Evaluating 19-Channel Z-score Neurofeedback:

Addressing Efficacy in a Clinical Setting

Submitted by

Nancy L. Wigton

A Dissertation Presented in Partial Fulfillment

of the Requirements for the Degree

Doctorate of Philosophy

Grand Canyon University

Phoenix, Arizona

May 15, 2014

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

GRAND CANYON UNIVERSITY

Evaluating 19-Channel Z-score Neurofeedback:

Addressing Efficacy in a Clinical Setting

by

Nancy L. Wigton

has been approved

May 15, 2014

APPROVED:

Genomary Krigbaum, Psy.D., Dissertation Chair

Daniel J. Smith, Ph.D., Committee Member

Eugenia Bodenhamer-Davis, Ph.D., Committee Member

ACCEPTED AND SIGNED:

Henry T. Radda, Ph.D.
Dean, College of Doctoral Studies

5/15/14
Date

Abstract

Neurofeedback (NF) is gaining recognition as an evidence-based intervention grounded in learning theory, and 19-channel z-score neurofeedback (19ZNF) is a new NF model. Peer-reviewed literature is lacking regarding empirical-based evaluation of 19ZNF. The purpose of this quantitative research study was to evaluate the efficacy of 19ZNF, in a clinical setting, using archival data from a Southwest NF practice, with a retrospective one-group pretest-posttest design. Each of the outcome measures framed a group such that 19ZNF was evaluated, as it relates to the particular neuropsychological constructs of attention (n = 10), behavior (n = 14), executive function (n = 12), as well as electrocortical functioning (n = 21). The research questions asked if 19ZNF improves these constructs. One-tailed t tests performed, compared pre-post scores for included clinical assessment scales, and selected quantitative electroencephalographic (QEEG) metrics. For all pre-post comparisons, the direction of change was in the predicted direction. Moreover, for all outcome measures, the group means were beyond the clinically significant threshold before 19ZNF, and no longer clinically significant after 19ZNF. All differences were statistically significant, with results ranging from p = .000 to p = .008; and effect sizes ranging from 1.29 to 3.42. Results suggest 19ZNF improved attention, behavior, executive function, and electrocortical function. This study provides beginning evidence of 19ZNF's efficacy, adds to what is known about 19ZNF, and offers an innovative approach for using QEEG metrics as outcome measures. These results may lead to a greater acceptance of 19ZNF, as well as foster needed additional scientific research.

Keywords: Neurofeedback, QEEG, z-score neurofeedback, 19ZNF, EEG biofeedback

Dedication

This dissertation is dedicated to my Lord and Savior, Jesus. From my first thoughts of considering a doctoral program being divinely inspired and directed, through to the last step I will take across a graduation stage, the Father, Son, and Holy Spirit are always the center point, the anchor. To that end, three Bible passages capture the experience of my journey.

> The way of God is perfect, the Lord's word has stood the test; He is the shield of all who take refuge in Him. What god is there but the Lord? What rock but our God? – the God who girds me with strength and makes my way blameless, who makes me swift as the deer and sets me secure on the mountains (Psalms 18:30-33, New English Bible).

"Commit your life to the Lord; trust in Him and He will act. He will make your righteousness shine clear as the day and the justice of your cause like the sun at noon" (Psalms 37:5-6).

"Not to us, O Lord, not to us, but to thy name ascribe the glory, for thy true love and for thy constancy" (Psalms 115:1).

Acknowledgments

It is only through the Lord's strength and wisdom that this dissertation came to fruition. Next, I acknowledge the man with whom the Lord has made me one, my husband. You are truly the wind beneath my wings, and without you I would not have had the wherewithal to complete this endeavor. Thank you for all your support and sharing your perseverance for my good. I also wish to acknowledge, with unbounded gratitude, the most perfect dissertation committee possible for this journey.

