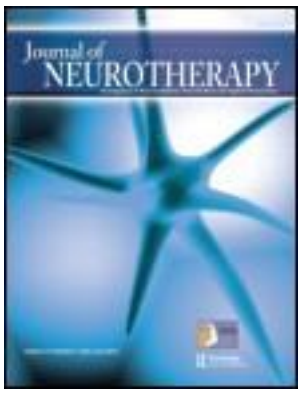

Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience

# Neurotherapy and the Challenge of Empirical Support: A Call for a Neurotherapy Practice Research Network

Lonnie A. Nelson MA [a]

[a] University of Arizona, Department of Psychology, Human Energy Systems Laboratory , Tucson, AZ, 85719 E-mail:

Published online: 08 Sep 2008.

**To cite this article:** Lonnie A. Nelson MA (2003) Neurotherapy and the Challenge of Empirical Support: A Call for a Neurotherapy Practice Research Network, Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience, 7:2, 53-67, DOI: 10.1300/J184v07n02_04

**To link to this article:** http://dx.doi.org/10.1300/J184v07n02_04

# Neurotherapy and the Challenge of Empirical Support: A Call for a Neurotherapy Practice Research Network

Lonnie A. Nelson, MA

**ABSTRACT.** This paper summarizes a review of the empirical support for neurotherapy (NT) as a treatment for psychological and neurological disorders according to the criteria for efficacious treatments set forth by Chambless and Hollon (1998). The review classifies the level of efficacy established for five of ten disorders examined according to the evidence reported in the literature. Specific suggestions for two distinct future research strategies are given. The first of these is in the area of mechanism research and efficacy; the second is a proposal for the creation of an observational study Practitioner Research Network (PRN) aimed at providing data on the effectiveness of neurotherapy as practiced in the field.

**KEYWORDS.** Efficacy, effectiveness, empirical support, evidence based medicine, observational studies

## *INTRODUCTION*

Electroencephalographic (EEG) biofeedback is among the most promising modalities for the treatment of psychological and neurological dis-

Lonnie A. Nelson is a Doctoral Candidate at the University of Arizona, Department of Psychology, Human Energy Systems Laboratory, Tucson, AZ 85719 (E-mail: lan@u.arizona.edu).

Journal of Neurotherapy, Vol. 7(2) 2003

10.1300/J184v07n02_04

orders in an enduring and effective (in terms of clinical utility) manner. Currently the research in support of the efficacy (as demonstrated by laboratory controlled clinical trials) of biofeedback for any specific disorder, though impressive in certain instances (e.g., epilepsy), still leaves much to be desired in most areas.

The object of this paper is three fold: (a) to briefly review the empirical basis for the clinical use of neurofeedback, (b) to suggest a research strategy that is both complete and cost effective, and (c) to request the participation of clinicians in the field in a large-scale observational study (or Practitioner Research Network). It is hoped that the effect of this paper will be to focus the neurotherapy (NT) community (practitioners and researchers) into a coherent whole that can face the challenges of evidence-based medicine in a unified manner in order to improve its position in the medical and psychological community at large.

## *A BRIEF REVIEW OF THE EVIDENCE FOR NEUROTHERAPY*

The PsycINFO database search was used with the keywords "EEG Biofeedback," "Neurotherapy," "Neurofeedback" and "Brainwave biofeedback" with limits of English language, and publication in or after 1980. In this review of the empirical basis for neurofeedback (NF), 39 articles were found that addressed specific, well-defined clinical disorders. These articles addressed a total of 10 disorders which were identified for review. These were: Attention Deficit/Hyperactivity Disorder (ADHD), Depression, Generalized Anxiety Disorder (GAD), Obsessive-Compulsive Disorder (OCD), Post-Traumatic Stress Disorder (PTSD), Schizophrenia, Substance Abuse, Traumatic Brain Injury (TBI), and Seizure Disorders (epilepsy). Reviewing the literature of these disorders, it was found that only five of these disorders had a sufficient research base to be evaluated. Those were ADHD, PTSD, Substance Abuse, TBI, and Seizure Disorders. This is primarily due to the lack of either (a) sufficient or appropriate control groups, (b) sufficient sample sizes, or (c) the lack of appropriate statistical analyses in the studies examining the effects of neurofeedback on the other disorders (Nelson, 2002).

