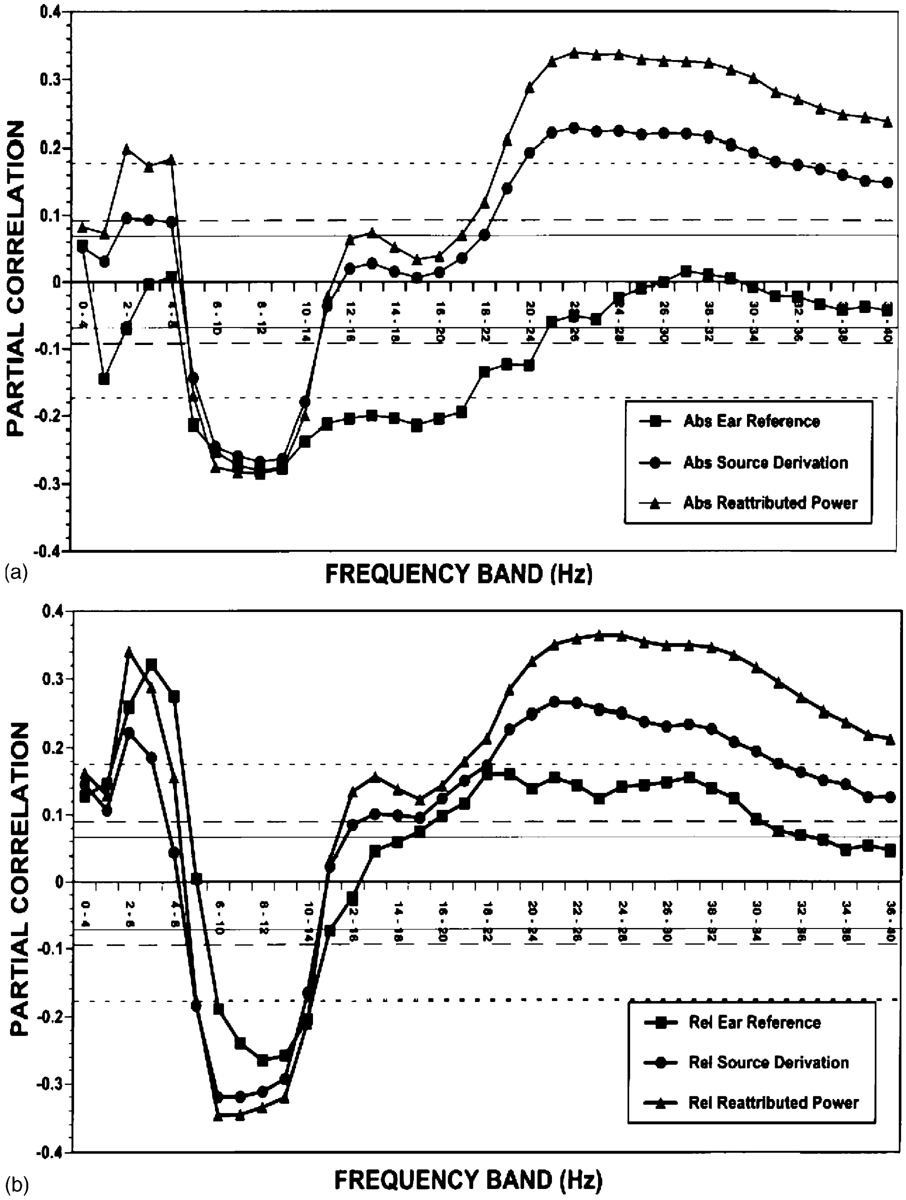


Fig. 8. Correlations between EEG power and PET perfusion values at tissue under the EEG scalp lead. Top panel depicts absolute power, and lower panel depicts relative power (i.e., power in the 4-Hz wide bin divided by total power across all spectral frequencies). “Ear reference” is a computer linked ears reference, “source derivation” is that described by Hjorth (1975) that weights immediate neighboring electrodes in the time domain prior to frequency-domain transformation, and “reattributed power” is a weighting of power derived from bipolar channels of nearest neighbors (Cook et al., 1998). Statistical significance is indicated by horizontal lines representing the magnitude at which a correlation coefficient attains significance: solid line for $P = 0.05$; large dashed line for $P = 0.01$; fine dashed line for $P = 0.001$. From Cook et al. (1998), reprinted with permission from Elsevier.

Worth noting from Cook et al. (1998), and as apparent in Fig. 8, is that the inverse relationship between cortical perfusion and EEG power is relatively constant throughout the entire alpha band in these adult subjects. Thus although the 8–13 Hz definition of alpha is somewhat arbitrary, the correspondence of activity in this range (and perhaps a bit of the 6–8 Hz portion of the upper Theta band) to underlying cortical activity at rest is relatively uniform in adult subjects. This observation does not suggest a strong functional distinction between smaller bandwidth divisions, such as upper or lower alpha, at rest in a psychiatrically and neurologically healthy population (although some investigators have made an argument for the utility of subdividing the alpha band for specific tasks and applications, e.g., Klimesch et al., 1997).

