Applied Psychophysiology and Biofeedback, Vol. 29, No. 4, December 2004 (© 2004)
DOI: 10.1007/s10484-004-0383-4

The Effectiveness of Neurofeedback and Stimulant Drugs in Treating AD/HD: Part II. Replication

Thomas Rossiter[1]

*This study replicated T. R. Rossiter and T. J. La Vaque (1995) with a larger sample, expanded age range, and improved statistical analysis. Thirty-one AD/HD patients who chose stimulant drug (MED) treatment were matched with 31 patients who chose a neurofeedback (EEG) treatment program. EEG patients received either office (*n = *14) or home (*n = *17) neurofeedback. Stimulants for MED patients were titrated using the Test of Variables of Attention (TOVA). EEG (effect size* [ES] = *1.01–1.71) and MED (*ES = *0.80–1.80) groups showed statistically and clinically significant improvement on TOVA measures of attention, impulse control, processing speed, and variability in attention. The EEG group demonstrated statistically and clinically significant improvement on behavioral measures (Behavior Assessment System for Children,* ES = *1.16–1.78, and Brown Attention Deficit Disorder Scales,* ES = *1.59). TOVA gain scores for the EEG and MED groups were not significantly different. More importantly, confidence interval and nonequivalence null hypothesis testing confirmed that the neurofeedback program produced patient outcomes equivalent to those obtained with stimulant drugs. An effectiveness research design places some limitations on the conclusions that can be drawn.*

KEY WORDS: neurofeedback; EEG biofeedback; AD/HD; stimulant drugs; active treatment control.

Attention-Deficit/Hyperactivity Disorder (AD/HD) is a behavioral disorder defined by inattention, hyperactivity, and/or impulsivity. Diagnosis is complicated by the fact that none of the core symptoms are exclusive to AD/HD and the majority of AD/HD patients suffer from at least one additional psychiatric disorder. AD/HD was originally thought to be limited to children and adolescents. However it is now recognized that in the majority of cases, AD/HD persists into the adult years. Stimulant drugs have been the treatment of choice for AD/HD for more than three decades. The use of neurofeedback (EEG biofeedback) to treat AD/HD dates from the 1970s (Lubar & Shouse, 1976). Nevertheless, it was not until the 1990s that neurofeedback became widely available as an alternative to stimulant drugs. To date, acceptance by the scientific community has been hindered by the paucity of well-designed outcome studies.

The purpose of this study was to compare the effectiveness of neurofeedback to that of stimulant medication in persons suffering from AD/HD. It was designed to replicate

[1] 1775 Highview Street, De Pere, Wisconsin 54115; e-mail: t.rossiter@worldnet.att.net.

Rossiter and La Vaque (1995) with a larger sample, an expanded age range including adults, more comprehensive collection of behavioral data for the neurofeedback group, and improved statistical analysis.

Rossiter and La Vaque (1995) as well as Monastra, Monastra, and George (2002) and Fuchs, Birbaumer, Lutzenberger, Gruzelier, and Kaiser (2003) previously reported positive effects of neurofeedback in treating AD/HD. This set of studies is unique in that they used an effectiveness research model (Kazdin, 2003) with nonrandom assignment of patients and an active treatment control. That is, patients chose their treatment(s) and the effectiveness of neurofeedback was evaluated by comparing it to stimulant drug therapy, a proven treatment for AD/HD. Rossiter (2004) reviews recent criticisms of neurofeedback outcome research and the relevant research design issues.

It was predicted that the neurofeedback (EEG) and stimulant medication (MED) groups will demonstrate statistically and clinically significant improvement on the Test Of Variables of Attention (TOVA; Leark, Dupuy, Greenberg, Corman, & Kindschi, 1996); the EEG group will show statistically and clinically significant improvement on the Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992) and the Brown Attention Deficit Disorder Scales (Brown, 1996); the null hypothesis of no statistically significant differences between the EEG and MED groups on TOVA gain scores will not be rejected; the proportion of EEG patients that demonstrate significant improvement will be equivalent, or noninferior, to that of the MED group (Rossiter, 2004).