In order to evaluate the empirical basis for any effect, it is useful to have a set of criteria to which one may compare the evidence. Fortunately, Chambless and Hollon (1998) have provided a reasonable set of criteria by which the efficacy of therapeutic approaches may be categorized according to the degree of empirical support for them reported in

the literature. The minimum criteria are as follows for specific hierarchical categories of efficacy.

*Possibly Efficacious:* This designation is given to treatments with only one study supporting their efficacy in comparison to a wait-list or assessment only control group or with all of the studies supporting efficacy coming from one researcher or group of researchers.

*Efficacious:* This designation is given to treatments that have been found efficacious in comparison to at least a wait-list or assessment only control group in at least two studies by independent researchers with an absence of conflicting findings with regard to efficacy.

*Efficacious and Specific:* This designation is given to treatments that have been demonstrated to be superior to a treatment condition that controls for nonspecific factors such as receiving attention and the expectation of change or a bona fide treatment. Comparisons to rival interventions are considered more highly in this category but the minimum requirement remains any "active" control condition. This could in theory be any control condition in which a single specific variable is controlled for in an active fashion (e.g., social support). However, as stated above, in studies using an established intervention as the control condition, in a soundly designed investigation intended to evaluate the relative efficacy of an unproven treatment, a finding of no difference from the established treatment is considered a positive finding for the unproven treatment in the absence of conflicting findings.

Due to the large number of case studies reported in the field of neurotherapy, a remark should be made regarding their status in affecting the categorization process. Chambless and Hollon (1998) suggest that case studies may be valuable in determining the efficacy of a treatment if they allow causal inference by way of their design. Designs such as the classic ABAB multiple baseline design in which a plateau or increase in symptomatology is demonstrated at the removal of treatment in the context of decreases in symptomatology with the introduction of treatment are examples of this type of design.

Unfortunately, there were no such case studies reported in the neurotherapy literature; all similar studies used an ABA design without the multiple baseline assessment. These reports are not to be ignored by any means, they simply do not allow the same certainty with regard to causal inference that a multiple baseline study affords.

It should also be noted that these classifications refer to the protocols tested in the research reports and not necessarily the EEG biofeedback component in isolation. The object of this sort of classification is simply to determine if an intervention is likely to be helpful for individuals with

a given disorder. Thus, the efficacy classification does not speak to the issue of mechanism.

A review of the literature on the efficacy of neurotherapy for ADD/ADHD revealed 10 research reports including both analog studies and actual clinical trials. Only one of these trials used an actual rival treatment as a control group (Rossiter & La Vaque, 1995). Two of the studies (Patrick, 1996; Linden, Habib, & Radojevic, 1996) used a wait-list control group. Five used a simple pre-post comparison, and two used a "learner," "non-learner" comparison. Of these 10 studies, seven showed positive results in favor of neurotherapy; however, the outcomes of the remaining three were unclear due to a lack of proper statistical analysis.

Even with this weakness in these studies, according to the Chambless and Hollon (1998) criterion above, neurotherapy for ADD/ADHD qualifies as an *efficacious* treatment. If, in the single study using a rival treatment as a comparison group, both groups had not received behavior therapy in addition to neurotherapy and psychostimulants (a major confounding factor) without changing the results, this classification would be *efficacious and specific.*

A single report examining the effects of neurotherapy on PTSD using a treatment as usual (TAU) comparison group showed positive results (reviewed in Moore, 2000). This single study classifies neurotherapy as a *possibly efficacious* treatment for PTSD.

Of a total of six clinical trials investigating the effects of neurotherapy on substance abuse, three of the studies (reviewed in Trudeau, 2000) used a TAU comparison group, two used no control group, and one used a TAU and sham feedback condition for comparison. Given that the substances abused by the subjects were variable (alcohol and stimulants), the findings of these studies are mixed for specificity, and positive for efficacy, classifying neurotherapy as an *efficacious* treatment for substance abuse.

