Psychophysiology, 35 (1998), 607–614. Cambridge University Press. Printed in the USA.

Anterior electrophysiological asymmetries, emotion, and depression: Conceptual and methodological conundrums

RICHARD J. DAVIDSON

Laboratory for Affective Neuroscience, University of Wisconsin–Madison, USA

Abstract

Two reports in the last issue of this journal attempted to replicate aspects of our previous studies on anterior electroencephalogram (EEG) asymmetry, affective style, and depression. In this commentary, an overview is provided of our model of anterior asymmetries, affective style, and psychopathology. Emphasis is placed on conceptualizing the prefrontal and anterior temporal activation patterns within a circuit that includes cortical and subcortical structures. The causal status of individual differences in asymmetric activation in the production of affective style and psychopathology is considered. Major emphasis is placed on EEG methods, particularly the need for multiple assessments to obtain estimates of asymmetric activation that better reflect an individual's true score. Issues specific to each of the two articles are also considered. Each of the articles has more consistency with our previously published data than the authors themselves suggest. Recommendations are made for future research to resolve some of the outstanding issues.

Descriptors: Affect, Depression, EEG, Asymmetry

In the last issue of this journal were reports from two groups that have attempted to replicate aspects of our previous studies on relations between resting electroencephalographic (EEG) measures of anterior activation asymmetry and depression (Reid, Duke, & Allen, 1998) and reactivity to positive and negative emotion-eliciting stimuli (Hagemann, Naumann, Becker, Maier, & Bartussek, 1998). These attempted replications have met with mixed success. In each of the articles, the authors have thoughtfully considered variables that might account for the discrepancies between their reported results and those in the published literature. The goal for this commentary is to first introduce some general remarks about issues that are common to both articles. I briefly review and update our model of anterior activation asymmetries and their relation to affective style and psychopathology. I also comment on some of the prior literature, both studies mentioned in these articles and some studies that were omitted. Methodological issues raised in each of the articles that pertain to the use of EEG to assess patterns of stable trait-like individual differences in activation of circuitry associated with affective style is considered. I then address the specific operational tests that are represented in the two articles under consideration and note a number of potentially important differences between the design of the present studies and those from my laboratory that these studies putatively replicated. I also comment on the analytic strategies adopted by the authors in each case and note how they compare with the original work they were intended to replicate and to the theoretical model they were designed to evaluate. In each case, serious concerns about the analytic strategy are raised in relation to both conceptual and methodological issues in this area of research. Some recommendations are made for future research to help clarify some of the questions that have been raised.

This research was supported by NIMH grants MH43454, MH40747, Research Scientist Award K05-MH00875 and P50-MH52354 to the Wisconsin Center for Affective Science (R. J. Davidson, Director), by a NARSAD Established Investigator Award and by a grant from the John D. and Catherine T. MacArthur Foundation. I wish to thank Chris Larson, Daren Jackson, and Steve Sutton for their helpful comments on earlier drafts of this article.

Address reprint requests to: Richard J. Davidson, Laboratory for Affective Neuroscience, University of Wisconsin–Madison, 1202 West Johnson Street, Madison, WI 53706. E-mail: rjdavids@facstaff.wisc.edu.

Anterior Activation, Affective Style, and Psychopathology: Clarifications of the Model

Both of the articles under consideration examine relations between individual differences in a measure of anterior activation asymmetry and individual differences in aspects of emotional behavior. In one case, relations between anterior asymmetry and depression are examined and in the other case, relations between anterior asymmetry and reactivity to positive and negative emotional stimuli in normal subjects are studied. Whereas each of the articles is meant to replicate a specific prior study, each is also predicated on a model of anterior asymmetry and emotion (hereafter referred to as either "the model" or the AAE model) that I have developed with my colleagues over the past 15 years (Davidson, 1984, 1992, 1994, 1998; Davidson & Tomarken, 1989). The model has undergone important changes over this period of time and has been continuously revised in light of new data, both from my laboratory and the work of many other investigators. The details of the model most relevant to the issues under consideration here will be described. Readers interested in more detail can consult the relevant publications.

We have proposed (see Davidson, 1998, for most recent review), as have others (e.g., Gray, 1994; Lang, Bradley, & Cuthbert,

1990) that two major forms of motivation and emotion are represented in separate neural circuits. The approach system facilitates appetitive behavior and generates certain forms of positive affect. The withdrawal system facilitates the withdrawal of an organism from sources of aversive stimulation and generates certain forms of negative affect. In my recent review (Davidson, 1998), many brain regions were identified that constitute the circuitry underlying these emotional systems. The brain regions and structures that comprise these systems include the dorsolateral prefrontal cortex, the ventral/medial prefrontal cortex, the nucleus accumbens (particularly the shell region) and other parts of the basal ganglia, the amygdala, the anterior temporal cortex, the parietal cortex, and the hypothalamus. Some of these regions participate in both hypothesized circuits, whereas other structures have been implicated more in one system than in the other (e.g., the nucleus accumbens in the approach system; the amygdala in the withdrawal system). I have hypothesized that a lateralization for valence exists in some, but not all components of this circuitry, primarily in the anterior cortical regions that have been implicated (see Davidson, 1998). There are undoubtedly complex interactions among structures within a circuit with both extensive excitatory and inhibitory influences. Moreover, it is likely that emotional behavior can be generated without activation of all components of the circuit.