A final issue with respect to interpreting alpha and cortical activity concerns how to conceptualize resting EEG data. Such resting data necessarily summarizes activity across several minutes, which will collapse across many variations in brain and psychological state during the recording period. Although investigators refer to such periods as resting periods, one might alternatively think of EEG activity during these periods as task-related, with individual differences in how subjects approach this “task” of resting for several minutes underlying the observed individual differences in EEG activity. The resting state is relatively uncontrolled, allowing for individual differences in mentation (broadly construed) during the resting period to influence the measure (cf. Schwartz et al., 1976 for a similar phenomenon when depressed and nondepressed subjects pondered a typical day while facial EMG was recorded).

4.2. Robustness or capitalization on chance: the impact of reference schemes, specific sites, and other variations

Reviewing the literature (see Tables 1–4 in Coan and Allen, 2004, this issue), one is impressed by the fact that significant relationships involving frontal EEG asymmetry: (1) derive from data analyzed under a variety of reference schemes, with different studies using different reference montages; (2) appear to involve different specific frontal regions in different studies; and (3) sometimes involve different frequency cutpoints for operationalizing alpha band activity. It is premature to know how best to interpret such a pattern of findings, but these observations suggest at least three non-mutually exclusive possibilities.

First, for those who like to see glasses as half-empty, this could reflect that this research field suffers—as all do to some degree—from significant inflation of the likelihood of Type I error, with alpha inflation resulting from the poor control for multiple comparisons, compounded by the many permutations of variables possible when recording from multiple regions under multiple reference montages, with the possibility of different operationalizations of alpha-band activity. In the absence of strong theory to suggest that investigators should find effects at one specific frontal region and not another, or under one reference scheme and not another, investigators wisely examine multiple sites and reference schemes, but the incumbent risk of this strategy is that the field as a whole may be inadequately protected against reporting spurious findings.

Second, for those who like to see glasses as half-full, the fact that relationships with frontal EEG asymmetry appear with data from different reference schemes at different times, and at different regions at different times, suggests that these observed variables are imperfect manifest representations of what should be considered a latent variable, i.e., functional frontal brain asymmetry. This argument would imply that only by recording multiple regions and under multiple reference schemes can one adequately assess an individual's true score on the latent variable. By aggregating over multiple measures, one gains power and reliability, thus enhancing the ability to find relationships between frontal brain asymmetry and criterion variables. This argument further implies that the impact of these factors, such as reference scheme and specific frontal region, should be tested explicitly in a statistical model, such as those detailed previously in the section on hierarchical general linear models. In the absence of such a model to protect from isolated chance findings appearing significant, the field may indeed suffer from inflation of Type I statistical error.

The latent trait argument would of course predict that the manifest variables should all show some modest correlations with the latent trait, and likely with one another. Although this is true of data recorded under the average reference and the computer averaged mastoids references, it is less true of data recorded using the Cz reference (Hagemann et al., 1998; Reid et al., 1998), which shows much lower correspondence with data using the other reference montages. It is worth noting, moreover, that by far the most common reference scheme used in frontal EEG asymmetry research is the Cz reference (see Coan and Allen, 2004, this issue), and that most studies do not employ multiple reference montages in their analysis of the data.

The third possibility is that there exists some systematic relationship between measured or unmeasured variables and asymmetry at specific sites or under specific reference schemes. This line of reasoning suggests the differential engagement of various frontal systems as a function of particular task demands, as a function of factors in the experimental environment, and as a function of various individual difference variables under study.

At present, it is difficult to assess the likelihood that such effects exist, as most studies do not assess a range of variables and attempt to relate them to asymmetry at specific regions. A notable exception comes from a recent study of Miller and Tomarken (2001), in which manipulations of expected reward or punishment produced changes in mid-frontal EEG asymmetry (that varied by sex), and manipulation of the required response produced changes in central asymmetry. These results suggest that there may indeed be task variables that will impact the specific region in which EEG alpha asymmetry effects are likely to appear. Nonetheless, there are many other non-task variables that may impact the regional specificity of EEG asymmetry effects, but such variables may not be known or assessed, thus making it impossible to discern whether there exists a systematic cause underlying the appearance of EEG asymmetry effects at some sites in some studies, and other sites in other studies.

4.3. Consistency and variability

Estimates of EEG alpha asymmetry are averages that summarize patterns of brain activity across several minutes, either contiguous time segments in the case of resting EEG asymmetry, or collapsed across numerous discrete but separated time segments in the case of state-manipulated EEG asymmetry. Although it has been adequately demonstrated that such estimates possess excellent internal consistency reliability (e.g., Allen et al., 2001; Coan and Allen, 2003b; Coan et al., 2001; Reid et al., 1998; Tomarken et al., 1992), these estimates of internal consistency are derived from several segments of data, each of which is an average reflecting the pattern of brain activity across many seconds of recorded data. There is a sense in which frontal EEG alpha asymmetry derived from these segments of data ignores variability on a finer temporal scale. Whether such variability will prove meaningful is ultimately an empirical question, but the utility of explicitly examining such variability appears promising. For example, Minnix and Kline (2004) examined the variance estimate associated with the average FFT from a resting assessment. Subjects who show more second-to-second variability in frontal EEG alpha asymmetry will have higher variance estimate across the entire recording epoch. Minnix and Kline (2004) found that increased variability of this sort was related to higher neuroticism. Thus a trait characterized by greater emotional lability was found to be associated with more lability in frontal EEG asymmetry as well.