To my chair, Dr. Genomary Krigbaum, words are insufficient to fully express the depth and breadth of my appreciation for your support, guidance, and direction. When I first read descriptions of what the ideal chair would be, with characteristics inclusive of mentor, advocate, role model, teacher, defender, guide, supervisor, coach, encourager, and friend, I wondered if it would ever be possible to find all those elements in one person. Yet in you, I found them all, and more. *Por siempre agradecida.* Moreover, thank you for encouraging me to build on the methodology you started. To Dr. Daniel Smith, I am grateful that you joined my dissertation team. I knew I could count on you for your statistical expertise, and you did not disappoint. Thank you for the many conversations prior to my dissertation journey, and in helping to pave the way for the best committee possible. To Dr. Genie Bodenhamer-Davis, as a *most respected* neurofeedback practitioner and educator, I am humbled and honored that you were willing to assist me in my dissertation journey. Thank you, so much, for your counsel over the last 3 years. To Dr. Ron Bonnstetter, thank you for your support in being my adjunct dissertation reader. Thank you for your compliments on my writing and your assurance I have what it takes to succeed as a scholar.

Table of Contents

List of Tables xi

List of Figures xii

Chapter 1: Introduction to the Study 1

Introduction 1

Background of the Study 2

Problem Statement 4

Purpose of the Study 5

Research Questions and Hypotheses 6

Advancing Scientific Knowledge 8

Significance of the Study 9

Rationale for Methodology 10

Nature of the Research Design for the Study 11

Definition of Terms 13

Assumptions, Limitations, Delimitations 19

Summary and Organization of the Remainder of the Study 22

Chapter 2: Literature Review 23

Introduction and Background to the Problem 23

Historical overview of EEG and QEEG. 24

Historical overview of NF 25

How problem/gap of 19ZNF research evolved into current form 28

Theoretical Foundations and/or Conceptual Framework 31

Foundations of EEG and QEEG 31

Learning theory as applied to NF....31
Traditional/amplitude-based models of NF....33
QNF model of NF....35
ZNF model of NF....38
Review of the Literature – Key Themes....39
QNF in the literature....39
4ZNF in the literature....47
19ZNF in the literature....50
Outcome measures for ZNF research....53
Summary....59
Chapter 3: Methodology....61
Introduction....61
Statement of the Problem....61
Research Questions and Hypotheses....62
Research Methodology....64
Research Design....65
Population and Sample Selection....66
Instrumentation....68
Validity....72
Reliability....74
Data Collection Procedures....76
Data Analysis Procedures....78
Ethical Considerations....81

Limitations 82

Summary 84

Chapter 4: Data Analysis and Results 86

Introduction 86

Descriptive Data 86

Data Analysis Procedures 93

Results 96

Summary 103

Chapter 5: Summary, Conclusions, and Recommendations 105

Introduction 105

Summary of the Study 106

Summary of Findings and Conclusion 107

Implications 113

Theoretical implications 114

Practical implications 115

Future implications. 116

Recommendations 117

Recommendations for future research. 117

Recommendations for practice. 118

References 120

Appendix A 136

Appendix B 137

Appendix C 138

Appendix D 139

List of Tables

Table 1.1. Research Questions and Variables.....8

Table 4.1. Descriptive Data for All Groups..... 91

Table 4.2. Shapiro-wilk Results for Difference Scores 95

Table 4.3. Summary of Results - All Groups.....104

List of Figures

Figure 1.1. Formation of Sample Groups 13

Figure 4.1. IVA Group Pre-Post Scores..... 97

Figure 4.2. DSMD Group Pre-Post Scores 99

Figure 4.3. BRIEF Group Pre-Post Scores 101

Figure 4.4. QEEG Group Pre-Post Scores 102

Chapter 1: Introduction to the Study

Introduction

Neurofeedback (NF) is an operant conditioning brainwave biofeedback technique, which is also referred to as electroencephalographic (EEG) biofeedback. This modality, dating back to the 1970s (Lubar & Shouse, 1976; Sterman, LoPresti, & Fairchild, 2010), trains electrical signals of targeted frequencies and involves recording EEG data from scalp sensors with an amplifier, which is subsequently processed by computer software. The software provides visual and sound display feedback to the trainee, thereby providing a reward stimulus when the brain is functioning in the target range. This reward process generates learning such that the brain's functioning is conditioned in the intended manner.