Whereas interactions undoubtedly occur between the approach and the withdrawal systems, these systems have been conceptualized as relatively orthogonal. In particular, we have continuously underscored the importance of examining activation levels in the separate right and left hemisphere components of these circuits, because some forms of psychopathology or affective style will likely involve bilateral differences in these systems. In the area of depression in particular, I have suggested that decreased activation in both the left and right prefrontal region is predicted to be associated with deficits in both the approach and withdrawal system, respectively, and as such, is likely to be associated with symptoms such as generalized blunted affect and anhedonia (see Davidson, 1994, 1998). Importantly, this view has critical methodological consequence, as it requires that the analytic strategy adopted permit the examination of the independent contributions of the two hemispheres to the behavior in question. Such methods were utilized in the two studies from my laboratory (Henriques & Davidson, 1991; Wheeler, Davidson, & Tomarken, 1993) that were the focus of the attempted replications, but were not used in the studies under question. More will be said about this issue below.

We have proposed that individual differences in some aspects of affective style are influenced by the activation levels of components of this circuitry. With particular regard to prefrontal activation asymmetry, we view individual differences in this component of the circuitry as a contributory cause of affective style. As a contributory cause, such differences in prefrontal activation are neither necessary nor sufficient for the production of a specific type of affective style or psychopathology. With regard to depression per se, we have unambiguously suggested that individual differences in prefrontal asymmetry are most appropriately viewed as diatheses that bias a person's affective style, and then in turn modulate an individual's vulnerability to develop depression. The proximal result of individual differences in this circuitry is variation in components of affective style, not depression or any other form of psychopathology per se. Our model was never intended as a model of depression or any other form of psychopathology for that matter. Moreover, it is likely that although we and others have uncovered relations between individual differences in prefrontal asymmetry and self-report measures that reflect components of affective style, self-report measures in the long run will prove to be inadequate to capture the core characteristics of affective style that are governed by the circuitry that I described above. This position is described in detail in a recent review (Davidson, 1998). We have used our model to help predict vulnerability to psychopathology because virtually all forms of major mental illness involve some abnormality of emotion. However, it is essential to keep in mind that traditional diagnostic categories are extremely heterogeneous and are not likely to map neatly onto brain circuitry. Moreover, there is considerable comorbidity among various disorders, particularly depression and anxiety, which share a core of common symptoms and genetic liability (Mineka, Watson, & Clark, 1998).

When an individual with left-sided prefrontal hypoactivation is exposed to negative life events over a prolonged period of time, we predict that there is an increase in the probability of developing depression. However, as a contributory cause, we would (1) not expect all subjects with relative right-sided anterior activation to be depressed; and (2) not expect all depressed subjects to show relative right-sided anterior activation, because we assume that there are multiple, complex routes to this disorder. We would predict, however, that depressed individuals would show abnormalities in at least some components of this circuitry.

Consistencies and Inconsistencies in Prior Literature

In each of the two articles under consideration, reference was made to studies in the literature that putatively bear on the validity of the model of anterior activation asymmetry and affective valence. Here I will comment briefly on those studies that have been interpreted to show failures to replicate the basic pattern of findings we have previously reported. Many of these issues have already been considered in previous reviews (Davidson, 1993; Davidson & Henriques, in press). First are the lesion data (e.g., House, Dennis, Warlow, Hawton, & Molyneux, 1990) that failed to replicate the earlier observations by Robinson and his colleagues (e.g., Robinson & Price, 1982; Robinson, Kubos, Starr, Rao, & Price, 1984; see review by Robinson & Downhill, 1995) of an increase in depression following left prefrontal lesions. As we have noted above and elsewhere (Davidson, 1993; Davidson & Henriques, in press), the diathesis-stress model acknowledges explicitly that patients with lesions in critical zones of the left hemisphere that have been implicated in the approach system circuitry will not necessarily become depressed. In other words, lesions in these locations are not sufficient for the production of depression. The lesions represent diatheses that will likely alter an individual's vulnerability to depression by affecting components of their affective style. As I have explained elsewhere (e.g., Davidson, 1993), the patient sample used by House et al. (1990) was different from that tested in the Robinson series. For example, fewer than half the patients in the House et al. (1990) study had been hospitalized as a result of their stroke whereas all of the patients examined by Robinson and Price (1982) were hospitalized. There were also major differences in the percentage of patients who lived alone. As was made clear in the report of another group (Astrom, Adolfsson, & Asplund, 1993), whether or not the patient is living alone and the number of social contacts are two key variables that predict the long-term course of depression in stroke patients. Whether or not the lesion is in the anterior left hemisphere was the single most important predictor of immediate poststroke depression according to this study. The basal ganglia are critical sites for poststroke depression and the effect appears to be strongly lateralized (see Robinson & Downhill, 1995, for review). As I noted above, the basal ganglia

are key component of emotional circuitry and left-sided regions within the basal ganglia are particularly important in the approach circuit. Lesions of this area have been found to increase the likelihood of major depression, including in a study that was listed by Reid and her colleagues as a failure to replicate the association between left anterior lesions and depression (Hermann, Bartels, Schumacher, & Wallesch, 1995). These latter authors reported that patients with lesions of the left hemisphere basal ganglia showed a significantly higher frequency of major depressive disorders and scored significantly higher on depression rating scales compared with patients with lesions in all other territories of the vascular supply.

As noted by Reid et al. (1998), a number of studies using both positron emission tomography (PET) and EEG methods have reported bilateral decreases in frontal activation in depressed patients compared with controls. As noted above, this pattern is not incompatible with our model. We would predict that such individuals have decreased activation in components of both the approach and withdrawal systems and should show correspondingly appropriate symptoms, including a relative lack of reactivity to both positive and negative incentives. As will be discussed in the final section of this commentary, studies that parse individuals on the basis of measures of brain function and then evaluate the consequences of these individual differences in patterns of brain function on objective measures of affective reactivity, are critically needed. Research that begins with traditional nosological categories from descriptive psychiatry is going to be ultimately fraught with problems because these categories are not likely to map cleanly onto brain circuitry.