Another way of assessing the stability of frontal EEG asymmetry was first reported by Baehr et al. (1998), who computed the percentage of time that right alpha is greater than left alpha at homologous leads. Baehr et al. (1998) found that the percent–time measure better discriminated psychometrically-defined depressed subjects from nondepressed subjects than the traditional asymmetry measure that averaged across the recording period. Allen et al. (2001) used the percent–time measure as well, finding that it produced comparable findings to the traditional asymmetry score. Thus although it may be premature to suggest that this metric has distinct advantages, the extant data suggest its promise and moreover suggest that it would not result in the elimination of significant findings with the traditional asymmetry score.

4.4. *Keeping straight the states and the traits*

Substantial data support the contention that frontal EEG asymmetry can serve as a relatively stable individual difference variable, yet also show predictable state-related fluctuations (see Coan and Allen, 2004, this issue). Evidence in support of the trait-like quality of frontal EEG asymmetry derives from studies specifically examining stability over time. Tomarken et al. (1992) assessed the psychometric properties of trait-like frontal EEG asymmetries, finding that frontal EEG asymmetry demonstrated acceptable test–retest stability (intra-class correlations ranging from 0.69 to 0.84 across 3 weeks). Similarly, Jones et al. (1997) found that frontal EEG asymmetry recorded at 3 months of age was highly correlated with asymmetry at 3 years ($r = 0.66$, $P < 0.01$). Similar figures come from Hagemann et al. (2002), who found that across four different measurement occasions, 60% of the variance in EEG asymmetry measures was due to individual differences in a temporally stable latent trait.

To enhance the ability to identify trait-related variance, some studies have specifically examined subjects who show the greatest cross-session consistency (e.g., Wheeler et al., 1993), reasoning that the strongest relationships to other traits should be shown by those who are consistent on the measure of trait EEG asymmetry (cf. Bem and Allen, 1974). Others have averaged data across multiple sessions to mitigate occasion-specific fluctuations and presumably derive a better estimate of the trait-related variance in EEG asymmetry (e.g., Sutton and Davidson, 1997).

When attempting to account for the nonstable variance in frontal EEG asymmetry, three sources must be considered: reliable changes from one session to the next, reliable and systematic changes within session, and unreliability of measurement. Because frontal EEG asymmetry demonstrates high internal consistency reliability at any given assessment session (Cronbach's alphas typically above 0.95; Reid et al., 1998), it is unlikely that attenuated test–retest reliability of frontal EEG asymmetry is due to random measurement error. Rather, a majority of the variance in EEG asymmetry can be accounted for reliable and systematic sources of variation (Coan and Allen, 2003b) due to:

(1) *Stable trait consistency* across multiple assessments, which is presumably indicative of temperamental style and a tendency to respond in a characteristic way when confronted with emotionally evocative situations. An individual's trait level of frontal EEG asymmetry represents a quality of that individual—a quality that the individual brings to a variety of situations and contexts. This trait level is necessarily estimated, by averaging across multiple occasions of measurement (e.g., Sutton and Davidson, 1997; Wheeler et al., 1993), by modeling it as a latent trait (e.g., Hagemann et al., 2002), or by accounting for it within the context of a generalizability analysis (as described below).
(2) *Occasion-specific* variance refers to reliable variations in frontal asymmetry that characterize the variation in resting EEG assessments across multiple sessions of measurement. Such variation may reflect systematic but unmeasured sources such as current mood, recent life events and/or factors in the testing situation.
(3) *State-specific* variance refers to changes within a single assessment that characterize the difference between two experimental conditions or between baseline resting levels and an experimental condition. State-specific changes as conceptualized here are proximal effects in response to specific experimental manipulations. Such manipulations should be reversible and of relatively short duration.

These state-related fluctuations stand in contrast to the occasion-specific fluctuations, which are assumed to characterize the individual throughout the measurement occasion, reflecting the high internal consistency reliability estimates such measurement occasions typically show. Occasion variance is hypothesized to reflect the effects of time- or context-limited individual difference variables (e.g., mood on the day of assessment, recent or imminent life events, daily hassles) or alternatively the interaction of the individual with the experimental milieu in a manner that varies from assessment to assessment (e.g., effects of experimental milieu or procedures, Blackhart et al. (2002), or experimenter effects, Kline et al. (2002)). Such effects would not be the result of purposeful state-related experimental manipulations, but would rather represent an interaction of the subject with other experimentally uncontrolled stimuli.