METHOD

Participants

Participants were 62 patients seen by the author on a fee for service basis in his outpatient practice in Green Bay, WI. They were White, predominantly middle class individuals whose health care needs were paid for by private health insurance and/or by the patient or parents in the case of minors. The patients were evaluated by the author and received a primary *DSM-IV* (American Psychiatric Association, 1994) diagnosis of Attention-Deficit/Hyperactivity Disorder Combined Type or Predominantly Inattentive Type (Table I). Patients with secondary psychiatric diagnoses were included.

Two groups of 31 patients each were formed. The first group was drawn from 33 consecutive patients treated by the author with neurofeedback (EEG). Patients currently treated with stimulant medications were included in the EEG sample. Two patients treated with antidepressant and/or antihypertensive medication were excluded. EEG patients received no collateral treatment except for the six patients being treated with stimulant drugs concurrent with neurofeedback training. Eight patients had been treated with stimulant drugs in the past but terminated drug therapy 6 months or more before starting neurofeedback. In three cases, the decision to terminate stimulants was related to ineffectiveness of the medication and/or unacceptable side effects.

The second group consisted of patients who chose treatment with stimulants (methyl- or dextroamphetamine). The medication group (MED) was drawn from a pool of 64 patients whose stimulant medications had been titrated using the TOVA. They were matched with the EEG group first by age and then, to the extent possible, by the sum of the four baseline TOVA scores, IQ, gender, and primary AD/HD diagnosis in that order. Maturational changes in the

Table I. EEG and MED Group Demographic Variables

	EEG	MED
Age		
Mean (*SD*)	16.6 (12.7)	16.7 (12.5)
Range	7–55	7–52
Gender		
Male	21	22
Female	10	9
Intelligence		
Mean (*SD*)	107.5 (14.1)	104.6 (10.8)
Range	80–139	89–130
Primary diagnosis		
ADHD combined	16	20
ADHD inattentive	15	11
Secondary diagnosis	17	14

Note. $n = 31$ for EEG and MED groups.

TOVA pattern for AD/HD dictated that EEG and MED patients be matched by age. Because of the relatively small and heterogeneous patient groups, it was not possible to match EEG and MED patients on each of the four TOVA outcome variables. As an alternative, same age EEG and MED patients were matched on the sum of the standard scores of the four TOVA measures. Matched EEG–MED pairs were comparable on the overall extent of their deficits even though the specific pattern of deficits might differ.

The options of neurofeedback and/or stimulant medication were discussed with all patients. The patient and/or their parents chose the treatment. Among the neurofeedback patients, unwillingness to take medication was the overriding reason cited for seeking an alternative treatment. Patients choosing stimulant drug therapy seldom offered a reason for their choices. However, low out-of-pocket cost, proven effectiveness, and convenience appear to have been factors.

Instruments

Intelligence estimates were obtained using the Kaufman Brief Intelligence Test (Kaufman & Kaufman, 1990) or Wechsler Scale. Intelligence estimates were used to interpret TOVA results but were not dependent variables.

Test of Variables of Attention (TOVA) scores (errors of omission, errors of commission, response time, variability of response time) were the dependent variables used to compare EEG and MED outcomes. The TOVA is a continuous performance test (CPT) that is computer administered and scored. This reduces the likelihood of human bias with respect to administration and scoring. The TOVA avoids some of the potential difficulties inherent in relying on subjective parent, teacher, and patient reports as the primary basis for diagnosing AD/HD and assessing treatment effects. Riccio, Reynolds, and Lowe (2001) reviewed the research literature on CPTs including the TOVA. They concluded that CPTs (1) have high levels of sensitivity and specificity when differentiating AD/HD from normal individuals; (2) objectively evaluate symptoms associated with disorders of self-regulation, particularly impulsivity and attention problems; (3) are sensitive to the effect of stimulants on attention, processing speed, and executive control; and (4) have moderate to high ecological validity.