Over the years, new models of NF have been developed, and the most current iteration is a style of NF which is termed z-score NF (ZNF). ZNF is different from more traditional NF models in that it incorporates into the NF session real-time quantitative EEG (QEEG) z-score metrics making it possible to combine operant conditioning with real-time assessment using a normative database (Collura, Thatcher, Smith, Lambos, & Stark 2009; Thatcher, 2012). In 2006, a 4-channel ZNF (4ZNF) technique was introduced, which in 2009 was expanded to include all 19 sites of the International 10-20 System (of electrode placement) to allow for a 19-channel ZNF (19ZNF). To date, case study and anecdotal clinical reports within the field indicate this new 19ZNF approach is an improvement over traditional NF models (J. L. Koberda, Moses, Koberda & Koberda, 2012a; Wigton, 2013). However the efficacy of this new model has not yet been established from empirical studies. This research is different from prior qualitative

studies; it has been completed as a quantitative analysis of pre-post outcome measures with group data, and thus, it is a beginning in establishing empirical evidence regarding 19ZNF.

The remainder of this chapter formulates this dissertation through a review of the study background, problem statement, purpose and significance, and how this research advances the scientific knowledge. Moreover the research questions and hypotheses are presented, together with the methodology rationale and the nature of the research design. An extended Definition section is included to review the many technical terms germane to this research. Readers unfamiliar with NF or QEEGs may find it helpful to review the definitions first. Finally, to establish the scope of the study, a list of assumptions, limitations, and delimitations are included.

Background of the Study

In recent years NF has seen increasing acceptance as a therapeutic technique. Current literature includes reviews and meta-analyses which establish a recognition of NF as effective for the specific condition of attention deficit hyperactivity disorder (ADHD) (Arns, de Ridder, Strehl, Breteler, & Coenen 2009; Brandeis, 2011; Gevensleben, Rothenberger, Moll, & Heinrich, 2012; Lofthouse, Arnold, Hersch, Hurt, & DeBeus, 2012; Niv, 2013; Pigott, De Biase, Bodenhamer-Davis, & Davis, 2013). However, the type of NF covered in these reviews is limited to the oldest NF model (theta/beta ratio) and/or slow cortical potential NF. Yet of note are reports in the literature of a different NF model which is informed by QEEG data. This QEEG-guided NF (QNF) is reported to be used for a much wider range of conditions; not only ADHD, but also behavior disorders, cognitive dysfunction, various mood disorders, epilepsy,

posttraumatic stress disorder, head injuries, autism spectrum disorders, migraines, learning disorders, schizophrenia, and mental retardation (Arns, Drinkenburg, & Kenemans, 2012; Breteler, Arns, Peters, Giepmans, & Verhoeven, 2010; Coben & Myers, 2010; J. L. Koberda, Hillier, Jones, Moses, & Koberda 2012; Surmeli, Ertem, Eralp, & Kos, 2012; Surmeli & Ertem, 2009, 2010, 2011; Walker, 2009, 2010b, 2011, 2012b).

Yet, all the aforementioned models are limited in their use of only one or two electrodes and they also require many sessions to achieve good clinical outcomes. For the above-cited studies the reported average number of sessions was 40.5. Moreover, Thatcher (2012, 2013) reports 40 to 80 sessions to be the accepted norm for these older style models; thus leading to a sizeable cost to access this treatment. However, one of the newest ZNF models shows promise to bring about positive clinical outcomes in significantly fewer sessions (Thatcher, 2013). With 4ZNF there have been reports of successful clinical outcomes with less than 25 sessions (Collura, Guan, Tarrant, Bailey, & Starr, 2010; Hammer, Colbert, Brown, & Ilioi, 2011; Wigton, 2008); whereas clinical reviews and recent conference reports (J. L. Koberda, Moses, Koberda, & Koberda, 2012b; Rutter, 2011; Wigton, 2009, 2010a, 2010b, 2013; Wigton & Krigbaum, 2012) suggest 19ZNF can result in positive clinical outcomes, as well as QEEG normalization, in as few as 5 to15 sessions. Therefore a NF technique which shows promise to bring clinical improvement in fewer sessions – thereby reducing treatment cost – deserves empirical study.

