Psychophysiology, 41 (2004), 269–280. Blackwell Publishing Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2003.00149.x

The stability of resting frontal electroencephalographic asymmetry in depression

JOHN J. B. ALLEN,[a] HEATHER L. URRY,[b] SABRINA K. HITT,[a] AND JAMES A. COAN[a]
[a]Department of Psychology, University of Arizona, Tucson, Arizona, USA
[b]Department of Psychology, University of Wisconsin–Madison, Madison, Wisconsin, USA

Abstract

Although resting frontal electroencephalographic (EEG) alpha asymmetry has been shown to be a stable measure over time in nonclinical populations, its reliability and stability in clinically depressed individuals has not been fully investigated. The internal consistency and test–retest stability of resting EEG alpha (8–13 Hz) asymmetry were examined in 30 women diagnosed with major depression at 4-week intervals for 8 or 16 weeks. Asymmetry scores generally displayed good internal consistency and exhibited modest stability over the 8- and 16-week assessment intervals. Changes in asymmetry scores over this interval were not significantly related to changes in clinical state. These findings suggest that resting EEG alpha asymmetry can be reliably assessed in clinically depressed populations. Furthermore, intraclass correlation stability estimates suggest that although some traitlike aspects of alpha asymmetry exist in depressed individuals, there is also evidence of changes in asymmetry across assessment occasions that are not closely linked to changes in depressive severity.

Descriptors: Frontal EEG asymmetry, Depression, Risk, Psychometric reliability and stability

Resting anterior electroencephalographic (EEG) asymmetry is an individual difference variable that has been associated with traitlike qualities or psychopathological conditions in over 40 studies (for a review, see Coan & Allen, 2003). With respect to depression in particular, a pattern of relatively less left than right frontal activity—inferred by relatively more left than right alpha band activity (see Allen, Coan, & Nazarian, in press)—appears to characterize depressed individuals both when symptomatic (Allen, Iacono, Depue, & Arbisi, 1993; Gotlib, Ranganath, & Rosenfeld, 1998; Henriques & Davidson, 1991; Schaffer, Davidson, & Saron, 1983), as well as during normothymic periods (Allen et al., 1993; Gotlib et al., 1998; Henriques & Davidson, 1990), although not without exception (Reid, Duke, & Allen, 1998). These findings raise the possibility that resting frontal EEG asymmetry may be a relatively stable traitlike marker that distinguishes depressed individuals—or a subset of depressed individuals (cf. Reid et al., 1998)—from never-depressed individuals, and one that does not vary simply as a function of whether individuals are symptomatic. The present study therefore examined whether frontal EEG asymmetry demonstrates stability over repeated sessions despite changes in clinical status among depressed individuals undergoing non-pharmacological treatment for depression.

If frontal EEG asymmetry proves to be a relatively state-invariant marker, it may hold promise to serve as a liability indicator (cf. Iacono & Ficken, 1989) for the development of depression and related psychopathology, identifying those at risk for the subsequent development of the illness. Considerable research is required before this possibility can be fully evaluated, including: (a) investigating whether relative left frontal hypoactivity characterizes a reasonably large subset of depressed persons (sensitivity); (b) investigating whether, in a longitudinal design, left frontal hypoactivity distinguishes depressed from nondepressed individuals and other psychiatric groups (specificity) both during episode and remission; (c) investigating the temporal stability of frontal asymmetry in both depressed and nondepressed individuals; and (d) investigating whether left frontal hypoactivity has prognostic value, to predict the future development or redevelopment of depression. Although the aforementioned studies address to some extent the aspects listed under points a and b above (for review, see Coan & Allen, 2003), relatively little research has addressed aspect c, the temporal stability of resting frontal asymmetry, and the only studies to examine the prognostic value of resting frontal asymmetry involve using asymmetry to portend emotional responses among nondisordered individuals (e.g., Davidson & Fox, 1989; Wheeler, Davidson, & Tomarken, 1993).

This research was supported, in part, by a grant from the National Alliance for Research on Schizophrenia and Depression (NARSAD) to John J. B. Allen and by an Exploratory/Development Grant from the National Institutes of Health (1 R21 RR09492). The authors wish to thank Rosa N. Schnyer for her assistance with treating the patients in this study. Portions of these data were presented at the annual meeting of the Society for Psychophysiological Research, Granada, Spain, October 1999.

Address reprint requests to: John J. B. Allen, Department of Psychology, University of Arizona, P.O. Box 210068, Tucson, AZ 85721-0068. E-mail: jallen@u.arizona.edu.

Addressing the issue of stability, Tomarken, Davidson, Wheeler, and Kinney (1992) assessed the psychometric properties of frontal EEG asymmetry in 85 unselected college students tested on two occasions separated by approximately 3 weeks. Frontal EEG asymmetry assessed at midfrontal and anterior-temporal regions, under both Cz-referenced and computer-averaged ears referenced data, showed high internal consistency (Cronbach's alphas ranging from .81 to .92) and acceptable test–retest stability (intraclass correlations ranging from .53 to .72). Convergent results are provided by Tomarken, Keener, and Neubauer (1994), who recorded EEG of right-handed females three times in 1 year and two times during the following year. The test–retest correlation for midfrontal EEG asymmetry among individual sessions was .57; however, when averaged across 1 year, the overall test–retest stability for midfrontal asymmetry was .82. Similarly, Jones, Field, Davalos, and Pickens (1997) found that Cz-referenced frontal EEG asymmetry recorded at 3 months of age was highly correlated with asymmetry at 3 years ($r = .66$, $p < .01$, at midfrontal leads) in 15 children. Similar figures come from Hagemann, Naumann, Thayer, and Bartussek (2002), who found that across four different measurement occasions in 59 subjects, 52% to 64% of the variance in frontal EEG asymmetry (using a current-source-density derivation) was due to individual differences in a temporally stable latent trait, and 35% to 45% of the variance in frontal asymmetry scores was due to occasion-specific fluctuations. Highly similar estimates were obtained with a computer-linked mastoids derivation.

Stability of EEG asymmetry is implied by studies finding that—similar to currently depressed patients—formerly depressed but currently euthymic depressed patients demonstrate relatively less left frontal activity compared to never depressed controls (Allen et al., 1993; Gotlib et al., 1998; Henriques & Davidson, 1990). Yet these studies do not provide a direct test within subjects. The only studies to address directly the stability of EEG asymmetry in depression have involved relatively small samples, 4 patients and 4 controls in a study of seasonal depression (Allen et al., 1993), and 15 patients and 22 controls in a study of depressed patients initiating pharmacotherapy (Debener et al., 2000). Both studies involved only data from a single reference scheme (Cz in Allen et al., 1993, and linked earlobes in Debener et al., 2000). Whereas the study of seasonal depression found evidence of stability in that depressed patients demonstrated relatively less left frontal activity than controls both when symptomatic and when remitted 2 weeks later, the study of Debener et al. (2000) found greater variability across 2 to 4 weeks among depressed patients than controls. Patients in the Debener et al. study received a variety of antidepressant compounds, most initiated prior to the first EEG assessment, and 11 of the 15 additionally received benzodiazepines. No *systematic* change in asymmetry across sessions was observed in the depressed patients, and no data were provided concerning symptomatic response, but asymmetry was not related to measures of daily mood; the asymmetry was simply more variable across sessions in these patients. Although this could reflect the acute effects of the initiation of a trial of medication, no evaluation of this possibility was performed.

The Present Study

The present study therefore investigated the internal consistency and temporal stability of resting asymmetry in a sample of depressed women receiving a nonpharmacological intervention (acupuncture; Allen, Schnyer, & Hitt, 1998). The efficacy of the intervention is not the focus of this study; rather the focus is to examine the extent to which resting frontal asymmetry remains stable despite changes in clinical status. As stability will necessarily be constrained by reliability of measurement at each time point, the present study also sought to evaluate internal consistency reliability. The present study involves a modestly larger sample than previous studies examining stability in depressed patients, and examines stability in the absence of the administration of a psychopharmacological agent. Additionally, the present study examined a larger scalp montage, compared results using three reference schemes, and addressed three important questions: (1) Can frontal EEG asymmetry be reliably measured in clinical populations, and what length of recording is desirable to obtain reliable estimates of asymmetry?, (2) Is frontal EEG asymmetry stable despite changes in clinical status?, and (3) Are any changes in frontal EEG asymmetry related to changes in clinical state?