The Behavior Assessment System for Children (BASC) assesses a range of child psychopathology in individuals from 4 to 18 years of age. The Hyperactivity, Attention Problems, Externalizing Problems, and Internalizing Problems Scales and the Behavior Symptoms Index were dependent variables for EEG group outcome. Only Parent scales completed by mothers were included in the study. Pre- and posttreatment BASCs completed by fathers, patients, and teachers were not available in sufficient numbers to be included in the analysis.

The Brown Attention-Deficit Disorder (ADD) Scales are symptom checklists with versions for adolescents (13–18 years) and adults (19 years and above). Five Cluster scores evaluate the ability to organize and activate for work, sustain attention and concentration, sustain energy and effort, manage affective interference, and utilize working memory. The Total score (sum of the Cluster scores) was a dependent variable for EEG group outcome. It indicates that a diagnosis of an attention-deficit disorder is "possible but not likely," "probable but not certain," or "highly probable."

EEG Group Evaluation

The EEG group baseline evaluation included the TOVA, intelligence testing if current IQ data were not available, the BASC (7–18 years old), and/or the Brown ADD Scales (13 years and older). TOVA testing was completed between 8 a.m. and 11:30 a.m. with the author in the room. Retesting was scheduled at the same time ($\pm$30 min) as the baseline.

Medications for the six EEG patients treated with stimulants were discontinued 2 days before the baseline and posttreatment evaluations. The 2-day washout period is sufficient to produce TOVA results not contaminated by medication effects. Methylphenidate and dextroamphetamine are completely metabolized within 12–24 hr (DuPaul, Barkley, & Connor, 1998) and produce behavioral effects for 12 hr or less (Barkley, 1990).

After baseline testing, medication was reinstated for the six EEG patients. Four of the six made sufficient improvement that they terminated stimulants midway through treatment. At posttreatment evaluations, these patients were medication-free for a minimum of 6 weeks.

Posttreatment reevaluations were carried out after 40 neurofeedback sessions for office patients and 3 months (60 + sessions) for home patients. Reevaluations included the TOVA, BASC, and/or the Brown ADD Scales.

MED Group Evaluation

The MED group baseline evaluation included the TOVA and intelligence testing if current IQ data were not available. Medication titration was usually scheduled 3 days (range 3–7) after the patient started medication. Patients were tested with the TOVA on as many as four doses of a stimulant. If the TOVA did not improve, the titration process was repeated with a different stimulant. MED patients were maintained on the medication and dose that maximized improvement on the TOVA. No additional evaluations were scheduled with MED patients.

Procedure

Office neurofeedback was provided by a Lexicor Neurosearch-1620 and home neurofeedback by Lexicor PODs (Lexicor Medical Technology, Boulder, CO). A sampling rate of 128 Hz with 2-s epochs was used. Twelfth order digital filters defined the steepness of the bands. Biolex version 2.38 software (NRS-1620) or Mental Conditioning software version 2.38 (POD) provided neurofeedback. The active electrode was at C3 or C4 (10–20 International System). The reference was on the earlobe ipsilateral to the active electrode with the ground on the contralateral earlobe. The patient's skin was prepared using Nuprep. Electrodes were filled with Ten20 electrode paste. Skin impedance was less than 10 KΩ.

Office patients ($n = 14$) were typically seen three times a week (range 3–5) for 40 treatment sessions over 3 1/2 months. Home patients ($n = 17$) received 60+ training sessions over 3 months (Rossiter, 1998). This included four office sessions used to teach the patient and/or parents to use the computerized biofeedback equipment and to understand the information provided. Two to four additional office neurofeedback sessions were scheduled during the 3-month home program. Supervision was provided through e-mail and telephone contacts.