Currently in the peer-reviewed published literature, there are a couple of descriptive and clinical review articles *about* the 19ZNF model (Thatcher, 2013; Wigton,

2013) and two single case study reports (Hallman, 2012; J. L. Koberda et al., 2012a); however rigorous scientific studies evaluating 19ZNF have not been found, which poses a gap in the literature. Therefore, before the question of efficiency and number of sessions is examined, first its *efficacy* should be established. NF and ZNF efficacy has been discussed in the literature as having the desired effect in terms of improved clinical outcomes (La Vaque et al., 2002; Thatcher, 2013; Wigton, 2013), a definition that fits well within the scope of this research. In this study, there are two types of clinical outcome measures; one type (clinical assessments) is a set of psychometric tests designed to measure symptom severity and/or improvement, the other type (QEEG z-scores) provides a representative measure of electrocortical dysfunction and/or improvement. Thus, this dissertation is intended to address efficacy of 19ZNF in a clinical setting, through a retrospective evaluation of clinical outcomes, as measured by clinical assessments and QEEG z-scores.

Problem Statement

It is not known, by way of statistical evaluation of either clinical assessments or QEEG z-scores, if 19ZNF is an effective NF technique. This is an important problem because 19ZNF is a new NF model currently in use by a growing number of practitioners, yet scientific research investigating its efficacy is lacking. According to an Efficacy Task Force, established by the two primary professional organizations for NF and biofeedback professionals,[1] anecdotal reports (regardless of how many) are insufficient as a basis for

[1]The primary professional societies for neurofeedback and biofeedback are the International Society for Neurofeedback and Research (ISNR; www.isnr.org) and the Association for Applied Psychophysiology and Biofeedback (AAPB; www.aapb.org).

determining treatment efficacy, and uncontrolled case studies are scientifically weak (La Vaque et al., 2002). Therefore, scientific evidence of efficacy for 19ZNF is needed.

The identified population for this study is made up of those seeking NF services (both adults and children), and those who become NF clients. These individuals may have an array of symptoms, which adversely affect their daily functioning; they may also have previously diagnosed mental health disorders. When seeking NF services these individuals must choose among a variety of NF models. However the dearth of scientific literature regarding 19ZNF limits the information available to inform that decision-making process. Therefore, it is vital that both NF clinicians and clients have empirically derived information regarding the clinical value and efficacy of this new NF technique. Consequently, the problem of this empirical gap impacts the NF clinician and client alike. The goal of this research is to contribute in providing a first step towards addressing this research gap.

Purpose of the Study

The purpose of this quantitative, retrospective, one-group, pretest-posttest study research was to compare the difference between pre and post clinical assessments and QEEG z-scores data, before and after 19ZNF sessions, from archived data of a private neurofeedback practice in the Southwest region of the United States. The comparisons were accomplished via statistical analysis appropriate to the data (i.e. paired *t* tests), and will be further discussed in the Data Analysis section of Chapter 3. The independent variable is defined as the 19ZNF, and the dependent variables are defined as the standard scaled scores of three clinical assessments and QEEG z-score data. The clinical assessments measure symptoms of attention, behavior, and executive function, whereas

the z-scores provide a representative measure of electrocortical function. The full scopes of the assessments are further outlined in the Instrumentation section of Chapter 3.

Given the retrospective nature of this study, there were no individuals, as subjects, with which to interact. However the target population group is considered to be adults and children with clinical symptoms of compromised attention, behavior, or executive function, who are interested in NF as an intervention for improvement of those symptoms. This pretest-posttest comparison research contributes to the NF field by conducting a scientific study, using quantitative group methods, to address the efficacy of the new 19ZNF model.

Research Questions and Hypotheses

If the problem to be addressed is a lack of scientific evidence demonstrating efficacy of 19ZNF, the solution lies in evaluating its potential for improving clinical outcomes as measured by clinical assessments and electrocortical metrics. Therefore research questions posed in terms of clinical symptomology and cortical function measures is a reasonable approach. For this research the independent variable is the 19ZNF and the dependent variables are clinical outcomes, as measured by the scaled scores from three clinical assessments and z-scores from QEEG data. The clinical assessments are designed to measure symptom severity of attention, behavior, and executive functioning, and the z-scores are a representational measure of electrocortical function. The data gathering, scores calculation, and, data analysis were conducted by the researcher.

The following research questions guided this study:

R1a. Does 19ZNF improve attention as measured by the Integrated Visual and Auditory continuous performance test (IVA; BrainTrain, Incorporated, Chesterfield, VA)?

H_a1a: The post scores will be higher than the pre scores for the IVA assessment.

$H_0$1a: The post scores will be lower than, or not significantly different from, the pre scores of the IVA assessment.