Most studies of trait frontal asymmetry are not designed to allow for the separation of trait variance and occasion variance, as most studies entail only a single occasion of measurement of resting frontal asymmetry. If occasion-specific fluctuations were not sizable, then a single assessment of trait levels would prove sufficient. Recent evidence (Hagemann et al., 2002), however, suggests that reliable occasion-specific fluctuations account for approximately 40% of reliable variance in resting frontal asymmetry, while the consistency across multiple sessions, presumably reflecting a stable trait, accounts for approximately 60%. Further, there may exist individual differences in the magnitude of occasion-specific fluctuations. For example, Wheeler et al. (1993) selected a subset of 26 from among 81 women (i.e., 32% of the sample) who were classified as possessing stable asymmetry, meaning that 68% of the sample was classified as having unstable asymmetry.

4.4.1. Methods for assessing state, occasion, and trait variance

To reliably delineate sources of variance in frontal EEG asymmetry, an analytic strategy for decomposing variance components is needed. One such promising general strategy is generalizability or “G” theory.

Generalizability theory (Cronbach et al., 1972), or “G-theory,” was developed for the purpose of identifying the *generalizability* and *dependability* of different independent variables thought to contribute to a measure’s score (e.g., Di Nocera et al., 2001). The *generalizability* of a measure is analogous to more conventional estimates of reliability, such as the intraclass correlation, while *dependability* of a measure refers to a measure’s reliability across contexts. Generalizability and dependability estimates may be obtained for each independent variable thought to contribute variance to a measure. In practice, an independent variable with high dependability is one that contributes variance that is relatively independent of other independent variables affecting the measure of interest. For example, an estimate of the dependability of trait variance in frontal EEG asymmetry would allow one to assess how independent and stable trait variance is from state manipulations and measurement occasions. In addition to estimates of generalizability and dependability, actual variance components may be estimated for each independent variable hypothesized to contribute to an individual’s score at any one time, including variance components attributable to the interaction of independent variables. G-theory is based fundamentally on an ANOVA model in the estimation of variance components. A critical difference between G-theory analyses and classical ANOVA models is that G-theory requires the computation of expected, as opposed to observed variance components. Expected variance components are estimated by using specific algorithms employed in very few statistical packages (e.g., SAS PROC VARCOMP).

As applied to questions of state, occasion and trait variance in frontal EEG asymmetry, such a model might be defined as follows (cf. Di Nocera et al., 2001):

$$\sigma_{\mathrm{y}}^2 = \sigma_{\mathrm{t}}^2 + \sigma_{\mathrm{o}}^2 + \sigma_{\mathrm{s}}^2 + \sigma_{\mathrm{to}}^2 + \sigma_{\mathrm{ts}}^2 + \sigma_{\mathrm{os}}^2 + \sigma_{\mathrm{tos}}^2$$

where σ_{y}^2 is the total variance for a given variable, in this case frontal EEG asymmetry, across all occasions and manipulations, σ_{t}^2 is the variance in frontal EEG asymmetry attributable to individuals (here considered trait variance), σ_{o}^2 is the variance in frontal EEG asymmetry attributable to measurement occasion, σ_{s}^2 is the variance in frontal EEG asymmetry attributable to experimentally manipulated states, σ_{to}^2 is the variance in frontal EEG asymmetry attributable to the interaction of trait and occasion variance, σ_{ts}^2 is the variance in frontal EEG asymmetry attributable to the interaction of trait and state variance, σ_{os}^2 is the variance in frontal EEG asymmetry attributable to the interaction of occasion and state variance, and σ_{tos}^2 is the variance in frontal EEG asymmetry attributable to the interaction of trait, occasion and state variance (confounded with error of measurement).

G-theory thus provides variance component estimates (percent of variance accounted for by each component), coefficients of generalizability (ρ^2), as well as coefficients for dependability (ϕ or *phi*), for each independent variable of interest (in this case, trait, occasion and state components). If, for example, trait variance in frontal EEG asymmetry shows high dependability in addition to high generalizability, such a finding would bolster the likelihood that it would prove useful as a liability indicator for risk for psychopathology, or index a trait-like affective style, as this finding would indicate that trait frontal EEG asymmetry can be assessed reliably, and that trait frontal EEG asymmetry is independent of state manipulations and occasion-related fluctuations.

Further, to date no researcher has examined the stability of state manipulations in frontal EEG asymmetry, over time or otherwise. G-theory provides useful estimates of reliability that mirror and extend approaches designed to understand intraindividual dynamics, such as advocated by Mischel and colleagues (e.g., Shoda and Mischel, 1996; Shoda et al., 1994). Using Shoda and Michel's approach, such patterns of behavior are represented as *if* ... *then* ... probabilities that vary from individual to individual and that presumably reflect an individual's underlying personality type (Mendoza-Denton et al., 2001). Effects of this type are also easily accommodated by a generalizability analysis, as they would be reflected in σ^2_{ts}, the trait by state interaction term. Thus, in addition to assessing the dependability of trait variance across states and occasions, the dependability of state manipulations across individuals is estimable, as is the dependability of occasion variance across individuals and states.