Patients presenting with inattention, daydreaming, poor sustained attention, and/or lack of motivation received left hemisphere training with the active electrode at C3 (International 10–20 System) using enhance 15–18 Hz protocols. The C3 default inhibit band was initially 4–7 Hz and later 2–7 Hz. When the baseline EEG showed excessive alpha (8–11 Hz), an 8–11 Hz or 2–10 Hz inhibit band was used. Patients with symptoms of impulsivity, distractibility, and/or stimulus-seeking received right hemisphere training with the active electrode at C4 using enhance 12–15 Hz protocols. The C4 default inhibit band was initially 4–7 Hz and later changed to 2–7 Hz. The neurofeedback software was programmed to control eye movement and EMG artifact. Inattentive type AD/HD patients ($n = 15$) received left hemisphere (C3) training. Combined type AD/HD patients ($n = 16$) started each session with left hemisphere (C3) training and finished with right hemisphere (C4) training.

EEG treatment sessions included 30 or 36 min of neurofeedback. Training was conducted eyes open. No cognitive challenges (e.g., reading, drawing, listening, etc.) were used. The patient received simultaneous visual and auditory feedback based on the ratio of the inhibit band to the enhance band. Rossiter (2002) provides details of the neurofeedback procedures.

Statistical Analysis

Group comparisons (EEG vs. MED) and treatment effects (pre- vs. posttreatment) were evaluated using one-way multivariate analysis of variance (MANOVA). Significant effects were followed by planned comparisons of TOVA, BASC, and Brown ADD Scale scores changes using one-tailed t tests for dependent measures. One-tailed tests were used because the direction of change was predicted on the basis of the results of earlier studies (Fuchs et al., 2003; Monastra et al., 2002; Rossiter & La Vaque, 1995). MANOVAs and t tests were conducted using ProStat Version 3 software (Poly Software International, 2002). Effect sizes were calculated for repeated measures (Kazdin, 2003, p. 446). The Holm adjustment (Stevens, 1999) for multiple tests was used to maintain the experiment-wise $\alpha = .05$.

Equivalence/noninferiority testing (Rossiter, 2004) was used to determine whether the proportion of EEG patients improved was equivalent, or noninferior, to that of the MED group. The high proportion of MED patients improving (84%) precluded superiority testing. Proportions were compared because the sample was too small ($n = 31$) and heterogeneous to make equivalence/noninferiority testing based on TOVA mean scores feasible. More importantly, the proportion of patients significantly improved is a more meaningful measure of clinical outcome

RESULTS

Baseline EEG and MED Matching

EEG and MED groups were matched by age, the sum of the four baseline TOVA scores, IQ, gender, and primary AD/HD diagnosis in that order. Inspection of the demographic variables for the two groups (Table I) indicates that matching was successful except perhaps for the distribution of the AD/HD diagnostic categories. Matching of patients on the sum of the four TOVA scores (Table II) produced groups that were virtually identical on this measure. This procedure also resulted in comparable group means for the four TOVA scores. A one-way MANOVA demonstrated no statistically significant differences at baseline between EEG and MED groups on the TOVA (Wilks' Criterion = 0.94; $df = 4$, 57; $F = 0.92$; $p = .45$). In general, the EEG and MED groups were successfully matched on demographic variables and AD/HD related cognitive deficits (TOVA scores).

EEG and MED TOVA

It was predicted that EEG and MED groups would improve significantly on TOVA outcome scores (Table II). A one-way MANOVA of EEG group TOVA scores demonstrated statistically significant improvement (Wilks' Criterion = 0.54; $df = 4$, 57; $F = 12.13$; $p < .001$). Planned comparisons of pre- and posttreatment TOVA scores using one-tailed t tests for dependent measures confirmed that the EEG group demonstrated improved attention ($t = 4.29, df = 30, p < .001$), reduced impulsivity ($t = 4.39, df = 30, p < .001$), increased processing speed ($t = 3.99$, $df = 30$, $p < .001$), and decreased variability in attention ($t = 4.62$, $df = 30$, $p < .001$).