R1b. Does 19ZNF improve behavior as measured by the Devereux Scale of Mental Disorders (DSMD; Pearson Education, Incorporated, San Antonio, TX)?

H_a1b: The post scores will be lower than the pre scores for the DSMD assessment.

$H_0$1b: The post scores will be higher than, or not significantly different from, the pre scores of the DSMD assessment.

R1c. Does 19ZNF improve executive function as measured by the Behavior Rating Inventory of Executive Functioning (BRIEF; Western Psychological Services, Incorporated, Torrance, CA)?

H_a1c: The post scores will be lower than the pre scores for the BRIEF assessment.

$H_0$1c: The post scores will be higher than, or not significantly different from, the pre scores of the BRIEF assessment.

R2. Does 19ZNF improve electrocortical function as measured by QEEG z-scores (using the Neuroguide Deluxe software, Applied Neuroscience Incorporated, St.

Petersburg, FL), such that the post z-scores are closer to the mean than pre z-scores?

H_a2: The post z-scores will be closer to the mean than the pre z-scores.

H_02: The post z-scores will be farther from the mean, or not significantly different from, the pre z-scores.

See as follows Table 1.1, outlining the research questions and variables.

Table 1.1

Research Questions and Variables

Research Questions	Hypotheses	Variables	Instrument(s)
1a. Does 19ZNF improve attention as measured by the IVA?	The post scores will be higher than the pre scores for the IVA assessment.	IV: 19ZNF DV: IVA standard scale scores	IVA computerized performance test
1b. Does 19ZNF improve behavior as measured by the DSMD?	The post scores will be lower than the pre scores for the DSMD assessment.	IV: 19ZNF DV: DSMD standard scale scores	DSMD rating scale
1c. Does 19ZNF improve executive function as measured by the BRIEF?	The post scores will be lower than the pre scores for the BRIEF assessment.	IV: 19ZNF DV: BRIEF standard scale scores	BRIEF rating scale
2. Does 19ZNF improve electrocortical function as measured by QEEG z-scores such that the post z-scores are closer to the mean than pre z-scores?	The post QEEG z-scores will be closer to the mean than the pre z-scores.	IV: 19ZNF DV: QEEG z-scores	QEEG z-score data generated from Neuroguide software

Advancing Scientific Knowledge

The theoretical framework of NF is the application of operant conditioning upon the EEG, which leads to electrocortical changes, and in turn, better brain function and clinical symptom improvement; moreover, studies evaluating traditional NF have

demonstrated its efficacy (Arns et al., 2009; Pigott et al., 2013). The 19ZNF model is new, and experiencing increased use in the NF field, yet efficacy has not been established via empirical investigation. There is a gap in the literature in that the only peer-reviewed information available to date, regarding 19ZNF, are reviews, clinical report presentations, and single case studies. Also noted as typically absent from traditional NF studies are analyses of pre-post QEEG data (Arns et al., 2009); this lack of pre-post QEEG data continues in the QNF literature as well. This, then, poses a secondary gap, in terms of methodology, which this study has the potential to fill.

The clinical condition most researched for demonstrating traditional NF efficacy is ADHD (Pigott et al., 2013), which includes cognitive functions of attention and executive function. These issues also lead to some associated behavioral problems with adverse impacts in instructional settings that are also treated with 19ZNF. Therefore, addressing efficacy of 19ZNF with clinical assessments designed to measure these constructs, will contribute to filling the gap of what is not known about this new NF model, within a framework related to cognition and instruction. If efficacy is demonstrated, the theory of operant conditioning, upon which NF is founded, may be expanded to include 19ZNF.

Significance of the Study

The 19ZNF model is theoretically distinctly different from traditional NF in that it targets real-time QEEG z-scores with a goal of normalizing QEEG metrics (as indicated by clinical symptom presentation) rather than only increasing or decreasing targeted brain frequencies. This model has been in existence for five years and its use by NF clinicians is rapidly growing. Thus far, other than two qualitatively-oriented, single case study

reports (Hallman, 2012; J. L. Koberda et al., 2012a), there are no empirical group studies, with a quantitative methodology, studying the efficacy of 19ZNF in peer-reviewed literature. The significance of this study is that it aims to fill this significant gap manifest as a dearth of 19ZNF efficacy studies.