Hagemann et al. (1998) review a number of EEG studies that they interpret as failing to replicate associations between either measures of resting anterior activation asymmetry and depression, or emotion-induced changes in asymmetric anterior activation. We have recently commented on many methodological issues in the EEG/depression research area (Davidson & Henriques, in press) and so will not repeat these issues here, though diagnostic heterogeneity and variability in the nature of the measures used are salient characteristics of this literature. Whereas most of the studies on quantitative EEG and depression were not explicitly designed to test the association between our model of anterior EEG asymmetry and depression, Hagemann et al. refer to several other EEG studies in normal subjects that were designed to be at least partial tests of the valence/asymmetry hypothesis. In these studies, emotion was experimentally induced and changes in anterior EEG asymmetries were examined. Collet and Duclaux (1987) used short film clips to induce happiness and sadness. After all of the clips had ended, they had subjects "mentally review" the clips they had been exposed to previously and the EEG was derived from a single 30-s period during the happy review period and compared with a single 30-s period during the sad review period (determined on the basis of facial electromyogram measures). They found a nonsignificant difference in the predicted direction in the mid-frontal scalp region (F3,4) on a measure of alpha power asymmetry using an average reference based on a total of nine electrodes (see below for discussion of the reference electrode issue). The effect for females was stronger though still not reliable. Although the data from this study do not support the basic hypothesis of a difference in frontal asymmetry during induced happiness and sadness, there are a sufficient number of differences between this study and that others we have performed in adults during which emotion has been elicited (e.g., Davidson, Ekman, Saron, Senulis, & Friesen, 1990) as to make comparison between them hazardous at best. Other methodological questions can raised about each of the other studies. For example, Schellberg, Besthorn, Klos, and Gasser (1990) examined EEG using a physically linked ears reference from a total of nine male subjects presented with short film clips. A small amount of data per subject per condition was used (under 45 s). A number of interactions with emotion condition, hemisphere and topography as factors were close to being significant although with the sample size used, did not reach significance. As we showed in our article on EEG changes induced in response to emotional films (Davidson et al., 1990b), only when specific epochs were extracted for analysis based upon the presence of objective facial signs of emotion did we find a reliable difference in anterior EEG asymmetry between positive and negative emotion conditions. Such procedures designed to objectively verify the presence of the intended emotion were not utilized in this study or related studies (e.g., Cole & Ray, 1985).

Finally, many published studies were not cited in either of the two articles under consideration, both from my laboratory and from several independent groups, that support our basic model. Among them include Sobotka, Davidson, and Senulis (1992), Drevets et al. (1997), Jones, Field, Fox, Lundy, and Davalos (1997), Field, Fox, Pickens, and Nawrocki (1995), Dawson, Klinger, Panagiotides, Hill, and Spieker (1992), and Tomarken and Keener (1998). In addition, we have performed several studies showing relations between individual differences in prefrontal asymmetry in rhesus monkeys and biological and behavioral indices that reflect affective style (Davidson, Kalin, & Shelton, 1993; Kalin, Larson, Shelton, & Davidson, 1998). These studies show that animals with greater relative left-sided activation have a more positive and less negative affective and biological profile.

General Issues Regarding EEG Methods

Each of the articles under consideration raised several issues of crucial importance for the use of EEG to make inferences about patterns of regional cortical activation. Although a detailed consideration of these issues is beyond the scope of this brief commentary (see Pivik et al., 1993; Davidson, Jackson, & Larson, in press, for detailed consideration of these issues), I will comment here on a few of the questions that were raised about EEG methodology because they bear directly on my more specific remarks about each article.

First is the issue of the reliability of EEG measures of asymmetry. With respect to reliability, data presented in each of the two reports under consideration indicate that measures of EEG asymmetry show good internal consistency reliability. However, in neither of these reports is test–retest stability examined. As we have demonstrated (Tomarken, Davidson, Wheeler, & Kinney, 1992), measures of frontal EEG asymmetry are only moderately stable over a short time interval (average r of approximately .6 over a three-week interval). Based upon the test–retest reliability estimates we have obtained, in all of the data we have collected since our 1992 study (Tomarken et al., 1992b), we have tested individuals on at least two occasions to examine stability and to obtain a more accurate estimate of an individual's true score on the asymmetry metric. In most of the work focused on individual differences in EEG measures of frontal asymmetry and their relation to affective style, we have utilized selection procedures to screen extreme and/or stable groups of subjects or we have used multisession measures to provide us with a better estimate of an individual's true score on an asymmetry measure. Thus, for example, in the Sutton and Davidson (1997) study that reported on relations between EEG asymmetry and scores on the Behavioral Activation and Inhibition scales (Carver & White, 1994), EEG was measured

on two separate occasions, separated by 6 weeks. Similarly, in studies examining relations between baseline measures of EEG asymmetry and cognitive performance, we have also tested subjects on two separate occasions (Davidson & Hugdahl, 1996). In our study of relations between individual differences in prefrontal EEG asymmetry and dispositional positive and negative affect (Tomarken, Davidson, Wheeler, & Doss, 1992), our major analytic strategy was to compare individuals who showed extreme and stable right-sided frontal activation with those showing extreme and stable left-sided frontal activation. However, in that study we also report on the correlations between measures of EEG asymmetry and dispositional positive and negative affect and we present the correlations for the group as a whole ($N = 72$), as well as subjects with stable frontal asymmetry between assessments (those whose Session 2 standardized asymmetry score was within one-third of a standard deviation of their Session 1 standardized score) and those with unstable frontal asymmetry between assessments. As we showed in that study, stronger relations between EEG asymmetry and measures of dispositional affect were found for the stable subjects compared with those whose asymmetry scores were less stable between assessments. These data clearly underscore the critical importance of adopting a multisession strategy in individual differences research that seeks to use EEG measures of asymmetry as traitlike indices.