Table II. TOVA Means, Standard Deviations, and Effect Sizes for EEG and MED Groups

TOVA variables	EEG				MED			
	Pre	Post	Change	ES	Pre	Post	Change	ES
Omission	85.3 (25.2)	103.7 (6.7)	18.4 (23.9)	1.09	87.7 (21.8)	102.8 (12.8)	15.1 (26.6)	0.80
Commission	97.9 (15.2)	109.1 (10.2)	11.2 (14.2)	1.12	94.1 (19.5)	104.4 (15.6)	10.3 (15.7)	0.93
Response time	87.8 (21.9)	101.6 (20.0)	13.8 (19.3)	1.01	85.4 (18.1)	94.9 (16.8)	9.6 (11.4)	1.19
Variability	83.6 (22.4)	103.5 (14.2)	22.4 (24.1)	1.17	86.4 (19.5)	106.1 (17.5)	19.7 (15.3)	1.82
Total	354.5 (50.3)	418.0 (32.5)	63.4 (52.6)	1.71	353.5 (54.2)	408.3 (40.1)	54.7 (43.0)	1.80

Note. TOVA scores with $M = 100$, $SD = 15$, $n = 31$ for EEG and MED groups.

A one-way MANOVA demonstrated statistically significant improvement for the MED group on the TOVA (Wilks' Criterion = 0.74; $df = 4, 57$; $F = 5.12$; $p = .001$). The MED group showed gains in attention ($t = 3.17$, $df = 30$, $p = .002$), impulse control ($t = 3.67$, $df = 30$, $p < .001$), processing speed ($t = 4.67$, $df = 30$, $p < .001$), and decreased variability in attention ($t = 7.17$, $df = 30$, $p < .001$).

It was predicted that differences between the EEG and MED groups on TOVA gain scores would not be statistically significant. A one-way MANOVA (Wilks' Criterion = 0.95 $df = 4, 57$, $F = 0.76$, $p = .55$) indicated that the null hypothesis could not be rejected. Power was $\geq .80$ (Kazdin, 2003).

Posttreatment mean TOVA scores were within the average range (standard score = 90–109) for both the EEG and MED groups. Prior to treatment, only the impulsivity (errors of commission) scores for both the EEG and MED groups were within the average range.

EEG BASC and Brown ADD Scales

It was predicted that the EEG group would demonstrate significant improvement on BASC and Brown ADD Scale scores (Table III). Posttreatment BASC data were available for 23 of 25 patients. A one-way MANOVA demonstrated statistically significant improvement on the BASC (Wilks' Criterion = 0.54; $df = 4, 50$; $F = 6.90$; $p < .001$). Planned comparisons confirmed significant improvement on the Hyperactivity ($t = 4.57$, $df = 22$, $p < .001$), Attention Problems ($t = 7.00$, $df = 22$, $p < .001$), Externalizing Problems ($t = 4.53$, $df = 22$, $p < .001$), and Internalizing Problems ($t = 6.56$, $df = 22$, $p < .001$) Scales, and Behavioral Symptoms Index ($t = 6.90$, $df = 22$, $p < .001$). Posttreatment group means for the scales were within normal limits. Prior to treatment, only the Internalizing Problems Scale was within the Average range.

Pre- and posttreatment Brown ADD Scale scores were available for all EEG patients ages 13 and older (Table III). As predicted, EEG patients experienced significant improvement on the Total score ($t = 6.42$, $df = 10$, $p < .001$) with the group means moving from the diagnosis of AD/HD "highly probable" range prior to treatment to the "possible but not likely" range following treatment.