Two case studies, one case series and one clinical trial (Byers, 1995; Rozelle & Budzynski, 1995; Thornton, 2000; and Schoenberger, Shiflett, Esty, Ochs, & Matheis, 2001; respectively) were reviewed that addressed the efficacy of neurotherapy in treating traumatic brain injury (TBI). The case studies reviewed used no comparison conditions, and therefore cannot be counted toward efficacy even though positive results were reported. The clinical trial reported used a wait-list control group and reported positive findings, classifying neurotherapy for TBI as a *possibly efficacious* treatment.

Seizure disorders (epilepsy) are the most strongly supported of the neurotherapy applications. Sterman (2000) reviewed 29 clinical trials,

four using an ABAB reversal paradigm, six using a sham feedback control group, eight using a pre-treatment baseline comparison, 10 using no control groups and one using a wait list control group. All reported positive results in favor of the efficacy of neurotherapy. Based on these reports, neurotherapy can clearly be classified as both *efficacious and specific* for the treatment of seizure disorders.

For a summary of the efficacy classifications for the studies reviewed in this paper see Table 1.

## *SUGGESTIONS FOR A COMPREHENSIVE RESEARCH STRATEGY*

There are several points that require consideration in the evaluation of neurotherapy as a treatment approach for neurological and psychological disorders. The two main areas of consideration have to do with efficacy (the demonstration of the effect of a treatment under laboratory conditions) and effectiveness (the demonstration of the effect of a treatment in the field). While these two areas are certainly related, they are separate issues, and must be addressed with different research strategies.

In terms of efficacy, one should consider the mechanism of accomplishing functional change in neurotherapy and whether this affects any durable structural (e.g., dendritic arborization, vascular channel formation, etc.) changes with prolonged training. One is then led to consider the neuropsychology of psychopathology, the specific mechanisms of plasticity and regeneration of neural tissue, both in specific developmental periods and across the lifespan in general, and how these factors might interact with the above mentioned mechanisms.

In addition to this, one might consider the characteristics of the social and therapeutic context in which the treatment is given, and the possibility that the interaction of all of these variables combines to explain the benefits that are observed in the actual clinical (or research) setting. In addition to this, painstaking effort should be put forth in future reports including detailed descriptions of the exact protocols used, the duration of actual training time of each protocol, the frequency (or intervals) of training and the complete number of sessions that specific protocols are employed. In addition to this, the criteria for deciding to change a protocol should be recorded, as well as specifics related to the training context, such as whether the client was left alone during training or "coached." These details are important not only for analysis of a

TABLE 1

| **Disorder** | **Number of studies reviewed** | **Control groups** | **Results** | **Chambless and Hollon (1998) Category?** |
|---|---|---|---|---|
| ADHD | 10 (analog and clinical trials) | Rival Tx: 1<br>Wait-List: 2<br>Pre-Post: 5<br>Learner vs. Nonlearner: 2 | Positive: 7<br>Unclear due to lack of statistics: 3 | *Efficacious* |
| Depression | 2 (analog studies)<br>3 (case studies) | N/A : 4<br>Reversed Asymm. : 1 | Positive: 2 (analog studies)<br>3 (case reports) | Insufficient data for actual clinical trials |
| GAD | 4 (analog studies)<br>1 (case study) | Sham feedback, EMG feedback, etc. & N/A | Mixed (analog)<br>Positive (case) | Insufficient data for actual clinical trials |
| Phobic Anxiety Disorder | 2 (analog studies) | Untreated control group | Positive: 2 | Insufficient data for actual clinical trials |
| OCD | 2 (clinical trials) | "Learners vs. Nonlearners" | Positive: 2 (no stats, small N) | Insufficient data for actual clinical trials |
| PTSD | 1 (clinical trial) | TAU | Positive: 1 | *Possibly Efficacious* |
| Substance Abuse | 6 (clinical trials) (see Trudeau, 2000) | TAU: 3<br>N/A: 2<br>TAU+sham: 1 | Mixed for specificity, positive for Efficacy | Efficacious |
| Schizotypy and Schizophrenia | 2 (analog studies) | N/A | Mixed | Insufficient data for actual clinical trials |
| Traumatic Brain Injury | 3 (case studies)<br>1 (clinical trial) | Case studies: N/A<br>Clinical trial: Wait-list control | Case studies: Positive<br>Clinical trial: Positive | *Possibly Efficacious* |
| Seizure Disorders | 28 (clinical trials)<br>12 (case studies) | Clinical trials:<br>4 ABA design<br>6 sham fdbck.<br>8 baseline<br>1 wait-list<br>10 N/A<br>Case studies: N/A | Clinical trials: All Positive<br>Case studies:<br>11 positive<br>1 no change | *Efficacious and Specific* |