With respect to the issue raised by Hagemann et al. (1998) regarding whether variations in alpha power can be used as a proxy for activation, relevant theory and data are extensively reviewed in Davidson, Jackson, & Larson (in press). Whereas definitive studies that examine relations between simultaneously acquired EEG and hemodynamic measures (PET or fMRI) have not yet been performed (but see Larson et al., 1998), other data suggest that variations in regional alpha power can be used in this way (see Davidson, Chapman, Chapman, & Henriques, 1990, for discussion).

A crucial issue raised in both articles is the choice of an appropriate reference. A commendable feature of each of these articles is the use of multiple references and the direct comparison among references, a strategy we have advocated and illustrated for many years (see e.g., Davidson et al., 1990a; Henriques & Davidson, 1990, 1991; Tomarken et al., 1992b). Both articles demonstrate that correlations between frontal EEG asymmetry measures derived with a Cz and a computer-averaged ears or mastoid reference are poor, whereas the Reid et al. article shows that correlations between an average reference and a computer-derived mastoid reference are good. In our recent work, we have consistently advocated for the use of either an average-ears or if a sufficiently large number of electrodes are available, an average reference. It is only in our earlier studies that we used a vertex reference because neither an average-ears nor an average reference was available in those early experiments (see footnote 2 in Tomarken et al., 1992a). Based upon theoretical grounds, an average reference is clearly the more appropriate (see Davidson, Jackson, & Larson, in press, for detailed discussion). As Hagemann et al. correctly note, the use of a Cz reference is potentially problematic because as an active site, variations in power at Cz can distort the magnitude and direction of asymmetry recorded from lateral sites.

Comments on Reid et al.

Of major importance in formulating a test of the AAE model as it pertains to depression is how to conceptualize posterior asymmetries. I have not taken a stand on this issue and have been favorably influenced by the work of Heller and colleagues (see Heller & Nitschke, 1998) on this question. According to Heller, individual differences in posterior activation asymmetry should vary with symptoms of anxious arousal, with those showing more anxiety having greater right-sided parietal activation (Heller, 1993). Thus, according to this view, individuals with symptoms of both depression and anxiety are likely to show right-sided activation in both frontal and parietal scalp regions. If an analysis of variance (ANOVA) is performed on metrics of activation asymmetry with group and site (frontal, parietal) as factors, we would specifically not predict an interaction if right-sided activation is expected to be present in both anterior and posterior scalp regions. The primary method used to test the hypothesis of differences in frontal EEG asymmetry between depressed and nondepressed subjects in this study was to perform an ANOVA on asymmetry scores (log right minus log left alpha power) with group and site (mid-frontal, lateral-frontal, and parietal) as factors. To establish support for the AAE model, Reid et al. required the Group $\times$ Site interaction to be significant. Moreover, when Group $\times$ Site ANOVAs were computed, they usually included three sites (mid-frontal, lateral-frontal, and parietal) and the conservative Greenhouse–Geisser correction was utilized rather than MANOVAs, which are considerably more powerful. As noted above, if the parietal region were to show right-sided activation along with the frontal region, the interaction would not be significant. It would have been more appropriate to test directly the group difference in frontal activation by computing a Group $\times$ Hemisphere ANOVA on the frontal alpha power measures, as we did in Henriques and Davidson (1991). This strategy also permits an assessment of the main effect for group across hemisphere to test whether bilateral differences in activation might differentiate between groups, as some have reported. Unfortunately none of the separate data for each hemisphere are presented in this article so the reader cannot evaluate whether any group differences might have been presented bilaterally.

In their Study II with the sample of clinically depressed patients, Reid et al. reported that for both the computer-derived linked mastoid (LM) and the average (AR) references, the Group $\times$ Site interaction was indeed significant. They attribute this interaction exclusively to the group difference in parietal asymmetry for the LM reference, although their Figure 2 clearly indicates greater relative right-sided frontal activation in F7/8 using both the LM and the average references. It seems clear that with either a less conservative, more direct test of the group difference in frontal asymmetry, or with a slightly larger sample size, the group difference would indeed be significant. When Reid et al. analyzed the data from Study II separately for each 2-min block of data, they did find significant differences in lateral frontal asymmetry in the predicted direction for the first block of data. It is noteworthy that the asymmetry scores for the depressed group were lower than those for the control group at each of the four time points, indicating consistently more relative right-sided frontal activation throughout the 8-min recording period.

The initial study presented in this article using a subclinical depressed student sample differed in one important respect from our study (Schaffer, Davidson, & Saron, 1983) using subclinically depressed students. In our study, we used the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) for screening subjects in a mass testing, as Reid et al. did. However, we imposed a dual criterion for selection into the depressed group. We required that at initial screening depressed subjects had to have a BDI score of 20 or above and just prior to electrophysiological testing, they were required to have a score of 14 or above. The mean BDI score of our depressed sample at initial testing was

29.7 and at the time of EEG assessment it was 25.7. Thus, our depressed sample had to remain consistently depressed for inclusion in the study. Forty-one percent of the sample who met the criterion for the first BDI assessment failed to meet the criterion for the second assessment approximately 6 weeks later. It is also the case that the subclinically depressed sample in the Schaffer et al. study had somewhat higher BDI scores than the sample tested by Reid et al.

In their Study II, Reid et al. compared DSM-III-R diagnosed depressed patients to normal controls. An important difference between our study of clinical depression (Henriques & Davidson, 1991) and that of Reid et al. is in nature of the control group. Our control group was required to have an absence of lifetime psychopathology in themselves and their first-degree relatives while no such requirement was imposed in the Reid et al. study.