Although no investigators have applied the model specified above to data collected across multiple occasions of measurement, Coan and Allen (2004, this issue) did assess the extent to which state changes in frontal EEG asymmetry were reliably elicited across subjects, and the extent to which trait levels of frontal EEG asymmetry were preserved across state manipulations. The results indicated that trait-specific variance, state-specific variance, as well as variance attributable to their interaction, each accounted for approximately 10% of the total explained variance in frontal EEG asymmetry. Trait stability as measured by the *g* coefficient (intraclass correlation) was estimated to be moderately high (0.47), whereas state stability was extremely high (*g* coefficient $= 0.92$). Although these results identify stability and sizable contributions of both trait and state frontal EEG asymmetry, trait variance as estimated from this single measurement occasion will necessarily include both stable trait and occasion specific influences. Indeed, while the state variance in response to the manipulation was highly stable, the trait variance was only moderately so. This may be due to the influence of unmeasured but relevant occasion specific factors, which future efforts might profitably explore.

A final note with respect to the G-theory approach concerns its flexibility to assess the impact of a variety of other factors, such as the effect of reference scheme and specific frontal region. By including terms to account for variance due to particular reference scheme, or to the particular frontal region (e.g., F4–F3 versus F8–F7 versus FTC2–FTC1), the magnitude of these sources of variance and their interactions with trait, occasion, and state variance can be assessed. Similarly, estimates of the stability of the effects across these factors can be quantitatively assessed.

5. Synopsis

Research on frontal EEG asymmetry and emotion now represents a substantial body of literature. There are numerous methodological issues to which the field may have paid insufficient attention, while at the same time paying potentially too much attention to other factors. The field may have been too concerned with recording at least 8 min of data to obtain reliable estimates of asymmetry, overly concerned about the impact of blink artifacts, and overly concerned with closely matching impedances at homologous leads. By contrast, too little concern has generally been given to assessing the impact of reference scheme, disentangling left from right hemisphere effects using appropriate statistical models, and discerning whether specific regions are differentially involved in various tasks or as a function of individual differences. It remains to be determined whether the impact of myogenic activity substantially influences findings involving EEG alpha asymmetry.

As one reviews the frontal EEG asymmetry and emotion literature (Coan and Allen, 2003a; 2004, this issue), it is apparent that many different data analytic approaches have been used, resulting in a collection of findings that converge despite rather dramatic differences in: (1) the conditions under which data were recorded; (2) the manner in which data were reduced; and (3) the manner in which data were subsequently analyzed. The optimist will see this as a testament to the robustness of the underlying systems reflected in frontal EEG asymmetry, and the curmudgeon will see this as representing considerable literature-wide alpha slippage due to the many permutations of data reduction and analysis. A conservative intermediate interpretation is that the larger enterprise of interpreting the data and theory building will benefit from a more solid empirical foundation, one that will require that careful attention be given to EEG data recording and analysis. The issues highlighted here may best be regarded as fundamentals that may inform future efforts, to assist in the creation of a more methodologically consistent and precise data base. Only with such a foundation can researchers then explore the underlying functional, anatomical and neurochemical systems that may be tapped by frontal EEG asymmetry.

Acknowledgements

This research was supported, in part, by a grant from the National Alliance for Research on Schizophrenia and Depression (NARSAD) and an Exploratory/Development Grant from the National Institutes of Health (1 R21 RR09492) to J.J.B.A., and by a Graduate Research Fellowship from the National Science Foundation to J.A.C.

References

Aiken, S.L., West, S.G., 1991. Multiple Regression: Testing and Interpreting Interactions. Sage Publications, Thousand Oaks, CA.

Allen, J.J.B., Harmon-Jones, E., Cavender, J.H., 2001. Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG. Psychophysiology 38, 685–693.

Allen, J.J.B., Urry, H.L., Hitt, S.K., Coan, J.A., 2004. The stability of resting frontal electroencephalographic asymmetry in depression. Psychophysiology 41, 269–280.

Andersen, P., Andersson, S.A., Lomo, T., 1967a. Nature of thalamo-cortical relations during spontaneous barbiturate spindle activity. Journal of Physiology 192, 283–307.

Andersen, P., Andersson, S.A., Lomo, T., 1967b. Some factors involved in the thalamic control of spontaneous barbiturate spindles. Journal of Physiology 192, 257–281.

Baehr, E., Rosenfeld, J.P., Baehr, R., Earnest, C., 1998. Comparison of two EEG asymmetry indices in depressed patients vs. normal controls. International Journal of Psychophysiology 31 (1), 89–92.

Bem, D.J., Allen, A., 1974. On predicting some of the people some of the time: the search for cross-situational consistencies in behavior. Psychological Review 81 (6), 506–520.

Berger, H., 1930. Uber das elektrenkephalogramm des Menschen II. Journal fur Psychologie und Neurologi 40, 160–179.

Berger, H., 1931. Uber das elektrenkephalogramm des Menschen III. Archiv fur Psychiatrie und Nervenkrankheiten 94, 16–60.

Berger, H., 1932. Uber das elektrenkephalogramm des Menschen IV. Archiv fur Psychiatrie und Nervenkrankheiten 97, 6–26.