specific mechanism, but also because such detailed documentation allows treatments to be "manualized" for replication by other interested parties.

Finally, one may wonder whether these mechanisms might be affected without the use of equipment, via specific mental exercises (e.g., sensorimotor integration exercises, behavioral stillness, timing exercises, etc.; Mulholland, 1995). This issue speaks to the cost effectiveness of neurotherapy versus other related techniques.

Are there particular types of EEG manipulation that *require* the actual apparatus and other forms of EEG change for which the feedback is unnecessary or even intrusive? It is not difficult to imagine that an individual attempting to produce increased levels of alpha might fair better without an intermittent "beep" that may serve primarily as a distraction (feedback), causing a form of task related cortical activation that could be counterproductive, and that some mental exercises such as various meditative techniques might prove more effective in this specific capacity (e.g., Tyson, 1987; Echenhofer & Coombs, 1987). When one wishes to manipulate a more specific aspect of the EEG, for example, the frontal alpha asymmetry of an individual, the apparatus may be absolutely required in the absence of some form of explicit cognitive retraining (Allen, Harmon-Jones, & Cavender, 2001).

Another question that arises is whether there might be critical developmental periods during which neurotherapy is likely to be maximally and minimally effective? One might expect that a child's brain may be somewhat more plastic (at least structurally) than an aged adult's. This would likely make a considerable difference in terms of the ability to build new functional (and possibly structural) pathways (Othmer, 2000). Findings reported above by Andrews and Schonfeld (1992) indicate that an early age of seizure onset was a valuable predictor in terms of the number of sessions required to bring seizures under control.

Similarly, are there points in the course of the development of a specific disorder in which neurotherapy is likely to prove maximally and minimally effective? That is, would neurotherapy be more effective in the early stages of a disorder, or is it equally effective with late stage and chronic conditions? Finally, what are the limits of neurofeedback under each of these conditions?

Without the proper research the answers to most of these questions can only be speculated upon. However, in order to employ neurotherapy in the most fruitful manner possible, answers to these questions are needed. The laboratory methods are available and straight forward for addressing these sorts of questions. This type of laboratory research

has been ongoing for decades and will hopefully be continued fruitfully. However, these questions are related to only one of the areas in which research is needed, the area of efficacy. The question of effectiveness in the field requires a fundamentally different approach.

## *SUGGESTIONS FOR AN "EFFECTIVENESS" RESEARCH STRATEGY*

In general, EEG biofeedback (neurotherapy) appears to be a promising treatment for a variety of disorders. However, the research base is limited and with a few noted exceptions, is not well formulated. Of the various strategies available to the contemporary researcher in the science of clinical psychology, the researchers of the neurotherapy community have used relatively few. The most popular today is the randomized clinical trial (RCT). Though there are few empirical reasons to accept this as the preferable research paradigm, nonrandomized designs that are sound in other ways yield very similar results, and depending on the specific design can be much less costly (Kaptchuk, 2001; Shadish, Navarro, Matt, & Phillips, 2000).