Method

Participants

Selection criteria. As described elsewhere (Allen et al., 1998), 38 women (aged 18–45) meeting DSM–IV (American Psychiatric Association, 1994) criteria for current Major Depression were enrolled in a treatment study designed to investigate the efficacy of acupuncture for depression. Diagnoses were determined on the basis of responses to the Structured Clinical Interview for DSM-III-R (SCID-P; Spitzer, Williams, Gibbon, & First, 1990), adapted to allow for diagnosing depression according to DSM–IV criteria. In ongoing research across several studies, these raters obtained interrater and intrarater reliabilities of $r_{icc} = .90$ and $r_{icc} = .88$, respectively.

Potential patients were excluded for: (1) dysthymia or chronic (>2 years) major depression; (2) any other current Axis I disorder; (3) history of psychosis or mania; (4) substance abuse or dependence within the past 4 months; (5) any current treatment; (6) endocrine abnormalities; (7) history of CNS lesions, head trauma, or any medical disorder or treatment that could cause depression; (8) active suicidal potential necessitating immediate treatment; or (9) pregnancy.

Exclusions and data loss. Participants were assessed at 4-week intervals, with the study design involving either three or five assessments depending on the treatment group (see below) to which participants were randomly assigned. Four women terminated after only a single EEG assessment, and were excluded from analysis. Among the remaining 34 women, 3 were excluded from the present study due to left- or mixed-handedness, and 1 due to history of concussion. This yielded a final sample of 30 women for analysis of resting EEG asymmetry data. EEG data from the baseline recording (see below) from 7 of the 30 participants reported here were also reported elsewhere (in Study 2 of Reid et al., 1998). Following the midpoint assessment, one of the three treatment groups ($n = 11$) terminated the study by design, and 1 of the participants from the other treatment groups terminated prior to the conclusion of treatment, leaving 18 participants with EEG data spanning all five assessment sessions.

Among the 30 participants with EEG data available for the first 8-week treatment phase, 1 was missing baseline data because it could not be retrieved from archival tape, leaving 29 participants at the baseline assessment. EEG recordings were not

obtained for 2 other participants at the 4-week assessment session, and were not obtained for 1 other participant at the 8-week assessment, leaving 28 and 29 participants with complete data at 4 and 8 weeks, respectively, and 26 with complete data at all three assessment sessions. After the 8-week assessment, among the 18 participants continuing in the study, all 18 had EEG recorded at the 12- and 16-week assessments, but 3 were among those who had not completed one of the first three assessment sessions. Thus 15 participants had complete data across all five assessment sessions.

Sample description. Depression severity was assessed with a 31-item modified version of the Hamilton Rating Scale for Depression (HRSD) modeled after the revision of the HRSD by Gelenberg et al. (1990), with 19 of the items in this version directly addressing DSM–IV symptoms of depression (for details, see Allen et al., 1998). These 19 items, reported in Allen et al. (1998) and used in the present study as well, will be referred to as the DepHRSD.[1] An assessment of interrater reliability, calculated on a random sample of 22 interviews conducted from this and other studies ongoing in the same laboratory during this time period, yielded an intraclass correlation of .96. The 30 women reported here had mild to moderate depression (DepHRSD mean $\pm$ *SD* of 25.1 $\pm$ 7.0), with a mean duration of the current episode of 9 months ($\pm$ 7.1 *SD*). Ten of these 30 women had previous episodes that were too numerous or indistinct to count, with the remaining 20 having a history of 3.7 ($\pm$ 3.0 *SD*) episodes, including the current episode. Sixty-five percent of the 30 patients reported that one or more first-degree relatives also suffered from depression of comparable severity.

Procedure

Participants were assessed on three or five occasions spanning 8 or 16 weeks. The design of the study involved randomly assigning participants to one of three experimental conditions: (1) 8 weeks of acupuncture treatment specifically designed to address depressive symptoms (termed SPECIFIC), after which these participants concluded the study; (2) 8 weeks of an active control acupuncture treatment designed to address symptoms that are not part of the depressive symptomatology (termed NONSPECIFIC), followed by 8 weeks of treatment specifically for depression; and (3) 8 weeks of wait list (termed WAITLIST), followed by 8 weeks of treatment specifically for depression. Participants were assessed prior to randomization and again monthly while enrolled in the study, yielding three assessments (at baseline, 4, and 8 weeks) for the first group, and five assessments (at baseline, 4, 8, 12, and 16 weeks) for the later two groups.

At each assessment, prior to EEG recording, participants were interviewed by a trained clinical rater blind to group assignment using the HRSD interview. Additionally, at the baseline assessment, participants completed additional questionnaires (handedness [Chapman & Chapman, 1987], footedness [modeled after Chapman, Chapman, & Allen, 1987], and a medical history and medical information questionnaire).

At the conclusion of the HRSD interview, participants were prepared for electroencephalographic (EEG) recording. Participants were fitted with a stretch-lycra cap with tin electrodes. Participants were fitted with three additional tin electrodes to monitor eye movements, placed on the nasion and directly below the pupil of each eye at a position equal to 20% of the nasion–inion distance below FP1 and FP2. Impedances at all sites were required to be less than 5 KΩ, with homologous sites within 1 KΩ of one another. Signals were recorded from 25 sites according to the International 10–20 system (Fz, Cz, Pz, FP1, FP2, F3, F4, F7, F8, C3, C4, P3, P4, T3, T4, T5, T6, O1, O2, FTC1, FTC2, TCP1, TCP2, PO1, PO2) in addition to the right mastoid (A2) and the three ocular sites (nasion, left inferior orbit, right inferior orbit). All EEG and ocular sites were referenced to the left mastoid (A1) online and were recorded with AC differential amplifiers (bandpass 0.1 to 100 Hz). Frontalis electromyographic (EMG) activity was recorded via a bipolar arrangement, with free tin electrodes placed on the right and left frontalis muscles referenced to the frontal-pole EEG sites (FP1 and FP2). Data were digitized continuously at 512 Hz.

After preparation for EEG recording, participants were seated in a comfortable chair in a sound-dampened chamber with dim (25-W bulb) illumination. Resting EEG was recorded for 8 min, in blocks with eyes open (O) and with eyes closed (C), in one of two counterbalanced orders (OCCOCOOC or COOCOCCO).

Electroencephalographic Data Processing

Each record was visually screened to remove epochs with movement and muscle artifacts, following which a computer-based blink rejection algorithm rejected any epoch with activity greater than $\pm$ 50 μV in amplitude (the prototypic minimum amplitude of blinks) in an off-line derived ocular channel (nasion-left inferior orbit). Each participant's data were then rereferenced off-line to Cz, to computer averaged mastoids (LM; "linked" mastoids), and to the average of EEG sites (AR; average reference). Although rational arguments have been levied in favor of one or another reference scheme (e.g., Hagemann, Naumann, & Thayer, 2001; Reid et al., 1998), it remains an empirical question which reference scheme may be preferred in assessing risk for depression and other psychopathology. Until this issue is resolved, it is prudent to report results from multiple reference montages so that data are available for future investigators to compare across studies. In the present study, analyses were thus conducted in parallel on data from each of the three reference schemes, and the results of each reference scheme will be reported separately.