Blackhart, G.C., Kline, J.P., Donohue, K.F., LaRowe, S.D., Joiner, T.E., 2002. Affective responses to EEG preparation and their link to resting anterior EEG symmetry. Personality & Individual Differences 32 (1), 167–174.

Borod, J.C., Haywood, C.S., Koff, E., 1997. Neuropsychological aspects of facial asymmetry during emotional expression: a review of the normal adult literature. Neuropsychology Review 7, 41–60.

Cacioppo, J.T., Tassinary, L.G., Fridlund, A.J., 1990. The skeletomotor system. In: Cacioppo, J.T., Tassinary, L.G. (Eds.), Principles of Psychophysiology. Cambridge University Press, New York, NY.

Campbell, D.T., Fiske, D.W., 1959. Convergent and discriminant validation by the multitrait–multimethod matrix. Psychological Bulletin 56, 81–105.

Coan, J.A., Allen, J.J.B., 2003a. Frontal EEG asymmetry and the behavioral activation and inhibition systems. Psychophysiology 40, 106–114.

Coan, J.A., Allen, J.J.B., 2003b. The state and trait nature of frontal EEG asymmetry in emotion. In: Hugdahl, K., Davidson, R.J. (Eds.), The Asymmetrical Brain, second ed. MIT Press, Cambridge, MA, pp. 565–615.

Coan, J.A., Allen, J.J.B., 2004. Frontal EEG asymmetry as a moderator and mediator of emotion. Biological Psychology 67, 7–49.

Coan, J.A., Allen, J.J.B., Harmon-Jones, E., 2001. Voluntary facial expression and hemispheric asymmetry over the frontal cortex. Psychophysiology 38, 912–925.

Cook, I.A., O'Hara, R., Uijtdehaage, S.H., Mandelkern, M., Leuchter, A.F., 1998. Assessing the accuracy of topographic EEG mapping for determining local brain function. Electroencephalography & Clinical Neurophysiology 107 (6), 408–414.

Cronbach, L.J., Gleser, G.C., Nanda, H., Rajaratnam, N., 1972. The Dependability of Behavioral Measurements: Theory of Generalizability of Scores and Profiles. Wiley, New York.

Cruickshank, R.M., 1937. Human occipital brain potentials as affected by intensity-duration variables of visual stimulation. Journal of Experimental Psychology 21, 625–641.

Davidson, R.J., 1988. EEG measures of cerebral asymmetry: conceptual and methodological issues. International Journal of Neuroscience 39 (1–2), 71–89.

Davidson, R.J., 1992. Anterior cerebral asymmetry and the nature of emotion. Brain & Cognition 20 (1), 125–151.

Davidson, R.J., 1998. Affective style and affective disorders: perspectives from affective neuroscience. Cognition & Emotion 12 (3), 307–330.

Davidson, R.J., 22 and 24 September 2002. Personal communication.

Davidson, R.J., Chapman, J.P., Chapman, L.J., Henriques, J.B., 1990. Asymmetrical brain electrical activity discriminates between psychometrically-matched verbal and spatial cognitive tasks. Psychophysiology 27 (5), 528–543.

Davidson, R.J., Jackson, D.C., Larson, C.L., 2000a. Human electroencephalography. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (Eds.), Handbook of Psychophysiology, second ed. Cambridge University Press, Cambridge, UK, pp. 27–52.

Davidson, R.J., Marshall, J.R., Tomarken, A.J., Henriques, J.B., 2000b. While a phobic waits: regional brain electrical and autonomic activity in social phobics during anticipation of public speaking. Biological Psychiatry 47 (2), 85–95.

Di Nocera, F., Ferlazzo, F., Borghi, V., 2001. G theory and the reliability of psychophysiological measures: a tutorial. Psychophysiology 38, 796–806.

Durup, G., Fessard, A., 1936a. L'electrencephalogramme de l'homme. Donnees quantitatives sure l'arret provoque par des stimuli visuels ou auditifs. Comptes Rendus des Seances de la Societe de Biologie et de ses Filiales 122, 756–758.

Durup, G., Fessard, A., 1936b. L'electrencephalogramme de l'homme observations psychophysiologiques relative a l'action des stimuli visuels et auditifs. Annee Psychologie 36, 1–32.

Eshel, Y., Witman, S., Rosenfeld, M., Abboud, S., 1995. Correlation between skull thickness asymmetry and scalp potential estimated by a numerical model of the head. IEEE Transactions on Biomedical Engineering 42 (3), 242–249.

Ferree, T.C., 27 November 2002. Personal communication.

Ferree, T.C., Luu, P., Russell, G.S., Tucker, D.M., 2001. Scalp electrode impedance, infection risk and EEG data quality. Clinical Neurophysiology 112 (3), 536–544.

Fox, N.A., Davidson, R.J., 1987. Electroencephalogram asymmetry in response to the approach of a stranger and maternal separation in 10 month old infants. Developmental Psychology 23, 233–240.