In terms of viable research strategies, there are various ways to go about the process of producing empirical support that offer highly valid and occasionally more reliable findings than the traditional placebo controlled RCT (Kaptchuk, 2001; Concato, Shah, & Horwitz, 2000). A good example of a valid alternative strategy for demonstrating the effectiveness of an intervention is the observational study (Feinstein, 1989). Critically represented historically, observational studies have many advantages over RCTs, and their downfalls are quickly minimized when one considers what has been learned about conducting research in clinical populations in the last 20 years (Benson & Hartz, 2000).

One of the major reasons for the widespread negative evaluation of observational studies in clinical areas is the often cited finding that the results of observational studies systematically exaggerate the differences between control and treatment groups. This observation is due primarily to the use of inappropriate controls in the studies reviewed by Sacks, Chalmers, and Smith (1983) and Chalmers, Celano, Sacks, and Smith (1983). At the time that these studies were conducted, inconsistent inclusion criteria, the assignment of patients with worse prognoses to control groups, and the susceptibility bias that resulted did in fact produce larger differences between treatment and control groups than

one would see in a randomized placebo controlled RCT. Often ignored is the fact that this difference was not due to an inflation of the rates of improvement in the treatment group, but a disproportionate likelihood of poor outcomes in the control groups (Benson & Hartz, 2000).

Similarly, all of the major problems cited with reference to observational studies in the current opinion can be remedied by (a) inclusion and exclusion criteria standards, (b) current statistical techniques (e.g., intention to treat analysis to avoid performance bias), along with (c) a standardization of assessment. The standardization of assessment would provide reliable and consistent data that could be used to determine levels of pre-treatment severity and post-treatment change, and appropriate control treatments in terms of subject activity and attention. In addition to these, drop-out rates must be documented in order to equate for "survival rates" across the groups (Feinstein, 1989; Heinsman & Shadish, 1996; Benson & Hartz, 2000).

Given these points, the neurotherapy community would do well to heed this information since observational studies are faster, cheaper, and similarly valid to RCTs (though more heavily weighted toward ecological validity) as well as having the capability of providing effectiveness data regarding treatments that are already being used in the field.

An additional concern in employing observational studies is that they do not address the issues of specific causal factors to the degree that a laboratory implemented design might. It should be noted that laboratory designs are valuable, but as explanatory and exploratory endeavors, rather than applied ones. An excellent example of this relationship is demonstrated in the above section reviewing the literature concerning neurotherapy and depression. The case reports reviewed were published before the excellently designed laboratory work of Allen et al. (2001) and may even have played some role in the planning of the research. Alternatively, the work of Baehr, Rosenfeld, and Baehr (1997) was based on the earlier laboratory findings of Henriques and Davidson (1991). This example of interplay between laboratory findings and applied findings is exactly the approach suggested here. The two types of research should be complementary.

For the applied aspects of the research, observational studies may provide the simplest, fastest and least expensive means. Borkovec, Echemendia, Ragusea, and Ruiz (2001), in a special issue of *Clinical Psychology*, report on the implementation of the idea of a Practitioner Research Network (PRN), in which practitioners would agree to conduct standardized assessments of their clients before, at certain time points during therapy, at termination, and at follow up periods. This ef-

fectively made the practice of every participating clinician part of a huge laboratory aimed at providing effectiveness data to find out how the interventions were working in the field. It is exactly this sort of approach that this author feels would provide the effectiveness data that the field of neurotherapy needs at this juncture.

This network could be of varying size, with varying numbers of practitioners and patients. The main prerequisite for the proposed approach on the part of the network of practitioners is first and foremost, the participation of as many clinicians in the field as possible, who are willing to provide high quality basic data. Beyond that, certain agreed upon procedures must be employed by these participants, these would include (a) uniform admission criteria for specific disorders or presenting complaints, (b) standardized assessments at standardized points along the course of treatment and at follow-up, (c) documentation of protocols employed, and (d) documentation of additional treatments the individuals may be receiving.

Uniform admission and exclusion criteria for inclusion in the data set are a fundamental requirement for interpretation of the eventual results. The specific criteria will be left unspecified here, to be considered on the basis of specific disorders to be investigated, hopefully by a joint group of researchers and clinicians. However, the uniformity of the sample across treatment groups is the main issue of concern. This uniform sample should be representative of the "real world" population insofar as this is possible with the understanding that in the "real world" comorbidity may be the rule for certain disorders or disorder "sets." Again, since the purpose of the research network is to evaluate the course of treatment in practice, the focus should be on what is seen in practice as a clinician.