Each 1-min EEG block was divided into 119 2-s epochs that overlapped by 1.5 s. A fast Fourier transform (FFT) was applied to all artifact-free epochs, after the data had been weighted with a Hamming window that tapered the distal 10% of each epoch. The power spectra for each of these epochs were then averaged for each minute of recording, and weighted averages across minutes (weighted by the number of epochs in each minute) were created. The percentage of rejected epochs was quite consistent across sessions: Baseline, 50%; 4-week, 49%; 8-week, 50%; 12-week, 53%; and 16-week, 54%. Average power in the 8–13 Hz band was taken as an index of alpha power. Finally, an asymmetry score was computed by taking the difference of natural log transformed scores for all sites that had symmetrical

[1]The 19 items included in the DepHRSD are depressed mood, suicidal thinking, suicidal behavior, difficulty falling asleep, middle of the night awakening, early morning awakening, hypersomnia, loss of interest, loss of pleasure, psychomotor retardation, psychomotor agitation, loss of energy, appetite change, weight change during the past month, loss of sexual interest, decreased concentration, helplessness, hopelessness, and worthlessness/failure. The nonincluded items on the HRSD reflect symptoms such as anxiety (psychic anxiety, somatic anxiety, obsessive-compulsive symptoms) and other nondepressive symptoms (e.g., paranoia, depersonalization).

left and right locations (FP1 and FP2, F3 and F4, F7 and F8, C3 and C4, P3 and P4, T3 and T4, T5 and T6, O1 and O2, FTC1 and FTC2, TCP1 and TCP2, PO1 and PO2). The asymmetry score was computed such that the left log transformed score was always subtracted from the right (i.e., ln[Right] – ln[Left]), with higher values on this index putatively reflecting relatively greater left activity (i.e., relatively greater right alpha). The natural log transformation is customary in research examining EEG asymmetry, as EEG power values tend to be positively skewed (e.g., Allen et al., in press; Tomarken, Davidson, Wheeler, & Kinney, 1992).

Results

Results will be presented first addressing the internal consistency and stability of EEG asymmetry, followed by an examination of whether changes in clinical status are related to changes in EEG asymmetry and whether a trait estimate of frontal EEG asymmetry can predict treatment response or relapse 6 months following the conclusion of treatment. In all cases, the focus will be primarily on three frontal regions: midfrontal (F4-F3), lateral frontal (F8-F7), and fronto-temporal-central (FTC2-FTC1). Means for all regions and reference schemes are presented in Table 1.

Internal Consistency

Internal consistency of EEG alpha asymmetry was assessed by treating each 1-min recording period as an item on an eight-item scale (cf. Tomarken, Davidson, Wheeler, & Kinney, 1992). Table 2 reveals that resting EEG alpha asymmetry scores demonstrated generally high internal consistency. Internal consistency coefficients (Cronbach's alpha) for alpha asymmetry scores across regions and reference schemes at baseline ranged from .76 to .94, with a median of .88, and across all five assessment sessions they ranged from .37 to .94, with a median of .85. Given the special interest in frontal alpha asymmetry, internal consistency coefficients for alpha asymmetry at frontal regions (F34, F78, and FTC12) across reference schemes ranged from .86 to .89, with a median of .89 at baseline, and from .61 to .92, with a median of .86 across five assessment sessions.

Reviewers often suggest that 8 min of resting EEG asymmetry are desirable to obtain adequate internal consistency reliability, as this was the number reported in the only other psychometric investigation of resting EEG alpha asymmetry (Tomarken, Davidson, Wheeler, & Kinney, 1992). As few as five 1-min samples, however, also produced acceptable estimates of internal consistency in that study (Tomarken, Davidson, Wheeler, & Kinney, 1992), and estimates based on even shorter time frames of 2 min have proven similarly reliable (Coan, Allen, & Harmon-Jones, 2001).

The approach taken by Tomarken, Davidson, Wheeler, and Kinney (1992) to assess the reliability of fewer than 8 min of data confounded the length of recording with the number of discrete items included in the calculation of coefficient alpha, that is, they used the Spearman–Brown prophecy formula to estimate the reliability for shorter recording periods, estimating based on the assumption of six asymmetry values for 6 min of data, seven values for 7 min, and eight values for 8 min.The best assessment of whether fewer minutes of recording would produce comparable estimates of internal consistency would involve keeping 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 & Novick, 1968). Thus in the present study, the first 2, 4, and 6 min of recorded data were divided into eight blocks each. Each block contained 2-s overlapping epochs that were subjected to Fourier analysis as detailed above. Each block contained 89 segments for the 6-min data set, 59 for the 4-min data set, and 30 for the 2-min data set.

Table 1. *Mean (± SE) Asymmetry Scores at Baseline by Region and Reference Scheme*

	Average reference	Cz reference	Linked mastoid reference
FP12	.037 (.007)	.004 (.008)	.044 (.008)
F34	– .008 (.027)	– .044 (.025)	.045 (.021)
F78	.037 (.019)	– .027 (.022)	.045 (.012)
FT12	.010 (.030)	.023 (.036)	– .001 (.021)
T34	– .019 (.039)	.002 (.035)	– .017 (.032)
C34	.009 (.043)	– .015 (.052)	.011 (.025)
TC12	– .027 (.041)	.007 (.045)	– .020 (.033)
T56	.148 (.058)	.098 (.046)	.105 (.050)
P34	.022 (.043)	.084 (.042)	– .021 (.021)
PO12	– .023 (.025)	– .005 (.021)	– .018 (.015)
O12	.062 (.042)	.049 (.031)	.039 (.038)

Note: Asymmetry scores calculated by ln(Right) − ln(Left). $N = 29$ for all sites except T3T4, where $N = 28$.

Table 2. *Internal Consistency Reliability (Cronbach's Alpha) of EEG Alpha Asymmetry Scores across 8 min as a Function of Region, Reference Scheme, and Assessment Session*

	Average reference					Cz reference					LM reference				
Session	Baseline	4-Week	8-Week	12-Week	16-Week	Baseline	4-Week	8-Week	12-Week	16-Week	Baseline	4-Week	8-Week	12-Week	16-Week
Region															
FP12	.85	.82	.82	.77	.65	.86	.84	.87	.87	.78	.88	.86	.83	.81	.75
F34	.89	.86	.84	.84	.75	.89	.91	.92	.89	.81	.86	.83	.83	.85	.85
F78	.89	.87	.86	.83	.76	.86	.88	.90	.86	.86	.89	.89	.83	.84	.70
FT12	.88	.82	.84	.83	.61	.89	.87	.86	.80	.84	.89	.89	.86	.89	.85
T34	.86	.87	.83	.79	.70	.89	.88	.87	.84	.88	.86	.89	.83	.84	.76
C34	.87	.88	.89	.85	.86	.92	.76	.92	.87	.78	.87	.85	.86	.84	.82
TC12	.90	.86	.87	.86	.69	.89	.74	.87	.84	.81	.85	.87	.84	.80	.67
T56	.83	.85	.81	.61	.71	.86	.81	.84	.74	.79	.85	.83	.78	.86	.46
P34	.84	.77	.72	.75	.50	.86	.76	.85	.80	.83	.76	.79	.74	.67	.37
PO12	.89	.89	.91	.83	.89	.90	.80	.88	.86	.92	.86	.91	.93	.81	.81
O12	.90	.87	.88	.83	.84	.88	.84	.86	.81	.81	.94	.90	.88	.88	.85

Notes: Asymmetry scores calculated as ln(Right) − ln(Left) alpha power. Due to bad electrodes or missing assessments, the ranges of the number of subjects at each time of assessment are: baseline, 24–25; 4-week, 26–27; 8-week, 24–25; 12-week, 15–16; 16-week, 17–18.