Friedman, B.H., Thayer, J.F., 1991. Facial muscle activity and EEG recordings—redundancy analysis. Electroencephalography and Clinical Neurophysiology 79, 358–360.

Gasser, T., Sroka, L., Mocks, J., 1985. The transfer of EOG activity into the EEG for eyes open and closed. Electroencephalography & Clinical Neurophysiology 61 (2), 181–193.

Gibbs, F.A., Davis, H., Lennox, W.G., 1935. The electro-encephalogram in epilepsy and in conditions of impaired consciousness. Archives of Neurological Psychiatry 34, 1133–1148.

Glaser, E.M., Ruchkin, D.S., 1976. Principles of Neurobiological Signal Analysis. Academic Press, New York.

Gratton, G., 1998. Dealing with artifacts: the EOG contamination of the event-related brain potential. Behavior Research Methods, Instruments, & Computers 30 (1), 44–53.

Gratton, G., 2000. Biosignal processing. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (Eds.), Handbook of Psychophysiology, second ed. Cambridge University Press, New York, pp. 900–923.

Hagemann, D., Naumann, E., 2001. The effects of ocular artifacts on (lateralized) broadband power in the EEG. Clinical Neurophysiology 112 (2), 215–231.

Hagemann, D., Naumann, E., Becker, G., Maier, S., Bartussek, D., 1998. Frontal brain asymmetry and affective style: a conceptual replication. Psychophysiology 35 (4), 372–388.

Hagemann, D., Naumann, E., Thayer, J.F., 2001. The quest for the EEG reference revisited: a glance from brain asymmetry research. Psychophysiology 38, 847–857.

Hagemann, D., Naumann, E., Thayer, J.F., Bartussek, D., 2002. Does resting electroencephalograph asymmetry reflect a trait? An application of latent state–trait theory. Journal of Personality & Social Psychology 82 (4), 619–641.

Harmon-Jones, E., Allen, J.J.B., 1997. Behavioral activation sensitivity and resting frontal EEG asymmetry: covariation of putative indicators related to risk for mood disorders. Journal of Abnormal Psychology 106 (1), 159–163.

Harmon-Jones, E., Allen, J.J.B., 1998. Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. Journal of Personality & Social Psychology 74 (5), 1310–1316.

Henriques, J.B., Davidson, R.J., 1990. Regional brain electrical asymmetries discriminate between previously depressed and healthy control subjects. Journal of Abnormal Psychology 99 (1), 22–31.

Henriques, J.B., Davidson, R.J., 1991. Left frontal hypoactivation in depression. Journal of Abnormal Psychology 100 (4), 535–545.

Hjorth, B., 1975. An on-line transformation of EEG scalp potentials into orthogonal source derivations. Electroencephalography & Clinical Neurophysiology 39 (5), 526–530.

Iacono, W.G., Lykken, D.T., 1981. Two-year retest stability of eye tracking performance and a comparison of electro-oculographic and infrared recording techniques: evidence of EEG in the electro-oculogram. Psychophysiology 18 (1), 49–55.

Jasper, H.H., Cruickshank, R.M., 1937. Electro-encephalography: II. Visual stimulation and the after-image as affecting the occipital alpha rhythm. Journal of General Psychology 17, 29–48.

Jones, N.A., Field, T., Davalos, M., Pickens, J., 1997. EEG stability in infants/children of depressed mothers, Child Psychiatry. Child Psychiatry & Human Development 28 (2), 59–70.

Kitamura, K., 1939. Die elektrencephalographische untersuchung der gerschmacksempfindlichkeit. Tohoku Psychologica Folia 7, 13–32.

Klimesch, W., Doppelmayr, M., Pachinger, T., Ripper, B., 1997. Brain oscillations and human memory: EEG correlates in the upper alpha and theta band. Neuroscience Letters 238 (1–2), 9–12.

Kline, J.P., Blackhart, G.C., Joiner, T.E., 2002. Sex, lie scales, and electrode caps: an interpersonal context for defensiveness and anterior electroencephalographic asymmetry. Personality & Individual Differences 33 (3), 459–478.

Knott, J.R., 1938. Reduced latent time of blocking of the Berger rhythm to light stimuli. Proceedings of the Society for Experimental Biology and Medicine 38, 216–217.

Knott, J.R., 1939. Some effects of "mental set" upon the electrophysiological processes of the human cerebral cortex. Journal of Experimental Psychology 24, 384–405.

Larson, C.L., Davidson, R.J., Abercrombie, H.C., Ward, R.T., Schaefer, S.M., Jackson, D.C., 1998. Relations between PET-derived measures of thalamic glucose metabolism and EEG alpha power. Psychophysiology 35 (2), 162–169.

Leissner, P., Lindholm, L.E., Petersen, I., 1970. Alpha amplitude dependence on skull thickness as measured by ultrasound technique. Electroencephalography & Clinical Neurophysiology 29 (4), 392–399.

Leuchter, A.F., Cook, I.A., Lufkin, R.B., Dunkin, J., Newton, T.F., Cummings, J.L., 1994. Cordance: a new method for assessment of cerebral perfusion and metabolism using quantitative electroencephalography. Neuroimage 1 (3), 208–219.