The specific details of this prerequisite would need to be decided jointly by a team of practitioners and clinical scientists in order to meet the balance between uniformity and reality. Since inclusion criteria simultaneously means exclusion criteria, it must be understood that individuals who do not meet these criteria will still be treated, and while data may be collected on them, they should be categorized (and analyzed) appropriately and separately from those who meet the criteria. A coding system for severity could be included to ensure that the appropriate statistical comparisons are made.

Standardized assessments are another fundamental requirement for clear interpretation of findings. One possibility would be to administer a "core battery" (Borkovec et al., 2001) of assessments at standard interval time points, while employing "session" instruments consisting of a

few items that could be administered easily at the beginning and conclusion of each session. This may allow a fine-grained (in terms of time) analysis that would detect rates of change that the "core battery" misses because of its low "sampling rate." The specifics of this set of assessment tools could be decided at some point in the future in the same joint manner as the other issues. However, there are some well-defined reasons for using a specific set of popular indicators that should be seriously considered.

The assessments chosen to compose the "core battery" should be the most reliable, inexpensive, time-effective, and valid indicators available for each variable under examination (Borkovec et al., 2001). Since the neurotherapy community is already equipped to use the QEEG in a diagnostic manner, and since the theoretical foundation of neurotherapy predicts that changes should be visible from this measure, this becomes an obvious choice for all of the disorders chosen for study in the neurotherapy PRN. However, there is substantial variability in data acquisition, processing and interpretation as practiced in the field (Kaiser, 2001). These factors would need to be standardized among the practitioners participating in the research network. Other assessments used should mirror as closely as possible the standard assessment tools used in other studies in the clinical psychology literature for use in meta-analytic techniques described below.

The implementation of these assessment tools at standard intervals in the course of treatment and at specific points in follow up are absolutely essential to the acquisition of reliable observations. Only with reliable observations can valid comparisons between time points be made. For a course of neurotherapy that requires 40 sessions, sensible time points might be (a) pre-treatment, (b) mid-point (20 sessions), (c) post-treatment, and (d) follow-up assessment at 6 or 12 months (or both). This would allow for a rough estimate of the course of change in neurotherapy. Also, it would provide data regarding the duration of changes achieved by way of neurotherapy for comparison with those obtained through other treatment modalities.

Another essential point in this effectiveness research design is the documentation of the EEG biofeedback training protocol or combinations thereof employed in the course of treatment as well as (as previously stated) the duration of actual training time that each protocol is used, the frequency (or intervals) of training, and the complete number of sessions that specific protocols are employed. In addition to this, the criteria for deciding to change protocol should be recorded, as well as specifics related to the training context, such as whether the client was

left alone during training or "coached." The importance of this factor cannot be overstated. These details are important not only for analysis of specific mechanisms, but also because such detailed documentation allows treatments to be "manualized" for replication by other interested parties.

A related issue of similar importance for applied effectiveness research is the painstaking documentation of all other treatments a client may be receiving while also receiving neurotherapy. Given that many of the clients to health services are receiving some other treatment simultaneously, it is necessary to know what other factors are affecting the course of their condition. This is perhaps the largest challenge (in terms of internal validity) to this type of research design. It is not at all uncommon for an individual with a severe disorder to seek alternative therapies (which neurotherapy debatably qualifies as), additional conventional treatment, or some mixture of the two. It may prove that neurotherapy in combination with a nutritional program and a detoxification period is far more effective than neurotherapy alone.

The central "hub" of this effort will face many challenging administrative and methodological issues. Primary among them will be the selection of an appropriate comparison group for any given disorder and the evaluation of improvement from pretreatment status. Given that the most (debatably) popular assessment technique in the neurotherapy community, Quantitative Electroencephalography (QEEG), is inherently reliant upon the Normative Reference Database, the question arises as to whether or not it is appropriate to compare an individual who has either a history of, or the current manifestation of, any mental illness or neurological disorder to a normative database composed of individuals who have never had such a disturbance for the purposes of gauging improvement.