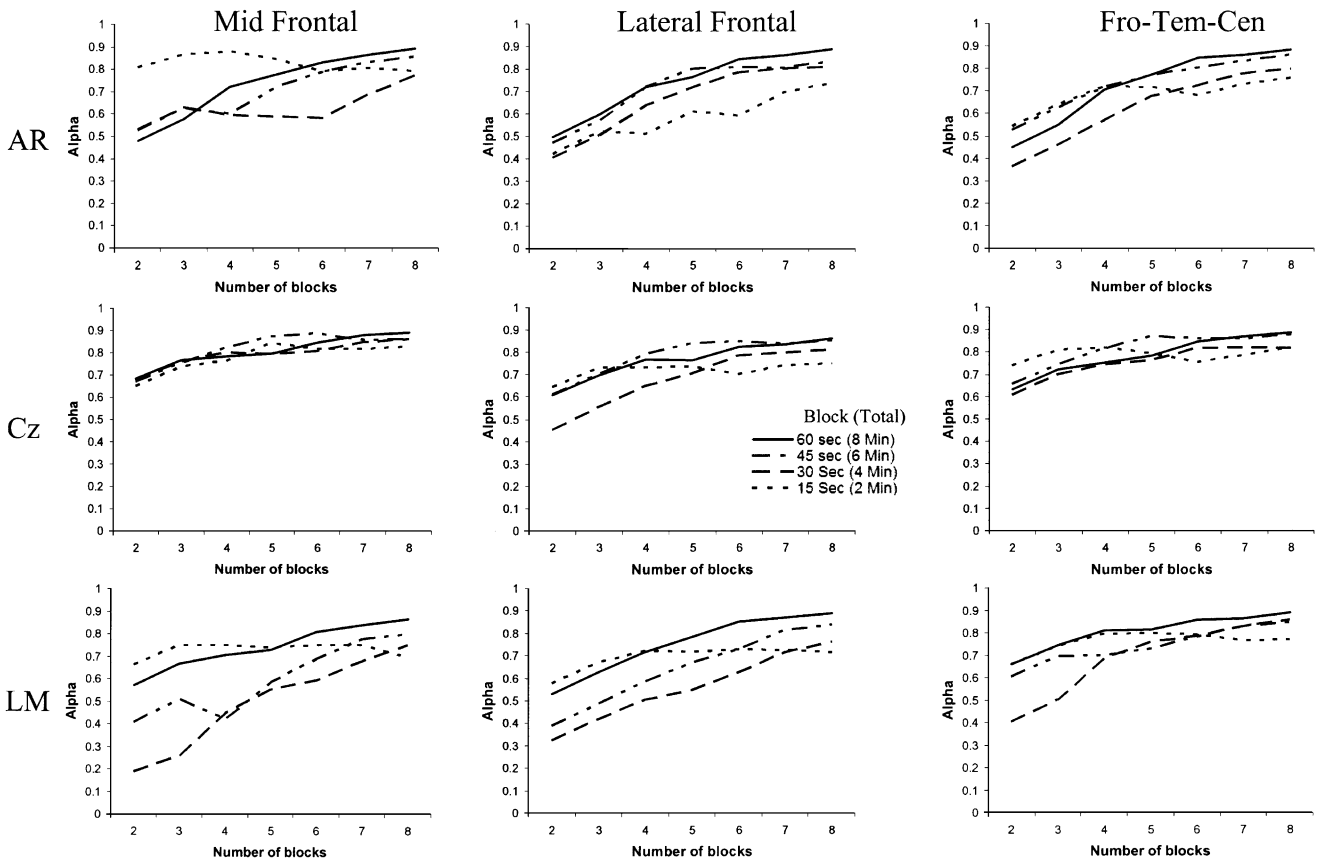

Figure 1. Cronbach's alpha internal consistency estimates for resting alpha asymmetry as a function of site, reference scheme, length of data recording, and number of blocks (items) used to calculate alpha. The number of participants ranges from 19 to 28, reflecting that some participants did not have 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 participant.

In each case, eight asymmetry values were obtained for each site, reflecting the asymmetry score averaged across one-eighth of the total time of recording (15 s for the 2-min data, 30 s for 4-min data, or 45 s for the 6-min data). These eight values were then treated as items on an eight-item scale to assess internal consistency reliability.

Figure 1 shows the results for frontal sites as a function of reference scheme. As can be seen, the number of minutes of recording does not affect the estimate of internal consistency as much as does 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 blocks 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.

Stability

Stability was measured by computing intraclass correlations (ICCs) across measurement occasions. Unlike the Pearson product moment test–retest correlation, the ICC allows for the inclusion of more than two occasions of measurement and is sensitive not only to changes in the rank ordering of participants across assessments (as is the Pearson correlation), but also to absolute differences in scores within persons across sessions.

ICCs were derived in the present study using a one-way random effects model, which corresponds to model (1,1) of Shrout and Fleiss (1979). This model, a bit more conservative than Shrout and Fleiss' model (3,1) implemented by Tomarken, Davidson, Wheeler, and Kinney (1992), assumes that the assessment sessions are randomly selected from a larger population of assessment times, and that every participant is rated at each time of assessment. This assumption allows one to generalize beyond the specific three or five observed assessment sessions to infer what the reliability would be at *any* three or five comparable assessment sessions. Model (3,1), by contrast, would only allow one to generalize to the particular observed assessment sessions, and will therefore produce ICCs that are slightly higher than those of model (1,1).[2]

[2]The primary difference between the two models is whether, in deriving the ICC, one uses the within-subjects mean square (in Model [1,1]) from a one-way ANOVA with subjects as a between-subjects random effect or the error mean square (in Model [3,1]) from a two-way mixed ANOVA with subjects as a between-subjects random effect and assessment session as a within-subjects fixed effect. Model (1,1) was used in the present study, providing a more conservative and generalizable estimate of stability:

$$ICC(1,1) = \frac{MS_B - MS_W}{MS_B + (k-1)MS_W}$$

$$ICC(3,1) = \frac{MS_B - MS_E}{MS_B + (k-1)MS_E},$$

where MS_B = mean square between subjects, MS_W = mean square within subjects, MS_E = mean square error, and k = number of assessment sessions.

An additional consideration in presenting ICCs concerns whether one is interested in each session as the unit of analysis or the average of all sessions as the unit of analysis. If one were interested in ultimately averaging data across all sessions to derive an estimate of trait EEG asymmetry, then one would be interested in the average-measure ICC, or models (1,k) or (3,k) in Shrout and Fleiss (1979). Such data were presented as ICC2 by Tomarken, Davidson, Wheeler, and Kinney (1992), and reflect the relatively common practice of amalgamating asymmetry scores from multiple sessions to derive a better estimate of trait-related asymmetry (see also Hagemann et al., 2002). If, by contrast, one is interested in estimating the extent to which EEG asymmetry is likely to change across any given assessment, asking how stable it remains across time and changes in clinical status, then one would be interested in the single-measure ICC (presented as ICC1 in Tomarken, Davidson, Wheeler, & Kinney, 1992), or models (1,1) and (3,1) of Shrout and Fleiss (1979).[3] Thus, with an interest in interpreting the ICCs to reflect the extent to which one might expect asymmetry scores to remain stable across any given set of assessments while participants experience changes in clinical status over time, a one-way random effect model (1,1) was used to derive ICCs in the present study. These estimates will be lower than those of the other methods presented, and thus reflect the most conservative estimate of stability.

Asymmetry scores showed reasonable stability over the course of 8 and 16 weeks. As Figure 2 indicates, ICCs ranged from .41 to .77 over 8 weeks, and from .33 to .85 over 16 weeks. Across the three assessments (8 weeks), the median ICC was .62 for average-referenced, .73 for Cz-referenced, and .57 for LM-referenced asymmetry scores. Across the five assessments (16 weeks), the median ICC was .61 for average-referenced, .54 for Cz-referenced, and .61 for computer-linked-mastoid-referenced asymmetry scores. Considering only the frontal regions (mid-frontal, lateral frontal, and fronto-temporal central), the median ICC across three assessments was .56, .76, and .41 for AR-, Cz-, and LM-referenced data; across five assessments, the comparable medians were .61, .60, and .61 for AR-, Cz-, and LM-referenced data.

To estimate how stable asymmetry scores would be given shorter recording periods, data were again analyzed for 2, 4, 6, and 8 min at each of the first three assessments (baseline, 4 weeks, 8 weeks). As shown in Figure 3, longer data-recording periods lead to modest increments in estimates of stability. To provide a more complete picture of recording period as a source of variance

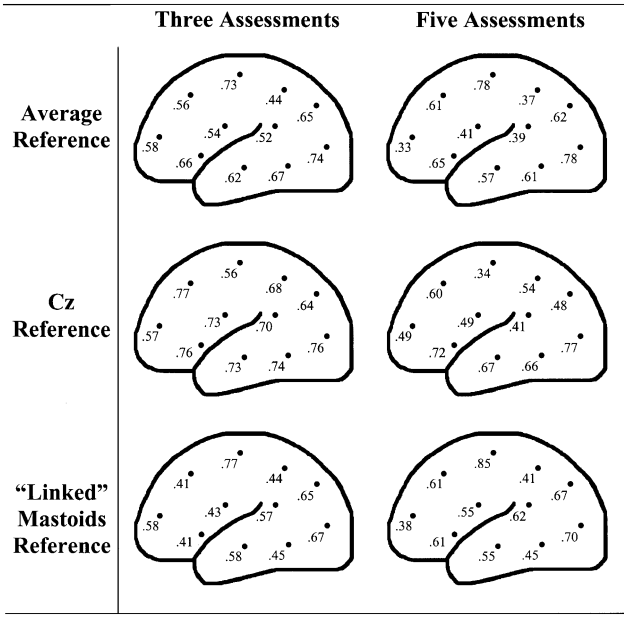

Figure 2. Intraclass correlations (calculated according to Shrout and Fleiss, 1979, model [1,1]) summarizing stability of EEG alpha asymmetry across three recording sessions (8 weeks, 26 patients) or five recording sessions (16 weeks, 15 patients) as a function of region and reference scheme.

across measurement occasions, a variance component analysis of asymmetry scores across individuals, recording length (2, 4, 6, or 8 min), measurement occasions (three), and reference scheme (AR, Cz, LM) was conducted. This analysis was repeated for data from each frontal region. In all regions, the percentage of variance accounted for by the length of the recording period was less than 1%. Thus, relative to trait individual differences in frontal asymmetry and occasion-specific deviations from trait variance, the impact of recording length (across the range of 2 to 8 min) is negligible.