Leuchter, A.F., Cook, I.A., Witte, E.A., Morgan, M., Abrams, M., 2002. Changes in brain function of depressed subjects during treatment with placebo. American Journal of Psychiatry 159 (1), 122–129.

Leuchter, A.F., Uijtdehaage, S.H., Cook, I.A., O'Hara, R., Mandelkern, M., 1999. Relationship between brain electrical activity and cortical perfusion in normal subjects. Psychiatry Research 90 (2), 125–140.

Livanov, M.N., 1940. Rhythmical stimuli and the interrelation between the areas of the cerebral cortex. Journal of Physiology 28, 172–194.

Lord, F.M., Novick, M.R., 1968. Statistical Theories of Mental Test Scores. Addison-Wesley, Reading, MA.

Mendoza-Denton, R., Ayduk, O., Mischel, W., Shoda, Y., Testa, A., 2001. Person × situation interactionism in self-encoding (I am . . . when . . .): implications for affect regulation and social information processing. Journal of Personality & Social Psychology 80 (4), 533–544.

Miller, A., Tomarken, A.J., 2001. Task-dependent changes in frontal brain asymmetry: effects of incentive cues outcome expectancies and motor responses. Psychophysiology 38, 500–511.

Minnix, J.A., Kline, J.P., 2004. Neuroticism predicts resting frontal EEG asymmetry variability. Personality and Individual Differences 36, 823–832.

Motokawa, K., Tosiada, M., 1941. Die elektrenkephalographische untersuchung uber den adaptationsmechanismus des zentralnervensystems. Japanese Journal of Medical, Science and Biophysics 7, 213–233.

Nunez, P., 1981. Electrical Fields of the Brain: The Neurophysics of EEG. Oxford University Press, New York.

Nyquist, H., 1928. Certain topics in telegraph transmission theory. Transactions of the American Institute of Electrical Engineers 47, 617–644.

Overall, J.E., Woodward, J.A., 1975. Unreliability of difference scores: a paradox for measurement of change. Psychological Bulletin 82, 85–86.

Pfefferbaum, A., 1990. Model estimates of CSF and skull influences on scalp-recorded ERPs. Alcohol 7 (5), 479–482.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson Jr., R., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37 (2), 127–152.

Reid, S.A., Duke, L.M., Allen, J.J.B., 1998. Resting frontal electroencephalographic asymmetry in depression: inconsistencies suggest the need to identify mediating factors. Psychophysiology 35 (4), 389–404.

Reilly, E.L., 1987. EEG recording and operation of the apparatus. In: Niedermeyer, E., Silva, F.H.L.D. (Eds.), Electroencephalogaphy: Basic Principles, Clinical Applications, and Related Fields, second ed. Urban & Schwarzenberg, Inc., Baltimore, pp. 57–77.

Schwartz, G.E., Fair, P.L., Salt, P., Mandel, M.R., Klerman, G.L., 1976. Facial muscle patterning to affective imagery in depressed and nondepressed subjects. Science 192 (4238), 489–491.

Shoda, Y., Mischel, W., 1996. Toward a unified, intra-individual dynamic conception of personality. Journal of Research in Personality Special Issue: The Future of Personality 30 (3), 414–428.

Shoda, Y., Mischel, W., Wright, J.C., 1994. Intraindividual stability in the organization and patterning of behavior: incorporating psychological situations into the idiographic analysis of personality. Journal of Personality & Social Psychology 67 (4), 674–687.

Steriade, M., Deschenes, M., Domich, L., Mulle, C., 1985. Abolition of spindle oscillations in thalamic neurons disconnected from nucleus reticularis thalami. Journal of Neurophysiology 54 (6), 1473–1497.

Stern, J.A., Walrath, L.C., Goldstein, R., 1984. The endogenous eyeblink. Psychophysiology 21 (1), 22–33.

Sutton, S.K., Davidson, R.J., 1997. Prefrontal brain asymmetry: a biological substrate of the behavioral approach and inhibition systems. Psychological Science 8 (3), 204–210.

Tomarken, A.J., Davidson, R.J., Wheeler, R.E., Kinney, L., 1992. Psychometric properties of resting anterior EEG asymmetry: temporal stability and internal consistency. Psychophysiology 29 (5), 576–592.

Travis, L.E., Barber, V., 1938. The effect of tactile stimulation upon the Berger rhythm. Journal of Experimental Psychology 22, 269–272.

Travis, L.E., Knott, J.R., Griffith, P.E., 1937. Effect of response on the latency and frequency of the Berger rhythm. Journal of General Psychology 16, 391–401.

Wheeler, R.E., Davidson, R.J., Tomarken, A.J., 1993. Frontal brain asymmetry and emotional reactivity: a biological substrate of affective style. Psychophysiology 30 (1), 82–89.

Zimmerman, D.W., Williams, R.H., Zumbo, B.D., 1993. Reliability of measurement and power of significance tests based on differences. Applied Psychological Measurement 17, 1–9.