There are a number of important issues that Reid et al. address in their thoughtful discussion that merit some comment here. First, it deserves emphasis that there were significant group differences in the lateral prefrontal region with depressed subjects showing greater relative right-sided activation during the first 2 min of the recording session. Second, it is also noteworthy that significant group differences were found in the anterior temporal region, again with depressed subjects showing greater relative right-sided activation (across the entire 8-min baseline recording period) compared with controls.

Reid et al. note in their discussion that in our study of clinically depressed patients (Henriques & Davidson, 1991), one of the patients had an extreme asymmetry value for one of the reference montages that was examined (average reference data). We agree that this data point was indeed extreme. We have gone back and re-examined the raw EEG record to ascertain whether it contained any artifact that might have been missed during the original inspection but were not able to find any. However, in light of the fact that this data point was indeed extreme, we reanalyzed the data without this subject, testing the Group $\times$ Hemisphere interaction as we had done in the original article. For the average reference, the interaction remains significant, $F(1,25) = 6.23$, $p < .02$, and for the Cz reference, the interaction also remains significant, $F(1,25) = 5.44$, $p < .03$, indicating that the depressed subjects still showed significantly greater relative right-sided activation even after removal of this one depressed subject who showed the most extreme right-sided data point.

Reid et al. remind readers that some of the depressed patients in Henriques and Davidson (1991) were on medication at the time of testing. In the original article, we compared specifically between depressed patients on versus off antidepressant medication (see Table 4 in Henriques & Davidson, 1991) and found that no difference existed. However, this issue requires more careful study in a longitudinal design following patients before and during and ideally after a course of antidepressant medication to separate effects that might be medication-dependent from those that are associated with changes in clinical status.

Comments on Hagemann et al.

The principal purpose of the experiment reported by these authors was to replicate and extend the findings we reported in Wheeler et al. (1993) showing that subjects who differ in measures of resting prefrontal activation asymmetry respond differently to positive and negative emotional stimuli with subjects displaying greater relative left-sided activation reporting more positive and less negative affect in response to film clips designed to elicit these classes of emotion. Hagemann et al. reported on a power analysis they performed to ascertain whether their sample size of 37 was sufficient to test their hypothesis adequately and used data from both Tomarken, Davidson, and Henriques (1990) and Wheeler et al. (1993) to compute an effect size. However, there is a serious error in the strategy they adopted to compute their effect size and in the logic they followed in designing their study. In the Wheeler et al. (1993) study, we assessed EEG on two separate occasions separated by 3 weeks. The effects we reported on relations between baseline prefrontal asymmetry and reactivity to emotion elicitors emerged *only for those subjects whose frontal EEG asymmetry remained stable across the two assessments*. We noted explicitly in this article that "no significant relations were found between midfrontal asymmetry (for Session 1 or Session 2 or across Sessions 1 and 2) and the aggregate rating measures for the total sample (all ps $> .47$)…" (p. 85). In fact, the highest correlation between midfrontal asymmetry and the aggregate rating measures for the entire sample was $r = .045$. If we combine this r-value with the r-value from the Tomarken et al. (1990) study that was performed on an unselected sample (using Fischer's r-to-Z transformation) as Hagemann et al. did, we arrive at an aggregate effect size of $r = .22$. To detect an effect of this magnitude at a power level of .8 using a .05, one-tailed test, an N of 140 is required, a sample size more than three and one-half times that used in the study under consideration.

Unlike our prior studies assessing this topic, Hagemann et al. used pictures from the IAPS series (Lang, Ohman, & Vaitl, 1988) to elicit positive and negative emotion. We too have used these pictures in both psychophysiological (e.g., Sutton, Davidson, Donzella, Irwin, & Dottl, 1997) and neuroimaging (e.g., Irwin et al., 1996) research. However, in the study under consideration, Hagemann et al. had subjects rate each stimulus immediately after it was presented. It is not clear whether this heavy demand for subjects to rate the stimuli might have interfered with their affective reactions to the pictures. The means presented by Hagemann et al. on a 0–9 point scale of intensity for positive and negative affect were lower than means we reported in Tomarken et al. (1990) on a 0–8 point scale. Thus, the true difference is likely to be even larger. Potentially even more important is the fact that the variability in our ratings was consistently greater for all measures than the variability reported by Hagemann et al. For studies of individual differences, it is imperative to maximize variability in responding to uncover meaningful relations between the EEG and rating measures.

A noteworthy feature of the Hagemann et al. study is the thorough evaluation of different reference placements for the EEG, and varying lengths of EEG recording time included in the different analyses. This evaluation enabled these authors to determine if there was a particularly optimal combination of methods that produced the strongest replication of the original finding. However, it must be emphasized that with a sample size more than double the sample size used by Hagemann et al., we reported that when the data from the unselected sample was examined, no significant relations between EEG baseline measures of frontal asymmetry and any of the affect rating measures were found. The fact that Hagemann et al. also found no such relation using their unselected sample is absolutely consistent with our prior data.

However, because Hagemann et al. did use five different databases for their EEG measures (30 s of eyes open; 30 s of eyes closed; 4 min of eyes open; 4 min of eyes closed; 8 min of combined eyes open and closed) and four different measures of global positive affect and global negative affect, each with two different reference methods (Cz and computer-averaged mastoids), a total of

40 separate correlations with the positive affect variables and 40 correlations with the negative affect variables were measured. Our model would predict that the relation between EEG frontal asymmetry and the positive affect variables would be positive, because greater relative left-sided frontal activation is reflected in higher asymmetry scores. Of the 40 correlations Hagemann et al. presented in their Table 6, 38 are positive. If EEG asymmetry were unrelated to the positive affect variables, we would expect half of these correlations to be positive and half negative. A binomial test reveals that the likelihood of obtaining 38 of 40 correlations of the same sign by chance is $z = 5.69$, $p < .00001$. Likewise, we would predict that the correlation between frontal EEG asymmetry and negative affect variables would be negative. According to their Table 7, 30 of the 40 correlations were positive (or zero). Again, this pattern is highly statistically significant by the binomial test, $z = 3.16$, $p < .0005$. These analyses indicate that although only a few of the individual effects they reported reached statistical significance, the overall pattern of data was highly consistent with our model.