Relationship of EEG to Clinical Status

The modest stability of resting anterior EEG asymmetry scores in this sample occurred despite a general tendency for participants to experience improvements in their clinical state over time. DepHRSD scores declined significantly across 8 weeks for the entire sample, $F(2,58) = 15.5$, $p < .001$, and across 16 weeks for the sample completing 16 weeks, $F(4,64) = 17.3$, $p < .001$. Reported significance values reflect Greenhouse–Geisser corrections for deviations from sphericity. Replicating the findings of the slightly larger sample ($n = 34$) reported in Allen et al. (1998) with the 30 participants retained for the present EEG asymmetry analysis, participants receiving treatment specifically for depression demonstrated larger change across the 8 weeks of treatment than participants receiving nonspecific treatment, as indexed by a significant Time × Treatment interaction, $F(4,54) = 2.9$, $p < .05$, in the 3 (treatment group) × 3 (time: baseline, 4-week, 8-week) repeated measures analysis of variance (ANOVA). The 12 participants receiving treatment specifically for depression experienced a drop of 11.7 points on the DepHRSD, significantly ($p < .05$) more than those receiving

[3]The computation of models (1,k) and (3,k) again differ from one another in terms of whether one uses the within-subjects mean square or the error mean square from ANOVAs described in footnote 2. These two average-measure models ([1,k] and [3,k]) differ from the single-measure models ([1,1] and [3,1]) computationally by whether one is "penalized" (in the single-measure models) as a function of the number of assessment sessions by including a term in the denominator. The computational formulae for the average measure estimates of stability thus do not include a term involving the number of sessions in the denominator, and are presented below for reference:

$$ICC(1,k) = \frac{MS_B - MS_W}{MS_B + MS_W}$$

$$ICC(3,k) = \frac{MS_B - MS_E}{MS_B + MS_E},$$

where MS_B = mean square between subjects, MS_W = mean square within subjects, MS_E = mean square error, and k = number of assessment sessions.

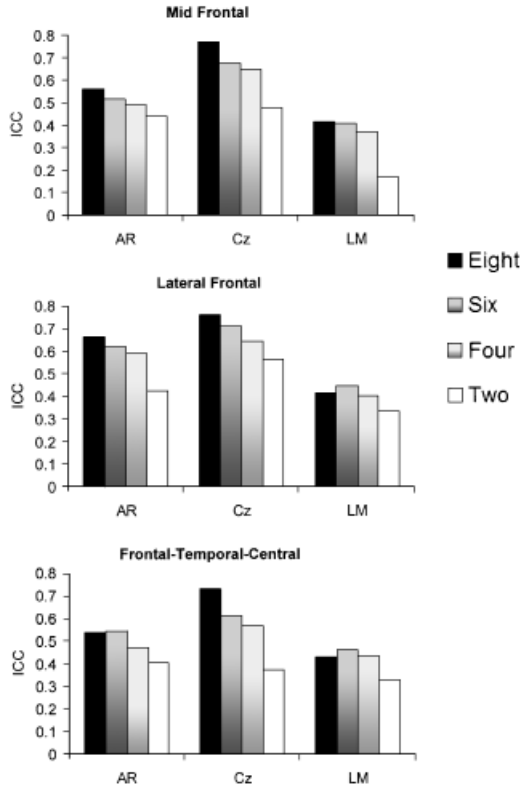

Figure 3. Intraclass correlations (calculated according to Shrout and Fleiss, 1979, model [1,1]) summarizing stability of EEG alpha asymmetry across three recording sessions (8 weeks) in 26 patients as a function of region, reference scheme, and length of data recording period (2, 4, 6, or 8 min).

nonspecific treatment (drop of 1.2 points), and nonsignificantly more than those on the waitlist (drop of 7.6 points).

Baseline measures. Frontal EEG asymmetry at baseline was unrelated to depressive severity at baseline for the 29 participants with complete data: at midfrontal leads (rs = .08, −.08, −.03 for AR, Cz, LM); at lateral frontal leads (rs = .14, .05, .02 for AR, Cz, LM); or at fronto-temporal-central leads (rs = .32, .03, .11 for AR, Cz, LM). Moreover, baseline frontal EEG asymmetry was not a significant predictor of change in depressive severity across 8 weeks as assessed by the change in DepHRSD scores (baseline minus 8-week score, with higher numbers reflecting greater improvement): at midfrontal leads (rs = −.12, −.12, −.18 for AR, Cz, LM); at lateral frontal leads (rs = .09, .10, .00 for AR, Cz, LM); or at fronto-temporal-central leads (rs = .14, .11, −.11 for AR, Cz, LM).

Change over time. To assess whether clinical changes were mirrored by changes in frontal asymmetry, frontal sites (midfrontal, lateral-frontal, and fronto-temporal-central) were examined as correlates, and potential mediators, of clinical change. Zero order correlations summarizing the relationship between the change in frontal EEG asymmetry and change in DepHRSD scores were not significant, ranging from −.14 to +.18, with the exception of one trend (r = .32, p = .06) for a decrease in depressive symptom severity to be associated with an increase in relative left frontal activity at lateral-frontal leads under the Cz reference scheme. Although this analysis has limited statistical power, no robust relationship exists between changes in frontal EEG asymmetry from session to session and current depressive severity, suggesting that changes in asymmetry from session to session reflect other factors unique to each occasion of measurement.

To provide a more robust test of whether a pattern of change in frontal EEG asymmetry might underlie change in depressive symptomatology, a multivariate repeated measures analysis of variance with changing covariates was implemented. DepHRSD scores at baseline, 4 weeks, and 8 weeks served as dependent variables as a function of time, with three changing covariates: midfrontal asymmetry, lateral frontal asymmetry, and fronto-temporal-central asymmetry across the three assessments. The analysis was repeated for each reference scheme, and in no case did the covariates account for a significant proportion of variance (all Fs(3,45) < 1.1, all ps > .39). Moreover, in each case, covariate-adjusted depression severity decreased significantly over time (all Fs(2,45) > 12.8, all ps < .001), a finding that would not have obtained if the pattern of frontal EEG asymmetry across the three assessments had mediated depressive severity over the same time frame.

Predictive Utility of EEG Asymmetry

To assess whether trait frontal EEG asymmetry would be a predictor of clinically relevant variables, a trait estimate (cf. Hagemann et al., 2002; Sutton & Davidson, 1997) of frontal asymmetry was computed by averaging asymmetry scores across the first three assessment sessions for each participant, separately for each region and reference scheme, yielding averaged values at midfrontal, lateral frontal, and fronto-temporal-central regions derived under each reference scheme. Because an average was used, missing data did not lead to a reduction in the number of included participants; all analyses in this section thus involve all 30 participants, except those for predicting relapse, which include only the 23 participants for which follow-up data were available (see Gallagher, Allen, Hitt, Schnyer, & Manber, 2001, for details on the follow-up protocol).