Table III. BASC and Brown ADD Scale Means, Standard Deviations, and Effects Sizes for EEG Group

	Pretreatment	Posttreatment	Change	Effect size
BASC				
Hyperactivity	64.8 (19.1)	51.8 (12.9)	13.0 (13.6)	1.16
Attention problems	70.9 (8.3)	57.6 (9.0)	13.3 (9.1)	1.78
Externalizing problems	61.6 (13.9)	52.1 (9.5)	9.5 (10.0)	1.15
Internalizing problems	58.8 (10.1)	47.0 (4.9)	11.8 (8.6)	1.67
Behavior Symptoms Index	66.1 (11.3)	51.4 (8.4)	14.7 (10.2)	1.75
Brown ADD Scales				
Total score	68.8 (14.2)	40.1 (13.5)	28.7 (14.8)	1.59

Note. BASC scores are t scores with $M = 50$, $SD = 10$, $n = 23$. Brown ADD Scale scores are t scores with $M = 50$, $SD = 10$, $n = 11$.

Equivalence Testing

It was predicted that the proportion of EEG patients demonstrating significant improvement on the TOVA would be equivalent, or noninferior, to the proportion of significantly improved MED patients (Rossiter, 2004). A patient was significantly improved if the number of TOVA scores improved (gain of >7.5 points) exceeded the number worsened (loss of >7.5 points) and the sum of the four TOVA scores increased by a minimum of 15 points over the pretreatment baseline. Twenty-six of the thirty-one patients in the EEG and MED groups improved significantly. Equivalence/noninferiority testing used both the confidence interval approach (Westlake, 1981) and the nonequivalence null hypothesis approach (Anderson & Hauck, 1983). The equivalence interval chosen was the de facto standard of 20% (Schuirmann, 1987) of the proportion of patients in the MED group who improved (84%). The confidence interval approach yielded a $CI_{90\%}$ (−0.154 to 0.154) that was contained within the Equivalence Interval (±0.168). The nonequivalence null hypothesis approach yielded both z_1 (1.796) and z_2 (−1.796) greater than $z_{p<.05}$ (1.645). Both methods confirm the hypothesis that outcomes for the EEG group were equivalent to those for the MED group. Using the Holm procedure (Stevens, 1999), planned comparisons for which significant differences were predicted met the adjusted alpha levels with the experiment-wise $\alpha = .05$.

DISCUSSION

The results of the current study confirm the hypotheses being tested. The EEG and MED groups demonstrated statistically significant improvement on TOVA scores. However, the fact that a treatment results in statistically significant improvement does not necessarily mean that the treatment effect is clinically significant or important. There is no consensus regarding what standards should be used to define clinical significance. Alternatives suggested include a high percentage of patients improving, elimination of the presenting problem, normal functioning by the end of treatment, a degree of change that is recognizable by significant others in the patient's life (Jacobson & Truax, 1991) and large effect sizes (Stevens, 2002).

Gains made by the EEG and MED groups on the TOVA were clinically significant. This conclusion is based on the percentage of patients showing significant improvement over baseline (84% each); large effect sizes for both treatments (EEG = 1.01–1.71; MED = 0.80–1.82); percentage of individual TOVA scores showing significant improvement (EEG = 55%, MED = 56%); and posttreatment mean scores for MED and EEG groups that fall within the average range.

The EEG group demonstrated statistically significant improvement on BASC and Brown ADD Scale scores. In addition, gains made by the EEG group on measures of behavioral change were clinically significant. This conclusion is based on the large effect sizes of the EEG group on the BASC (1.15–1.75) and Brown ADD Scales (1.59); and posttreatment BASC and Brown ADD Scale mean scores that fall within the average range.

Differences between TOVA gain scores for the EEG and MED groups were not statistically significant. More importantly, the proportion of EEG patients that significantly improved was equivalent to that of the MED group.

Data from Rossiter and La Vaque (1995) were reanalyzed using the Holm procedure (Stevens, 1999) to control the experiment-wise alpha level for multiple comparisons. All planned comparisons for which significant differences were predicted met their adjusted alpha levels for significance with the experiment-wise $\alpha = .05$. Equivalence/noninferiority testing indicated that the proportion of the EEG group patients significantly improved was noninferior, but not equivalent (Rossiter, 2004), to that of the MED group.