For diagnostic purposes this approach is appropriate; however, the same electrophysiological characteristics that provided the predisposition to a given condition (or the effects on the brain of having had the condition in the case of TBI or severe epilepsy) may remain even after restoration of normal function and the relief of symptoms (e.g., frontal alpha asymmetry and persons with a history of depression; Henriques & Davidson, 1990). Of course, this is only one issue of consideration, as there are many other factors to take into account when gauging EEG changes as a marker of disorder remission, many of which are imposed by the limitations of our analysis techniques themselves.

This concern may be addressed by using, as a measure of improvement, not the deviation from "normal" that the individual shows at the

point of post-treatment and follow-up, but the amount of change in the characteristics of the QEEG from pre-treatment to post-treatment (and follow-up).

For other assessment measures, particularly popular symptom relief indicators, it may be appropriate to use quasi-meta analytic techniques to obtain comparison groups if sufficient statistical power can be approximated (e.g., Shadish et al., 2000). This type of meta-analytic technique may eventually be used in this type of research endeavor to eliminate the need for a control group of individuals who simply go through the assessments on a wait list, or are given an alternative treatment in the place of neurotherapy, as would be done in a randomized clinical trial. Having the ability to compare effect sizes with the effect sizes of treatments that are already empirically supported gives this approach a cost and time effectiveness that is virtually unheard of in clinical research. In this particular capacity, the "effect size" statistic (of which there are many) is particularly useful, as normalized effect size provides a common metric between comparable measurements on diverse treatment groups. This metric is exactly the unit that "matters" in terms of clinical utility of a technique as it is essentially a measure of "how far" (in terms of population standard deviations) a group's scores have been moved by an intervention. Having the ability to compare "how far" neurotherapy can move a group with a uniform disorder with "how far" therapy "X" can move a group with the same uniform disorder would be a very useful argument to make for the effectiveness of neurotherapy in the field.

Along the above delineated line, the neurotherapy Research Practice Network could join with another (perhaps the Pennsylvania network?) using a rival intervention comparison. In either of these cases, the comparisons would be valid and provide estimates of effect sizes that could be used to demonstrate effectiveness and predict outcomes as well as answering many of the theoretical questions raised earlier regarding optimal ages, and stages of disorder development for response to neurotherapy, as well as issues of plasticity in relation to age (as exploratory, rather than comparison analyses, of course). Combined with strong laboratory analog studies exploring mechanism, component analysis and physiological markers of disorders, this strategy could potentially be very fruitful.

Whether clinical researchers in the field of applied EEG biofeedback (neurotherapy) use this empirical approach or another, it is clear from the above review that more applied research is both warranted and required in this promising area.

## REFERENCES

Allen, J. J. B., Harmon-Jones, E., & Cavender, J. H. (2001). Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG. *Psychophysiology*, *38* (4), 685-693.

Andrews, D. N., & Schonfeld, W. H. (1992). Predictive factors for controlling seizures using a behavioural approach. *Seizure*, *1*, 111-116.

Baehr, E., Rosenfeld, J. P., & Baehr, R. (1997). The clinical use of an alpha asymmetry protocol in the neurofeedback of depression: Two case studies. *Journal of Neurotherapy*, *2* (2), 10-23.

Benson, K., & Hartz, A. J. (2000). A comparison of observational studies and randomized, controlled trials. *New England Journal of Medicine*, *324* (25), 1878-1886.

Borkovec, T. D., Echemendia, R. J., Ragusea, S. A., & Ruiz, M. (2001). The Pennsylvania practice research network and future possibilities for clinically meaningful and scientifically rigorous psychotherapy effectiveness research. *Clinical Psychology Science and Practice*, *8*, 155-167.

Byers, A. P. (1995). Neurofeedback therapy for a mild head injury. *Journal of Neurotherapy*, *1* (4), 22-37.