Hagemann et al. note in their discussion that of all the procedural combinations used in their study, EEG aggregated across the 8 min of baseline trials referenced to the computer-averaged mastoids and the use of the affective style index that removes that portion of variance that is due to a global rating bias "can be considered optimal to assess the relation between frontal asymmetry and affective style." We would add further the importance of also statistically removing the contributions of baseline mood as we did in Wheeler et al. and as Hagemann et al. did in their study, too. As we demonstrated in Wheeler et al., the strongest relations between the asymmetry measure and the affect variables was with the measure of affective bias. The correlation between the positive affect variable and frontal asymmetry was .31 for the Cz reference and .25 for the averaged-ears reference. For negative affect, the comparable correlations were −.13 and −.06. Hagemann et al. did not report the correlations with the affective bias variable for this procedural combination of EEG and rating measure, but these values will necessarily be higher than the highest correlations for positive affect because the negative affect correlations are in the opposite direction. Thus, using the procedural combination that most closely resembles that used in Wheeler et al., despite having assessed EEG on only one occasion and not examining only those subjects who demonstrated reliable frontal asymmetry across assessment, Hagemann et al. reported effects that we interpret as highly consistent with our original report. Finally, it should also be noted that in the Wheeler et al. study we showcased a method for examining relations between power at individual left and right hemisphere electrode sites and measures of affective reactivity. The measure we developed was a residualized power measure that removed the contributions of extraneous nonspecific influences that presumably reflect anatomical variables such as skull thickness (see Pivik et al., 1993; Davidson, Jackson, & Larson, in press, for detailed descriptions of the method). Using this measure, we reported correlations for the stable asymmetry subjects that reached .62. Unfortunately, this method was not used in the Hagemann et al. report.

One final issue raised by Hagemann et al. deserves comment. They reported that relations between measures of anterior temporal asymmetry and the affect rating indices were opposite to prediction, showing that individuals with greater relative right-sided anterior temporal activation reporting more positive affect. This finding represents an important methodological challenge. We have not consistently observed relations between affect measures and anterior temporal asymmetry (see Davidson, 1998, for review). The anterior temporal electrode sites (T3 and T4) are in very close geographic proximity to both the mastoid reference and the Cz reference. As such, recordings from T3 and T4 with either reference represent an essentially bipolar recording, with closely spaced electrodes. Thus activity that is in phase at the two electrode sites will not be amplified because of common mode rejection. We believe that for these electrode sites in particular, it is important to use an average reference. With either a Cz or a mastoid or ear reference, recordings from the anterior temporal region will be more bipolarlike than recordings from elsewhere on the scalp. Finally, Reid et al. do report significant differences between depressed and control subjects in the anterior temporal region in the predicted direction using both a computer-averaged ears reference (likely to be better for this purpose than a mastoid reference, which is closer to T3/4) and the average reference. Interestingly, there was no difference in T3/4 asymmetry between depressed and control subjects for the Cz-referenced data.

Summary and Conclusions

Science is a collective affair that benefits from the diverse and multiple inputs from the community of scholars. Replication is one of the most important components of scientific progress and it is with much appreciation that I view the efforts of the investigators whose studies are considered herein. Each of these reports is thoughtful and raises many important conceptual and methodological questions that should sharpen future research endeavors on these problems. In each of these studies, some effects that the investigators predicted on the basis of our prior work were replicated and some were not. In the Reid et al. report, they demonstrated that with clinically depressed subjects, the first 2 min of baseline recording revealed the predicted difference in EEG measures of frontal asymmetry. Moreover, recordings from the anterior temporal region showed that depressed subjects had greater relative right-sided activation than nondepressed subjects for measures aggregated over the entire recording period. Hagemann et al. found that correlations between baseline frontal asymmetry measures from an unselected sample and measures of affective reactivity to emotional pictures were mostly in the direction predicted by our prior research, though few of the effects reached independent statistical significance. As I showed above, however, the overall pattern of associations they reported was highly statistically significantly different from chance in the direction predicted by our prior work. In this commentary, I focused on methodological and conceptual limitations of this work and differences, in some cases significant differences, between methods used in the studies under consideration and those adopted by myself and my colleagues in the published research that these studies were designed to replicate. In the Hagemann et al. case, I do not view their data as a failure to replicate. Based upon the fact that they used an unselected sample whereas we found relations between frontal asymmetry and affective reactivity only for individuals who demonstrated stable frontal asymmetry across a 3-week time interval and not for the sample as a whole, their findings are actually consistent with ours. The Reid et al. study used methods that differed in other ways from the ones we used in our depression research, though we would not have expected such small differences to have such effects. Reid et al. have performed a valuable service by calling attention to the several methodological issues that require attention in future research.