The results of the current study and the statistical reanalysis of data from Rossiter and La Vaque (1995) support the view that a treatment program with neurofeedback as the primary component produces patient outcomes that are equivalent to, or noninferior to, those obtained with stimulant drugs. The improvement demonstrated by the EEG patients was not limited to reduction of the core AD/HD symptoms, that is, inattention, impulsivity, and hyperactivity. They also manifested significant decreases in internalizing and externalizing symptoms and psychopathology more generally. Posttreatment mean scores for the EEG group on the TOVA, BASC, and Brown ADD Scales were within normal limits.

It should be noted that when six patients treated with stimulants were removed from the EEG group, the TOVA gains were unchanged and did not differ from those of the MED group. In addition, patients who received home and office neurofeedback made comparable posttreatment gains on the TOVA.

The conclusions that can be drawn from the study are limited somewhat by the choice of an effectiveness research design. Effectiveness research is typically conducted in a clinical setting and utilizes patients seeking treatment and expecting improvement. Their presenting problems may include multiple diagnostic categories. Because the research is conducted in a clinical setting, some compromises in research methodology and experimental controls have to be made for practical and ethical reasons. Treatment may be tailored to meet the needs of the individual patient. The result is that not all patients receive exactly the same treatment. Furthermore, it is the patient, not the clinician, who is ultimately responsible for choosing the treatment. In essence, an effectiveness study can evaluate a treatment as it is actually provided in clinical practice. Effectiveness studies place greater emphasis on external validity than do efficacy studies for which internal validity is of paramount importance. Therefore, effectiveness research has the potential for broad applicability to the real world spectrum of patients as they present for treatment in clinics and hospitals (Clarke, 1995). Because of the less stringent experimental controls, an effectiveness study can demonstrate that a treatment program is clinically effective, but it may not be possible to establish to what extent various elements (e.g., the treatment under study, patient expectations, therapist characteristics, placebo, etc.) contribute to the positive outcomes.

This study allowed patients to choose between treatments (stimulant drug therapy and/or office or home neurofeedback), used heterogeneous patient groups with co-morbid disorders, and tailored individual neurofeedback protocols based on presenting symptoms and baseline EEG patterns. These deviations from strict experimental controls would be serious flaws in an efficacy study but are acceptable variations in an effectiveness study. However, they do preclude attributing the improvement in the EEG group solely to neurofeedback. The influence of nonspecific factors cannot be ruled out. This is not problematic if the goal is to assess the "real world" effectiveness of a treatment program with neurofeedback as the primary component. It is significant that Fuchs et al. (2003) and Monastra et al. (2002) independently obtained similar results with different clinicians, settings, patient populations, and treatment protocols.

The most significant design weakness in the study is that different testing schedules were used for the EEG and MED groups. Ideally, both the MED and EEG patients would have been reevaluated after 3 to 3 1/2 months of their respective treatments. However, testing schedules were based on different clinical needs of the two groups. The MED group was retested with the TOVA to determine the most effective dose of methylphenidate or dextroamphetimine. Medication titration was completed in 3–10 days after instituting stimulant drug therapy. Once the maintenance dose was established, no additional evaluations were scheduled. Reevaluation 3 months later was not clinically necessary or feasible. Adjustments in maintenance medication levels are seldom needed in less than 6–12 months barring significant change in the patient's weight, health, or behavior. Methylphenidate and dextroamphetimine are immediately effective and do not demonstrate incremental behavioral improvement or tolerance over time (DuPaul et al., 1998). AD/HD does not wax and wane and there is no evidence that it can be "outgrown" in 3 months. The time disparity between reevaluations of the MED and EEG groups, although not desirable, does not invalidate comparison of the EEG and MED TOVA scores. If anything, it may overstate the effectiveness of the stimulants. Compliance with taking stimulants is typically poor (DuPaul et al., 1998). Firestone (1982) found that 20% of AD/HD patients terminated stimulant drugs within 4 months. Furthermore, effectiveness beyond 4 weeks of treatment has not been demonstrated (Schachter, Pham, King, Langford, & Moher, 2001).