Moreover, few NF studies include analysis of EEG measures as an outcome measure (Arns et al., 2009). Therefore demonstrating how z-scores from QEEG data can be used for group comparison studies, in a way not previously explored, will benefit the scientific community. Thus, this research has the potential for opening doors for further research.

It was expected the findings would demonstrate 19ZNF results in improved clinical outcomes, as measured by clinical and QEEG assessments; thus demonstrating efficacy. Potential NF clients will benefit from this contribution of what is known about 19ZNF by having more information upon which to base decisions for what type of NF they wish to pursue. The potential effect of these results may provide the start of an evidence-based foundation for its use. This foundation may lead to a greater acceptance of what may be a more efficient (and thereby more economical) NF model, as well as foster the needed additional scientific research of 19ZNF.

Rationale for Methodology

The field of clinical psychophysiology makes use of quantifiable variables and the associated research should include specific independent variables, as well as dependent variables that relate to treatment response (e.g. clinical assessments) and the measured physiological component (e.g. EEG metrics) (La Vaque et al., 2002). Yet, many NF studies do not use the EEG metric as a psychophysiologic measure, but rather provide

reports, which are more qualitative in nature. Therefore, there is a need for NF research, with sound quantitative methodologies, using QEEG data as an outcome measure.

Currently, the available 19ZNF studies are in the form of qualitative research (Hallman, 2012; J. L. Koberda et al., 2012a). This literature entails presenting data, from single case studies, in the form of unstructured subjective reports of symptom improvement and graphical images of before and after QEEG findings, where the improvement is represented by a change in color on the picture (without statistical analysis of data). However, for this dissertation, the goal is to explore statistical relationships between the variables under investigation. The strength of quantitative methodologies, including quasi-experimental research, is that they provide sufficient information, regarding the relationship of the investigation variables, to enable the study of the effects of the independent variable upon the dependent variable (Carr, 1994); this is suitable in the evaluation of a quantitative technology such as 19ZNF.

As previously stated, for this research the independent variable is specified as 19ZNF. The dependent variables in this study are continuous variables in the form of standard scores from clinical assessments (IVA, DSMD, and BRIEF) and z-scores from QEEG data. The alternative hypotheses for all research questions predict a directional significant difference between the means of the pre and the post values for all dependent variables. Therefore, a quantitative methodology is appropriate for this dissertation.

Nature of the Research Design for the Study

This quasi-experimental research used a retrospective one-group, pretest-posttest design. When the goal of research is to measure a modification of a behavior pattern, or internal process that is stable and likely unchangeable on its own, the one-group pretest-

posttest design is appropriate (Kerlinger, 1986). In this type of design the dependent variable pretest measures are compared to the posttest values for each subject, thus comparing the members of the group to themselves rather than to a control or comparison group (Kerlinger, 1986). Consequently, the group members become their own control, hence reducing the potential for extraneous variation due to individual-to-individual differences (Kerlinger & Lee, 2000). Moreover, the size of the treatment effect can be estimated by analyzing the difference between the pretest to the posttest measures (Reichardt, 2009). Therefore, this design as well as a quantitative methodology, is well suited to evaluate the pre-post outcome measures from a clinical setting.

The rationale for this being a retrospective study is based on the fact that data available for analysis came from pre-existing archived records, which frequently provides a rich source of readily accessible data (Gearing, Mian, Barber, & Ickowicz, 2006). Within the pool of available data, a sample group was gathered for which various pre and post assessments were performed during the course of 19ZNF treatment. As depicted in Figure 1.1, an initial group was formed for which pre-post QEEG assessments and z-scores were available, and for which *either* the IVA, DSMD, *or* BRIEF pre-post assessment data was also available (n = 21). From this collection three additional groups were formed: One group for the IVA data (n = 10), a second group for the DSMD data (n = 14), and a third group for the BRIEF data (n = 12). Therefore, using a one-group pretest-posttest design with these identified groups is fitting. The independent variable is the 19ZNF and the dependent variables are the data from the clinical assessments and QEEG files (IVA, DMSD, BRIEF, and z-scores).