One of the strategies that we have used in our research is to parse individuals on the basis of differences in prefrontal activation

asymmetry, and on other biological parameters theoretically related to affective style (e.g., activation of the amygdala) and then examine how the individuals might differ behaviorally and biologically. This strategy has proven to be extremely effective in uncovering relations between frontal asymmetry and immune function (e.g., Kang et al., 1991), basal cortisol levels (Kalin et al., 1998), child temperament (Davidson, 1994), and self-report measures of affect and personality (Tomarken & Davidson, 1994; Tomarken et al., 1992a). Diagnostic categories from the descriptive nosology of psychiatry and personality dimensions from personality theory were derived, for the most part, without reference to the nervous system and do not therefore necessarily honor the distinctions among major circuits in the brain. Moreover, as I have suggested above, in the case of psychiatric diagnostic categories, considerable comorbidity and diagnostic heterogeneity is the rule rather than the exception. For these reasons, it is imperative that we move beyond the traditional nosology we have inherited and instead obtain objective measures of affective reactivity and affective chronometry (see Davidson, 1998). In this way, we can begin to parse specific parameters of individual differences in basic features of affective style such as latency to the peak of the response, latency to recover, and so forth. These parameters of affective responding can be linked theoretically to specific components of the circuitry for emotion and its regulation, making the search for underlying neural substrates a much more tractable problem. Given the extraordinary heterogeneity of depression it is actually remarkable that any consistency has been observed in measures of baseline regional brain activation. It is imperative that we move beyond the phenomenological domain and begin to parse the components of affective reactivity using modern objective measures. In the study of cognitive deficits in psychopathology, it would now be unthinkable to relate patterns of baseline brain function to self-report or clinician ratings of thought disorder. Rather, psychopathologists have excelled in their use of laboratory tasks designed to systematically decompose cognitive processes into more elementary mental operations that can then yield to a search for underlying neural substrates.

In the study of psychophysiological concomitants of individual differences and psychopathology, we often unwittingly assume that because we are using biological measures, they are necessarily reliable. Both articles under consideration here have devoted more attention than is typical in psychophysiological studies to the issue of reliability of measures. Reliability is particularly crucial for studies of individual differences assumed to have trait status because traits are assumed to reflect homogenous characteristics that are stable over time. We first reported on the computation of internal consistency reliability for metrics of EEG asymmetry (Tomarken et al., 1992b) and demonstrated that such indices were indeed reliable when several minutes of data were taken into account. Both Hagemann et al. and Reid et al. have strongly replicated our prior finding. However, in the same article we demonstrated that the test–retest reliability of frontal EEG asymmetry measures was only moderate and in many articles since the reliability article appeared, we have repeatedly illustrated the importance of collecting baseline data on more than one occasion so that metrics of asymmetry can be computed that are more accurate estimates of a subject's true score. Although our earlier work did not use multiple sessions, most of the individual difference studies were based on the use of specially selected extreme groups (e.g., based on depression or temperament measures). A significant limitation of the studies under consideration is that fact that they used only a single session to compute estimates of anterior EEG asymmetry. It is imperative that in future research using these methods at least two sessions (and ideally more than two) of baseline EEG data collection be included.

Finally, as I noted in the first section of this commentary, the circuitry underlying appetitive and aversive emotion and motivation in the human brain is complex and involves many interconnected structures. The prefrontal cortex is clearly a significant component of the circuitry but it is important that we not view it as the "location" of these complex processes. Rather, it appears that these affective systems are implemented in circuits, only part of which are cortical. If we are serious about pursuing our understanding of the underlying neural substrates of these individual differences, it is essential that we complement our EEG methods with other tomographic neuroimaging methods that allow for the examination of activity in subcortical structures with which the prefrontal cortex is interconnected. As I suggested earlier in this commentary, some of the heterogeneity in affective style and psychopathology may relate to where in the circuitry individual differences may reside. For example, I have suggested that individual differences in prefrontal function may be particularly associated with variations in the time course of emotional responding or affective chronometry, whereas individual differences in amygdala function may account for differences in the behavioral and/or autonomic signs of negative affect (see Davidson, 1998 for more detailed discussion). Self-report and interview-based measures are unlikely to be sufficiently sensitive to reflect these differences in subcomponents of emotional reactivity, requiring that we examine relations between our measures of brain function and objective, laboratory-based measures of affective reactivity. The use of these procedures together in the same subjects is likely to yield data that truly advance our understanding of affective neuroscience.

REFERENCES

Astrom, M., Adolfsson, R., & Asplund, K. (1993). Major depression in stroke patients: A 3-year longitudinal study. *Stroke*, *24*, 976–982.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561–571.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319–333.

Cole, H. W., & Ray, W. J. (1985). EEG correlates of emotional tasks related to attentional demands. *International Journal of Psychophysiology*, *3*, 33–41.

Collet, L., & Duclaux, R. (1987). Hemispheric lateralization of emotions: Absence of electrophysiological arguments. *Physiological Behavior*, *40*, 215–220.

Davidson, R. J. (1984). Hemispheric asymmetry and emotion. In K. Scherer and P. Ekman (Eds.), *Approaches to emotion* (pp. 39–57). Hillsdale, NJ: Erlbaum.

Davidson, R. J. (1992). Emotion and affective style: Hemispheric substrates. *Psychological Science*, *3*, 39–43.

Davidson, R. J. (1993). Cerebral asymmetry and emotion: Conceptual and methodological conundrums. *Cognition and Emotion*, *7*, 115–138.

Davidson, R. J. (1994). Asymmetric brain function, affective style and psychopathology: The role of early experience and plasticity. *Development and Psychopathology*, *6*, 741–758.

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

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

Davidson, R. J., Ekman, P., Saron, C., Senulis, J. A., & Friesen, W. V.

(1990b). Approach–withdrawal and cerebral asymmetry: Emotional expression and brain physiology 1. *Journal of Personality and Social Psychology, 58,* 330–341.

Davidson, R. J., & Henriques, J. B. (in press). Regional brain function in sadness and depression. In J. Borod (Ed.), *The neuropsychology of emotion*. New York: Oxford University Press.

Davidson, R. J., & Hugdahl, K. (1996). Baseline asymmetries in brain electrical activity predict dichotic listening performance. *Neuropsychology, 10,* 241–246.