As shown in Table 3, the trait estimates of frontal asymmetry were largely unrelated to baseline depressive severity, change in depression severity across the 8 weeks, or the likelihood of being depressed among those who were contacted 6 months following the completion of the study. Thirteen of the 23 individuals contacted 6 months following the completion of the study had experienced full remission at the conclusion of treatment. Among these 13, 3 had relapsed, again meeting the full DSM–IV criteria for depression. These numbers were too small to permit an adequate test of whether frontal EEG asymmetry predicts relapse among responders, but in the interest of exploration, these 13 participants were use to examine the point-biserial correlation between the dichotomous variable relapse status and frontal EEG asymmetry. No correlations were significant, ranging from −.32 to .12.

The Contribution of Overall Power to Asymmetry Findings

To assess whether the asymmetry score (ln[Right] – ln[Left]) may have been confounded by overall differences in alpha power, three tests were conducted, all using the baseline assessment data. First, the sum of power at homologous leads (ln[Right]+ln[Left]) was correlated with the asymmetry difference score (ln[Right] − ln[Left]), at each of 11 scalp regions under all three reference schemes. Only 1 of these 33 correlations was significant (temporal-central-parietal under linked mastoids, r = .44, p < .05). In no other case was there even a trend for significance. Across all 33

Table 3. *Correlations between Trait Estimates of Frontal EEG Asymmetry and Clinical Measures*

Region	Clinical Measure	Reference		
		AR	Cz	LM
F8–F7	Baseline	.17	.09	.16
	Change	.15	.20	.04
	Relapse	.02	.11	−.17
F4–F3	Baseline	−.04	−.03	−.12
	Change	−.16	.00	−.17
	Relapse	−.17	−.09	.00
FTC2–FTC1	Baseline	.31*	.15	.12
	Change	.11	.18	−.04
	Relapse	.10	.02	.21

Notes: Trait estimates of EEG asymmetry represent averages across the first three assessments (baseline, 4-week, and 8-week). Baseline: Baseline DepHRSD score; Change: Baseline DepHRSD score minus DepHRSD score at 8 weeks; Relapse: dichotomous value reflecting relapse after 6 months (coded 1) or no relapse after 6 months (coded 0).
*$p < .10$, two tailed. $N = 30$ for baseline and change measures, $n = 23$ for relapse measure.

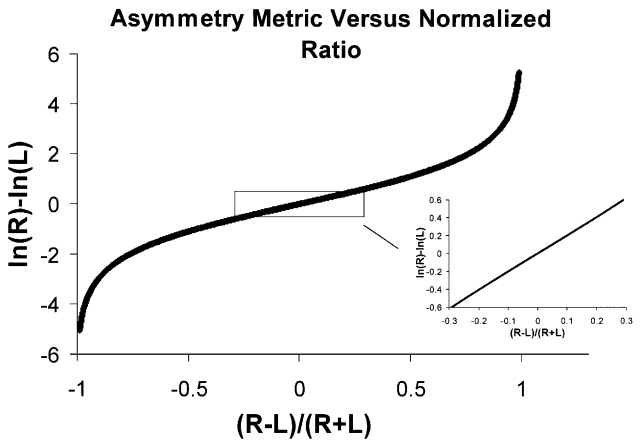

Figure 4. The relationship of the asymmetry metric (ln[Right] − ln[Left]) and a metric normalized for overall power ([R − L]/[R + L]), over a large range of scores. In asymmetry research, the ln(Right) − ln(Left) metric produces scores that typically are in the range of ± 0.5, the range demarcated by the two lines, where the relationship is linear.

correlations, the arithmetic average correlation was .046, and the median was .023, further suggesting no relationship between total alpha power at constituent sites and the asymmetry difference metric.

To address more specifically whether large overall alpha power may have contributed to asymmetry scores, a total alpha power score was computed for each reference scheme, representing the sum of the 22 sites involved in the asymmetry metrics. Only 2 of the 33 correlations between this total power score and the asymmetry metric were significant (temporal-central-parietal region under linked mastoids reference, $r = .42$, $p < .05$, and also under the average reference, $r = .38$, $p < .05$).

Finally, a "normalized" difference score was computed as suggested by an anonymous reviewer. This normalized score, operationalized as (R − L)/(R + L), correlated over .99 with the asymmetry metric (ln[Right] − ln[Left]). 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 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 Figure 4.

These three tests provide little support for the hypothesis that overall alpha power contributes significantly to the widely used asymmetry metric (ln[Right] − ln[Left]). Nonetheless, it remains a possibility that overall alpha power may influence the *stability* of the asymmetry score over time. To assess this possibility, stability across the baseline, 4-week, and 8-week assessments was examined, as fewer participants have data after the 8-week assessment, which would result in a test with less statistical power. Across these three assessment sessions, stability was operationalized as the root-mean-square (RMS) deviation from the mean of the three sessions. This was calculated for each asymmetry score (11 regions × 3 reference schemes), RMS values near zero indicate very similar scores across the three sessions, with larger RMS values reflecting greater deviation from session to session, regardless of direction. These RMS scores were then correlated with total alpha power at baseline, and only 1 of the 33 correlations was significant. Under the average reference scheme, total power correlated with mid-frontal asymmetry, $r = -.40$, $p < .05$, indicating greater total power was associated with less variation across sessions; similar correlations were not observed for the midfrontal region under the Cz reference ($r = .23$) or the computer linked mastoids reference ($r = -.17$). Across all 33 correlations, the arithmetic average correlation was −.059, and the median was −.071, further suggesting no strong or systematic relationship between total alpha power at baseline and consistency over time.

Discussion

Resting anterior EEG alpha asymmetry demonstrates stability in depressed patients that is comparable in magnitude to that seen in nonclinical samples (e.g., Hagemann et al., 2002; Tomarken, Davidson, Wheeler, & Kinney, 1992). At mid-frontal, lateral-frontal, and fronto-temporal-central sites, almost 60% of the variance is stable, as the median ICC at these sites (across reference schemes) was .56 across three assessments and .61 across five assessments. This stability is apparent despite rather substantial improvements in clinical status over the same interval. Nonetheless, superimposed on this traitlike stability are occasion-specific fluctuations, and unreliability of measurement, the latter accounting for comparatively little variance. One might have predicted that depressive severity would be an important occasion-specific factor that would be related to occasion-specific fluctuations in frontal asymmetry, but several tests failed to support this possibility. Although a small effect size would be difficult to statistically support given the present sample size, it is clear that if a relationship exists between clinical state and resting EEG asymmetry, it is not large.

Severity

The lack of a relationship of baseline frontal asymmetry to depressive severity within this clinically depressed sample is worth elaborating upon. Although Schaffer et al. (1983) found that frontal EEG asymmetry distinguished participants with quite high Beck Depression Inventory (BDI) scores from those with very low scores, investigations examining the relationship between severity and frontal EEG asymmetry *within* a clinical sample have found that asymmetry is unrelated to severity (Henriques & Davidson, 1991). A similar pattern is found among

nonclinical samples, in that Tomarken, Davidson, Wheeler, and Doss (1992) found differences in frontal asymmetry as a function of dispositional positive affect using extreme groups, but in samples not selected for extreme levels of self-reported emotion, frontal EEG asymmetry has been found to be unrelated to current unprovoked mood (Sutton & Davidson, 1997).

In the case of Schaffer et al. (1983), the BDI was used not as a measure of severity, but as a selection tool to identify those with considerable depression. In the case of Henriques and Davidson (1991), by contrast, the selection was accomplished by clinical interview to derive extreme groups, with BDI scores examined within groups. Henriques and Davidson (1991) did not compute correlations with the BDI for the sample as a whole (combining depressed and control participants), as participants were selected on the basis of depressive symptoms. Had they split their clinical sample (Henriques & Davidson, 1991) into high and low BDI scores, group membership would have remained unchanged and they would have, of course, found the same relationship as found by Schaffer et al. (1983), with high-BDI participants characterized by relatively lower left frontal activity. The present study did not include a nondepressed control group, so only within-group correlations could be examined.