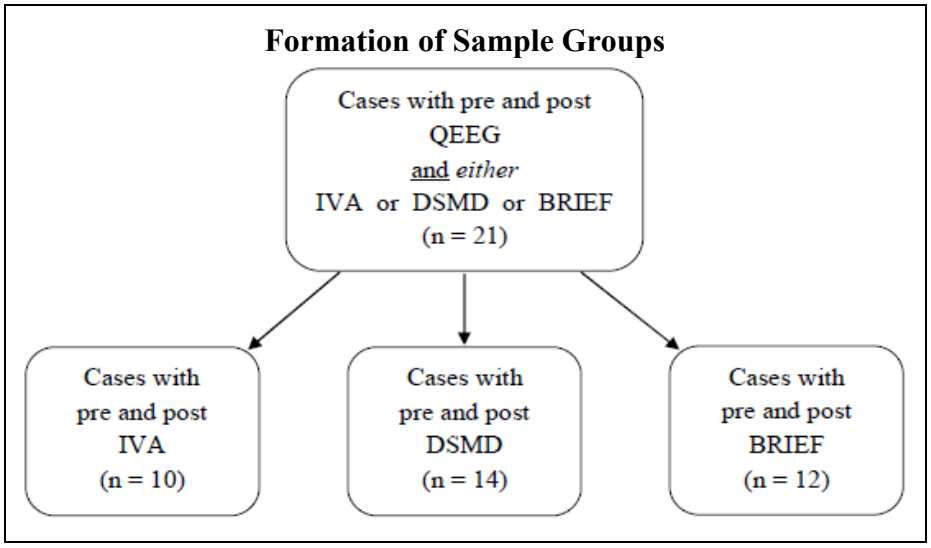


Figure 1.1. Illustration of how the sample groups were formed. The total number of subjects in the sample is 21. However, out of those 21, some may have multiple assessments, therefore subjects may be in more than one clinical assessment group.

Definition of Terms

The following terms were used operationally in this study.

19ZNF. 19-channel z-score NF is a style of NF using all 19 sites of the International 10-20 system, where real-time QEEG metrics are incorporated into the NF session in the form of z-scores (Collura, 2014). The goal is for the targeted excessive z-score metrics (whether high or low) to normalize (move towards the mean). The 19ZNF cases included in this study are those for which the assessed clinical symptoms corresponded with the z-score deviations of the QEEG findings, such that a treatment goal of overall QEEG normalization was clinically appropriate. While the 19ZNF protocols are individually tailored to the clinical and QEEG findings, the same treatment goal always applies, that is the overall QEEG normalization. Therefore, the underlying 19ZNF protocol of overall QEEG normalization is consistent for all cases.

Absolute power. A QEEG metric which is a measure of total energy, at each electrode site, for a defined frequency band (Machado et al., 2007); may be expressed in terms of microvolts, microvolts squared, or z-scores when compared to a normative database (Collura, 2014).

Amplifier. The equipment that detects, amplifies, and digitizes the brainwave signal (Collura, 2014). The term is more correctly referred to as a differential amplifier because the electrical equipment measures the difference between two signal inputs (brainwaves from electrode locations) (Collura, Kaiser, Lubar, & Evans, 2011).

Amplitude. A measure of the magnitude or size of the EEG signal; and is typically expressed in terms of microvolts (uV) (Collura et al., 2011). This can be thought of as *how much* energy is in the EEG frequency.

Biofeedback. A process of learning how to change physiological activity with the goal of improving health and/or performance (AAPB, 2011). A simple example of biofeedback is the act of stepping on a scale to measure one's weight.

Behavior Rating Inventory of Executive Functioning (BRIEF). The BRIEF, published by Western Psychological Services, Incorporated (Torrance, CA), is a rating scale. It has forms for both children and adults, and is designed to assess behavioral, emotional, and metacognitive skills, which broadly encompass executive skills, rather than measure behavior problems or psychopathology (Donders, 2002). The test results are expressed as T scores for various scales and sub-scales (with clinically significant scores ≥ 65), and lower scores indicate improvement upon re-assessment. The composite and global scales of Behavior Regulation Index, Metacognition Index, and Global Executive Composite were included in this study.

Coherence. A measure of similarity between two EEG signals, which also reflects the degree of shared information between the sites; computed in terms of a correlation coefficient, which varies between .00 to 1.00 (Collura et al., 2011).