Davidson, R. J., Jackson, D., & Larson, C. (in press). Human electroencephalography. In J. T. Cacioppo, L. Tassinary, & G. Berntson (Eds.), *Principles of psychophysiology* (2nd ed.). New York: Cambridge University Press.

Davidson, R. J., Kalin, N. H., & Shelton, S. E. (1993). Lateralized response to diazepam predicts temperamental style in rhesus monkeys. *Behavioral Neuroscience, 107,* 1106–1110.

Davidson, R. J., & Tomarken, A. J. (1989). Laterality and emotion: An electrophysiological approach. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (pp. 419–441). Amsterdam: Elsevier.

Dawson, G., Klinger, L. G., Panagiotides, H., Hill, D., & Spieker, S. (1992). Frontal lobe activity and affective behavior of infants of mothers with depressive symptoms. *Child Development, 63,* 725–737.

Drevets, W. C., Price, J. L., Simpson, J. R., Todd, R. D., Reich, T., Vannier, M., & Raichle, M. E. (1997). Subgenual prefrontal cortex abnormalities in mood disorders. *Nature, 386,* 824–827.

Field, T., Fox, N. A., Pickens, J., & Nawrocki, T. (1995). Relative right frontal EEG activation in 3- to 6-month-old infants of "depressed" mothers. *Developmental Psychology, 3,* 358–363.

Gray, J. A. (1994). Three fundamental emotion systems. In P. Ekman, & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 243–247). New York: Oxford University Press.

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

Heller, W. (1993). Neuropsychological mechanisms of individual differences in emotion, personality, and arousal. *Neuropsychology, 7,* 476–489.

Heller, W., & Nitschke, J. B. (1998). The puzzle of regional brain activity in depression nd anxiety: The importance of subtypes and comorbidity. *Cognition and Emotion, 12,* 421–447.

Henriques, J. B., & Davidson, R. J. (1990). Regional brain electrical asymmetries discriminate between previously depressed subjects and healthy controls. *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.

Herrmann, M., Bartels, C., Schumacher, M., & Wallesch, C. W. (1995). Poststroke depression: Is there a pathoanatomic correlate for depression in the postacute stage of stroke? *Stroke, 26,* 850–856.

House, A., Dennis, M., Warlow, C., Hawton, K., & Molyneux, A. (1990). Mood disorders after stroke and their relation to lesion location: A CT scan study. *Brain, 113* (Pt 4), 1113–1129.

Irwin, W., Davidson, R. J., Lowe, M. J., Mock, B. J., Sorenson, J. A., & Turski, P. A. (1996). Human amygdala activation detected with echoplanar functional magnetic resonance imaging. *NeuroReport, 7,* 1765–1769.

Jones, N. A., Field, T., Fox, N. A., Lundy, B., & Davalos, M. (1997). EEG activation in 1-month-old infants of depressed mothers. *Developmental Psychopathology, 9,* 491–505.

Kalin, N. H., Larson, C., Shelton, S. E., & Davidson, R. J. (1998). Asymmetric frontal brain activity, cortisol, and behavior associated with fearful temperament in Rhesus monkeys. *Behavioral Neuroscience, 112,* 286–292.

Kang, D. H., Davidson, R. J., Coe, C. L., Wheeler, R. W., Tomarken, A. J., & Ershler, W. B. (1991). Frontal brain asymmetry and immune function. *Behavioral Neuroscience, 105,* 860–869.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention and the startle reflex. *Psychological Review, 97,* 377–398.

Lang, P. J., Ohman, A., & Vaitl, D. (1988). Thc international affective picture system [photographic slides]. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

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

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49,* 377–412.

Pivik, T., Broughton, R., Coppola, R., Davidson, R. J., Fox, N. A., & Nuwer, R. (1993). Guidelines for quantitative electroencephalography in research contexts. *Psychophysiology, 30,* 547–558.

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.

Robinson, R. G., & Downhill J. E. (1995). Lateralization of psychopathology in response to focal brain injury. In R. J. Davidson, & K. Hugdahl (Eds.), *Brain asymmetry* (pp. 693–711). Cambridge, MA: MIT Press.

Robinson, R. G., Kubos, K. L., Starr, L. B., Rao, K., & Price, T. R. (1984). Mood disorders in stroke patients: Importance of location of lesion. *Brain, 107,* 81–93.

Robinson, R. G., & Price, T. R. (1982). Post-stroke depressive disorders: A follow-up of 103 patients. *Stroke, 13,* 635–641.

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

Schellberg, D., Besthorn, C., Klos, T., & Gasser, T. (1990). EEG power and coherence while male adults watch emotional video films. *International Journal of Psychophysiology, 3,* 279–291.

Sobotka, S., Davidson, R. J., & Senulis, J. A. (1992). Anterior brain electrical asymmetries in response to reward and punishment. *Electroencephalography and Clinical Neurophysiology, 83,* 236–247.

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.

Sutton, S. K., Davidson, R. J., Donzella, B., Irwin, W., & Dottl, D. A. (1997). Manipulating affective state using extended picture presentation. *Psychophysiology, 34,* 217–226.

Tomarken, A. J., & Davidson, R. J. (1994). Frontal brain activation in repressors and non-repressors. *Journal of Abnormal Psychology, 103,* 339–349.

Tomarken, A. J., Davidson, R. J., & Henriques, J. B. (1990). Resting frontal brain asymmetry predicts affective responses to films. *Journal of Personality and Social Psychology, 59,* 791–801.

Tomarken, A. J., & Keener, A. D. (1998). Frontal brain asymmetry and depression: A self-regulatory perspective. *Cognition and Emotion, 12,* 387–420.

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

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

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.

(Received Month 00, 1998; Accepted Month 00, 1998)