The finding that depressed patients differ from nondepressed controls in several (Allen et al., 1993; Baehr, Rosenfeld, Baehr, & Earnest, 1998; Gotlib et al., 1998; Henriques & Davidson, 1991; Schaffer et al., 1983), but not all (Reid et al., 1998), studies is consistent with a diathesis stress model, with relatively less left frontal activity tapping a nonnecessary risk factor for the genesis of depression. According to this reasoning, individuals with relatively less left frontal activity are at risk for depression, but not all depressed participants will demonstrate a pattern of relative left hypoactivity, as depression is highly heterogeneous. Further, such a model would suggest that, among those with relative left frontal hypoactivity, current clinical state should not alter the diathesis tapped by frontal asymmetry. Investigations finding a pattern of relative left frontal hypoactivity in formerly depressed but currently euthymic patients provide evidence in support of this interpretation (Allen et al., 1993; Gotlib et al., 1998; Henriques & Davidson, 1990), as do findings of the present study and others (Henriques & Davidson, 1991) that clinical severity is unrelated to frontal asymmetry.

An additional interpretation, and one that need not be mutually exclusive with the diathesis stress model, is that the relationship between frontal EEG asymmetry and mood/severity emerges only when a sample with the full range of scores on both variables is obtained. This psychometric argument assumes that there is an underlying linear relationship between mood/severity and frontal EEG asymmetry, but that a sample with a truncated range of mood scores attenuates the magnitude of the population correlation (Lord & Novick, 1968). Whether the diathesis putatively tapped by frontal EEG asymmetry bears a linear relationship to risk or whether it functions more as a threshold to be surpassed remains an empirical question, and one best answered with a prospective study of a large sample of individuals spanning the full range of mood/severity.

The argument that asymmetry is functioning as a stable diathesis would be bolstered by findings that asymmetry in this sample differs from that of never-depressed control participants. The present data only allow the interpretation that frontal EEG asymmetry can be reliably measured in depressed subjects, and that the degree to which it remains stable over time is comparable to that observed in nondepressed samples (Hagemann et al., 2002; Tomarken, Davidson, Wheeler, & Kinney, 1992). The present results thus support a necessary but nonsufficient condition for inferring that frontal EEG asymmetry can serve as a stable diathesis tapping risk for depression. The degree to which one is willing to accept the proposition that frontal asymmetry taps a diathesis for depression will depend on one's assessment of the literature examining frontal EEG asymmetry and depression. Table 4 presents an abbreviated summary of this literature. As summarized in the table, there is mixed support for concluding that frontal EEG asymmetry may serve as any type of marker, be it episode, liability, or genetic. Clearly the extant data (for a review, see Coan & Allen, 2003) do not support the contention that frontal EEG asymmetry is a diathesis that characterizes all depressed individuals, as not all depressed individuals are characterized by the asymmetry (Reid et al., 1998), but data are quite consistent with the proposition that frontal EEG asymmetry may index risk for depression in at least a subset of those that develop depression. Further, although indirect, support for this proposition derives from studies of infants of depressed mothers, who show evidence of a relative left frontal hypoactivity in comparison to infants of mothers who are not similarly burdened (e.g., Dawson et al., 1999; Field, Fox, Pickens, & Nawrocki, 1995). These findings suggest the possibility that frontal brain asymmetry may be associated with some forms of depression as a function of heredity, parenting, or a combination, even in very early stages of development.

Several studies suggest that relative left frontal hypoactivity may not be specific to depression, characterizing those with significant anxiety as well (Davidson, Marshall, Tomarken, & Henriques, 2000; Heller, Nitschke, Etienne, & Miller, 1997; Wiedemann et al., 1999). Thus frontal EEG asymmetry may index a more general risk for emotional psychopathology. Whether frontal EEG asymmetry represents a nongenetic liability marker or even genetic liability marker, however, has been largely unexplored (Coan, 2003).

Predictive Utility

The diathesis-stress model would suggest that frontal EEG asymmetry may hold predictive utility in identifying those at increased risk of developing depression at some point during life. Although not explicitly a part of the model, it is conceivable, but not supported in the present study, that frontal EEG asymmetry may hold prognostic value for predicting aspects of clinical course, including response during treatment or likelihood of relapse. Because long-term follow-up was not an explicit aspect of the present study design, the follow-up sample included only 23 of the 30 participants with complete EEG data. From among them, 13 were in full remission at the conclusion of treatment, and only 3 relapsed over the course of the subsequent 6 months. Clearly a stronger test is needed before dismissing the possibility that frontal EEG asymmetry may hold prognostic utility for predicting relapse.

The diathesis-stress model would not necessarily predict that trait frontal EEG asymmetry would predict response during treatment. In fact, if frontal EEG asymmetry is a liability marker, it should be relatively unaffected by changes in clinical status. One could imagine that if a diathesis indexed not simply risk for developing depression, but factors related to prognosis such as severity or chronicity (e.g., Beekman et al., 2002), then those with more of the diathesis should have more severe depression and be less responsive to treatment. Neither of these possibilities was observed in the present data, thus suggesting

Table 4. *Characteristics of Psychophysiological Markers as Applied to Frontal EEG Asymmetry for Depression*

Episode	Liability	Genetic
Characterizes most depressed persons (sensitivity)[1,4,5,8,–9,11]	Characterizes most depressed persons (sensitivity)[1,4,5,8,–9,11]	Characterizes most depressed persons (sensitivity)[1,4,5,8,–9,11]
Differentiates depressed from nondepressed (specificity)[1,–3,4,5,–6,–13]	Differentiates depressed from nondepressed, not only in episode but in remission as well[1,–3,7]	Differentiates depressed from nondepressed, not only in episode but in remission as well[1,–3,7]
Changes with variations in clinical state[10]	Demonstrates stability in both depressed and nondepressed individuals[1,–4,12,present report]	Demonstrates stability in both depressed and nondepressed individuals[1,–4,12,present report]
	Predicts the future development of depression in individuals currently not depressed[NA]	Predicts the future development of depression in individuals currently not depressed[NA]
		Is heritable within the normal population[2]
		Is more common in depressed persons with a strong family history of depression than those without a such a history[NA]
		Is more prevalent in families of depressed individuals than in families of nondepressed individuals[NA]
		Identifies those family members at risk for depression[NA]

Notes: Numerical superscripts refer to studies listed below. Positive numbers indicate that the study is consistent with the characteristic, and negative numbers indicate the study is inconsistent with the characteristic.
NA: None available. List of characteristics is after that of Iacono and Ficken (1989).
[1]Allen et al., 1993
[2]Allen, Reiner, Katsanis, and Iacono, 1997
[3]Davidson et al., 2000
[4]Debener et al., 2000
[5]Gotlib et al., 1998
[6]Heller et al., 1997
[7]Henriques and Davidson, 1990
[8]Henriques and Davidson, 1991
[9]Reid et al., 1998
[10]Rosenfeld, Baehr, Baehr, Gotlib, and Ranganath, 1996
[11]Schaffer et al., 1983
[12]Tomarken, Davidson, Wheeler, and Kinney, 1992
[13]Wiedemann et al., 1999

that to the extent that frontal EEG asymmetry may index risk for depression, it serves as a liability marker and not an episode marker or a prognostic sign.

Integrating the Present Findings with Previous Work

The present findings are consistent with much of the literature, including findings that frontal asymmetry is relatively stable in nonclinical populations (Hagemann et al., 2002; Jones et al., 1997; Tomarken, Davidson, Wheeler, & Kinney, 1992), and that previously depressed but euthymic depressed patients still differ from never-depressed controls (Allen et al., 1993; Gotlib et al., 1998; Henriques & Davidson, 1990). On the other hand, the present results contrast with those of Debener et al. (2000), in which considerable variability in asymmetry over time in depressed patients was observed. The most obvious difference between the study of Debener et al. (2000) and the present study is that patients in the present study received no medication, whereas those in the study of Debener et al. (2000) received a variety of antidepressant compounds and most additionally received benzodiazepines. Clearly a systematic investigation of the impact of antidepressant and antianxiety medications on frontal EEG asymmetry would be desirable, both in terms of understanding the extent to which they may alter asymmetry and also in terms of informing questions related to mechanism of action. An additional difference worth considering is that whereas the present sample included only women, the sample of Debener et al. (2000) included 5 men and 10 women. Although gender differences in the relationship of frontal asymmetry to factors putatively tapping risk have been reported (e.g. Kline, Allen, & Schwartz, 1998), the fact that the Debener et al. (2000) depressed sample included two-thirds women makes it less likely that differences in gender composition between the two studies are responsible for the discrepant findings.