REFERENCES

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

Anderson, S., & Hauck, W. W. (1983). A new procedure for testing equivalence in comparative bioavailability and other clinical trials. *Communications in Statistics—Theory and Methods*, *12,* 2663–2692.

Barkley, R. A. (1990). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment*. New York: Guilford Press.

Brown, T. E. (1996). *Brown Attention-Deficit Disorder Scales*. San Antonio, TX: Psychological Corporation.

Clarke, G. N. (1995). Improving the transition from basic efficacy research to effectiveness studies: Methodological issues and procedures. *Journal of Consulting and Clinical Psychology*, *63,* 718–725.

DuPaul, G. J., Barkley, R. A., & Connor, D. E. (1998). Stimulants. In R. A. Barkley (Ed.), *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed., pp. 263–293). New York: Guilford Press.

Firestone, P. (1982). Factors associated with children's adherence to stimulant medication. *American Journal of Orthopsychiatry*, *52*(3), 447–457.

Fuchs, T., Birbaumer, N., Lutzenberger, W., Gruzelier, J. H., & Kaiser, J. (2003). Neurofeedback treatment for attention-deficit/hyperactivty disorder in children: A comparison with methylphenidate. *Applied Psychophysiology and Biofeedback*, *28*(1), 1–12.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59,* 12–19.

Kaufman, A., & Kaufman, N. (1990). *Kaufman Brief Intelligence Test*. Circle Pines, MN: AGS Publishing.

Kazdin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Boston: Allyn and Bacon.

Leark, R. A., Dupuy, T. R., Greenberg, L. M., Corman, C. L., & Kindschi, C. L. (1996). *Test of Variables of Attention Professional Manual Version 7.0*. Available from Universal Attention Disorders, 4281 Katella, Suite 215, Los Alamitos, CA 90720.

Lubar, J. F., & Shouse, M. N. (1976). EEG and behavioral changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (SMR): A preliminary report. *Biofeedback and Self-Regulation*, *3,* 293–306.

Monastra, V. J., Monastra, D. M., & George, S. (2002). The effects of stimulant therapy, EEG biofeedback, and parenting style on the primary symptoms of attention-deficit/hyperactivity disorder. *Applied Psychophysiology and Biofeedback*, *27*(4), 231–249.

Poly Software International. (2002). *ProStat Version 3* (Computer software). Pearl River, NY.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior Assessment System for Children manual*. Available from American Guidance Service, 4201 Woodland Road, Circle Pines, MN 55014.

Riccio, C. A., Reynolds, C. R., & Lowe, P. A. (2001). *Clinical applications of continuous performance tests: Measuring attention and impulsive responding in children and adults*. New York: Wiley.

Rossiter, T. R. (1998). Patient directed neurofeedback for AD/HD. *Journal of Neurotherapy*, *2,* 54–64.

Rossiter, T. R. (2002). Neurofeedback for AD/HD: A ratio feedback case study and tutorial. *Journal of Neurotherapy*, *6,* 9–35.

Rossiter, T. R. (2004). The effectiveness of neurofeedback and stimulant drugs in treating AD/HD: Part I. review of methodological issues. *Applied Psychophysiology and Biofeedback*, *29,* 95–112.

Rossiter, T. R., & La Vaque, T. J. (1995). A comparison of EEG biofeedback and psychostimulants in treating attention deficit hyperactivity disorders. *Journal of Neurotherapy*, *1,* 48–59.

Schachter, H., Pham, B., King, J., Langford, S., & Moher, D. (2001). How efficacious and safe is short-acting methylphenidate for the treatment of attention-deficit disorder in children and adolescents? A meta-analysis. *Canadian Medical Association Journal*, *165,* 1475–1488.

Schuirmann, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, *15,* 657–680.

Stevens, J. S. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum.

Stevens, J. S. (1999). *Intermediate statistics: A modern approach* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Westlake, W. J. (1981). Bioequivalence testing—A need to rethink. *Biometrics*, *37,* 591–593.