Chalmers, T. C., Celano, P., Sacks, H. S., & Smith, H., Jr. (1983). Bias in treatment assignment in controlled clinical trials. *New England Journal of Medicine*, *309*, 1538-1561.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66* (1), 7-18.

Concato, J., Shah, N., & Horwitz, R. I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine*, *324* (25), 1887-1892.

Echenhofer, F. G., & Coombs, M. M. (1987). A brief review of research and controversies in EEG biofeedback and meditation. *The Journal of Transpersonal Psychology*, *19* (2), 161-171.

Feinstein, A. R. (1989). Epidemiological analyses of causation: The unlearned scientific lessons of randomized trials. *Journal of Clinical Epidemiology*, *6*, 481-489.

Heinsman, D. T., & Shadish, W. R. (1996). Assignment methods in experimentation: When do nonrandomized experiments approximate answers from randomized experiments? *Psychological Methods*, *1* (2), 154-169.

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

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

Kaiser, D. A. (2001). QEEG: State of the art, or state of confusion. *Journal of Neurotherapy*, *4* (2), 57-75.

Kaptchuk, T. J. (2001). The double blind, randomized, placebo-controlled trial: Gold standard or golden calf? *Journal of Clinical Epidemiology*, *54*, 541-549.

Linden, M., Habib, T., & Radojevic, V. (1996). A controlled study of the effects of EEG biofeedback on cognition and behavior of children with attention deficit disorder and learning disabilities. *Biofeedback and Self-Regulation*, *21* (1), 35-49.

Moore, N. C. (2000). A review of EEG biofeedback for anxiety disorders. *Clinical Electroencephalography*, *31* (1), 1-6.

Mulholland, T. (1995). Human EEG, behavioral stillness and biofeedback. *International Journal of Psychophysiology*, *19*, 263-279.

Nelson, L. (2002). A review of the empirical basis for applied electroencephalographic (EEG) biofeedback (neurotherapy) for the treatment of neurological and psychological disorders. Manuscript submitted for publication.

Othmer, S. (2000). EEG biofeedback: An emerging model for its global efficacy. In J. R. Evans, & A. Abarbanel, (Eds.), *Quantitative EEG and neurofeedback* (pp. 243-310). New York: Academic Press.

Patrick, G. (1996). Improved neuronal regulation in ADHD: An application of fifteen sessions of photic driven EEG neurotherapy. *Journal of Neurotherapy*, *2* (4), 27-36.

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* (1), 48-59.

Rozelle, G. R., & Budzynski, T. H. (1995). Neurotherapy for stroke rehabilitation: A single case study. *Biofeedback and Self-Regulation*, *20* (3), 211-228.

Sacks, H. S., Chalmers, T. C., & Smith, H., Jr. (1983). Sensitivity and specificity of clinical trials: Randomized vs. historical controls. *Archives of Internal Medicine*, *143*, 753-755.

Schoenberger, N. E., Shiflett, S. C., Esty, M. L., Ochs, L, & Matheis, R. J. (2001). Flexyx neurotherapy system in the treatment of traumatic brain injury: An initial evaluation. *Journal of Head Trauma Rehabilitation*, *16* (3), 260-274.

Shadish, W. R., Navarro, A. M., Matt, G. E., & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin*, *126* (4), 512-529.

Sterman, M. B. (2000). Basic concepts and clinical findings in the treatment of seizure disorders with EEG operant conditioning. *Clinical Electroencephalography*, *31* (1), 45-55.

Thornton, K. (2000). Improvement/rehabilitation of memory functioning with neurotherapy/QEEG biofeedback. *Journal of Head Trauma Rehabilitation*, *15* (6), 1285-1296.

Trudeau, D. L. (2000). The treatment of addictive disorders by brainwave biofeedback: A review and suggestions for future research. *Clinical Electroencephalography*, *31* (1), 13-22.

Tyson, P. D. (1987). Task related stress and EEG alpha biofeedback. *Biofeedback and Self-Regulation*, *12* (2), 105-119.

RECEIVED: 10/15/02
REVISED: 12/02/02
ACCEPTED: 12/06/02