Methodological Considerations

The present results involved three reference montages commonly used in frontal EEG asymmetry research. 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. Especially problematic is the Cz reference. The Cz reference may result in under- or overestimation of activity at the target site (Hagemann et al., 2001), and 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). Nonetheless, much of the literature has reported results with the Cz reference scheme (see Coan & Allen, 2003). The fact that many studies have successfully identified predicted relationships using the Cz references suggests at least two possibilities: (1) Significant results using the Cz reference to date derive, in part, not only from a 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) the Cz reference scheme introduces more error variance with respect to asymmetry per se, and may therefore result in a somewhat inconsistent pattern of empirical relationships with motivation, emotion, and

psychopathology. An evaluation of Cz to date suggests the latter, although differentiating these possibilities is possible only by reporting results with multiple reference schemes, as was done in the present study. 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 & Fiske, 1959).

Another finding to emerge from the present analyses concerns the number of minutes required to obtain reliable (in terms of internally consistent) measures of frontal asymmetry. Although the adage ''more is better'' clearly applies, the data depicted in Figure 1 demonstrate that 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 assessed with a sufficient number of constituent blocks. To highlight this point, consider a comparison of two comparable data points from Figure 1: four 60-s blocks or eight 30-s blocks, which correspond to identical timepoints from the EEG record. In all nine cases (3 sites $\times$ 3 reference schemes), the internal consistency of the latter is higher than the former, by an average of .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 less 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.

In terms of assessing EEG across time, greater stability was evident with longer recording blocks (Figure 3). The effect appeared incremental, with each longer recording period demonstrating descriptively higher reliability than the shorter recording period preceding it. Thus if one is interested in assessing long-term stability or change in EEG asymmetry, one would be advised to collect longer resting recording sessions as feasible, although compared to other sources of variance such as trait individual differences, the impact of recording length is very small, accounting for about 1% of overall variance.

Conclusion

The present study is consistent with many in the literature that imply stability of resting EEG alpha asymmetry over time in depressed patients. The present study, however, provides the only direct examination of stability in a nonmedicated population of depressed patients over time. These findings bolster the interpretation that resting anterior EEG alpha asymmetry may—for a subset of at-risk individuals—serve as a trait marker of risk for emotion-related psychopathology, including but not limited to depression, and one that is subject to occasion-specific fluctuations that are unrelated to clinical status.

REFERENCES

Allen, J. J., Iacono, W. G., Depue, R. A., & Arbisi, P. (1993). Regional electroencephalographic asymmetries in bipolar seasonal affective disorder before and after exposure to bright light. *Biological Psychiatry*, *33*, 642–646.

Allen, J. J. B., Coan, J. A., & Nazarian, M. (in press). What's the difference? Issues and assumptions in the use of difference scores and other metrics of anterior brain asymmetry in emotion. *Biological Psychology*.

Allen, J. J. B., Reiner, J., Katsanis, J., & Iacono, W. G. (1997). When it is and when it is not: The heritability of frontal EEG asymmetry. *Psychophysiology*, *34*, S17.

Allen, J. J. B., Schnyer, R. N., & Hitt, S. K. (1998). The efficacy of acupuncture in the treatment of major depression in women. *Psychological Science*, *9*, 397–401.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

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*, 89–92.

Beekman, A. T., Geerlings, S. W., Deeg, D. J., Smit, J. H., Schoevers, R. S., de Beurs, E., Beekman, A. T., Geerlings, S. W., Deeg, D. J., Smit, J. H., Schoevers, R. S., de Beurs, E., Braam, A. W., Penninx, B. W., & van Tilburg, W. (2002). The natural history of late-life depression: a 6-year prospective study in the community. *Archives of General Psychiatry*, *59*, 605–611.

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

Chapman, L. J., & Chapman, J. P. (1987). The measurement of handedness. *Brain & Cognition*, *6*, 175–183.

Chapman, J. P., Chapman, L. J., & Allen, J. J. (1987). The measurement of foot preference. *Neuropsychologia*, *25*, 579–584.

Coan, J. A. (2003). *The heritability of trait frontal EEG asymmetry and negative emotionality: Sex differences and genetic nonadditivity*. Doctoral dissertation, University of Arizona.

Coan, J. A., & Allen, J. J. B. (2003). The state and trait nature of frontal EEG asymmetry in emotion. In K. Hugdahl & R. J. Davidson (Eds.), *The asymmetrical brain* (2nd ed., pp. 565–615). Cambridge, MA: MIT Press.

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.

Davidson, R. J., & Fox, N. A. (1989). Frontal brain asymmetry predicts infants' response to maternal separation. *Journal of Abnormal Psychology*, *98*, 127–131.

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

Dawson, G., Frey, K., Panagiotides, H., Yamada, E., Hessl, D., & Osterling, J. (1999). Infants of depressed mothers exhibit atypical frontal electrical brain activity during interactions with mother and with a familiar, nondepressed adult. *Child Development*, *70*, 1058–1066.

Debener, S., Beauducel, A., Nessler, D., Brocke, B., Heilemann, H., & Kayser, J. (2000). Is resting anterior EEG alpha asymmetry a trait marker for depression? Findings for healthy adults and clinically depressed patients. *Neuropsychobiology*, *41*, 31–37.

Field, T., Fox, N., Pickens, J., & Nawrocki, T. (1995). Relative right frontal EEG activation in 3- to 6-month-old infants of ''depressed'' mothers. *Developmental Psychology*, *31*, 358–363.

Gallagher, S., Allen, J. J. B., Hitt, S. K., Schnyer, R. N., & Manber, R. (2001). Six-month depression relapse rates among women treated with acupuncture. *Complementary Therapies in Medicine*, *9*, 216–218.

Gelenberg, A. J., Wojcik, J. D., Falk, W. E., Baldessarini, R. J., Zeisel, S. H., Schoenfeld, D., & Mok, G. S. (1990). Tyrosine for depression: A double-blind trial. *Journal of Affective Disorders*, *19*, 98–103.

Gotlib, I. H., Ranganath, C., & Rosenfeld, J. P. (1998). Frontal EEG alpha asymmetry, depression, and cognitive functioning. *Cognition & Emotion*, *12*, 449–478.

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*, 619–641.

Heller, W., Nitschke, J. B., Etienne, M. A., & Miller, G. A. (1997). Patterns of regional brain activity differentiate types of anxiety. *Journal of Abnormal Psychology*, *106*, 376–385.

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

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

Iacono, W. G., & Ficken, J. W. (1989). Research strategies employing psychophysiological measures: Identifying and using psychophysiological markers. In G. Turpin (Ed.), *Handbook of clinical psychophysiology* (pp. 45–70). New York: John Wiley and Sons.

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

Kline, J. P., Allen, J. J. B., & Schwartz, G. E. (1998). Is left frontal brain activation in defensiveness gender specific? *Journal of Abnormal Psychology*, *107*, 149–153.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

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*, 389–404.

Rosenfeld, J. P., Baehr, E., Baehr, R., Gotlib, I. H., & Ranganath, C. (1996). Preliminary evidence that daily changes in frontal alpha asymmetry correlate with changes in affect in therapy sessions. *International Journal of Psychophysiology*, *23*, 137–141.

Schaffer, C. E., Davidson, R. J., & Saron, C. (1983). Frontal and parietal electroencephalogram asymmetry in depressed and nondepressed subjects. *Biological Psychiatry*, *18*, 753–762.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1990). *Structured clinical interview for DSM III R (SCID)*. Washington DC: American Psychiatric Press.

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

Tomarken, A. J., Davidson, R. J., Wheeler, R. E., & Doss, R. C. (1992). Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. *Journal of Personality & Social Psychology*, *62*, 676–687.

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*, 576–592.

Tomarken, A. J., Keener, A., & Neubauer, D. (1994). Long-term stability of frontal brain asymmetry: It helps to aggregate. *Psychophysiology*, *31*, S97.

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

Wiedemann, G., Pauli, P., Dengler, W., Lutzenberger, W., Birbaumer, N., & Buchkremer, G. (1999). Frontal brain asymmetry as a biological substrate of emotions in patients with panic disorders. *Archives of General Psychiatry*, *56*, 78–84.

(Received September 12, 2002; Accepted September 17, 2003)