Devereux Scale of Mental Disorders (DSMD). The DSMD, published by Pearson Education, Incorporated (San Antonio, TX), is a rating scale. It is designed to assess behavior problems and psychopathology in children and adolescents (Cooper, 2001). The test results are reported in the form of T scores for various scales and sub-scales (with clinically significant scores ≥ 60), and lower scores indicate improvement upon re-assessment. The composite and global scales of Externalizing, Internalizing, and Total were included in this study.

Electrode. Central to NF is the detection and analysis of the EEG signal from the scalp. In order to record brainwaves it is necessary to attached metallic sensors (electrodes) to the scalp and/or ears (with a paste or gel) to facilitate this process (Collura, 2014).

Electroencephalography (EEG). A recording of brain electrical activity (i.e. brainwaves) using differential amplifiers, measured from the scalp (Collura et al., 2011). The information from each site or channel is digitized to be viewed as an oscillating line, such that all channels can be viewed on a computer screen at one time.

Fast Fourier transform (FFT). The conversion of a series of digital EEG readings into frequency ranges/bands, which can be viewed in a spectral display. Just as different frequencies of light can be seen when filtered through a prism, so too can EEG elements be isolated when filtered through a FFT process into different frequency bands (Collura, 2014).

Frequency / frequency bands. The representation of how fast the signal is moving, expressed in terms of Hertz (Hz) (Collura, 2014) and commonly arranged in bandwidths, also referred to as bands. Generally accepted frequency bands are delta (1-4 Hz), theta (4-7 Hz), alpha (8-12 Hz), beta (12-25 Hz), and high beta (25-30 Hz); the beta band may be broken down into smaller bands of beta1 (12-15 Hz), beta2 (15-18 Hz), and beta3 (18-25 Hz), and the alpha band may be divided into alpha1 (8-10 Hz) and alpha2 (10-12 Hz).

Gaussian. Referring to the normal distribution and/or normal curve (Thatcher, 2012).

Hedges' *d*. An effect size, belonging to the *d* family indices (along with Hedges' *g*), which use the standard score form of the difference between the means; therefore it is similar to the Cohen's *d*, with the same interpretation (Hunter & Schmidt, 2004). However, when used with small sample sizes, both the Cohen's *d* and Hedges' *g*, can have an upward bias and be somewhat over-inflated; however the Hedges' *d* includes a correction for this bias (Hunter & Schmidt, 2004). Therefore, in studies with smaller sample sizes, the use of the Hedges' *d* provides a more conservative, and likely more accurate effect size. Also, complicating this issue is confusion in the literature regarding the use of the designator *g* or *d* for which particular Hedges index, and/or which calculation does or does not include the correction factor (Hunter & Schmidt, 2004). For example, frequently Hedges' *g* is described as adjusting for small samples sizes; however, this is only true if the calculation used includes the correction factor. Moreover, there are even variations in the literature of the correction equation which is applied. As a result, the only way to know which calculation is actually being used is for the Hedges'

index equation to be explicitly reported. To that end, for this study, the Hedges' *d* definition/calculation will be that used in the Metawin 2.1 meta-analysis software (Rosenberg, Adams, & Gurevitch, 2000). In this context the Hedges' *d* is calculated by multiplying the Hedges' *g* by the correction, which is sometimes referred to as J.

Where $g=\frac{\left(\bar{X}^{E}-\bar{X}^{C}\right)}{s}$ and $J=1-\frac{3}{4\left(N^{C}+N^{E}-2\right)-1}$ therefore $d=g\left[1-\frac{3}{4(n_e+n_c-2)-1}\right]$.

Hertz (Hz). The number of times an EEG wave oscillates (moves up and down) within a second; commonly expressed as cycles per second (Collura, 2014).

International 10-20 System. A standardized and internationally accepted method of EEG electrode placement locations (also referred to as *sites*) on the scalp. The nomenclature of 10-20 derives from electrode locations being spaced a distance of either 10% or 20% of the measured distance from certain landmarks on the head. The system consists of a total of 19 sites, with eight locations on the left, eight on the right, and three central sites found on the midline between the right and left side of the head (Collura, 2014).

Visual and Auditory + Plus Continuous Performance Test (IVA). The IVA, developed and published by BrainTrain Incorporated (Chesterfield, VA), is a computerized interactive assessment. It is normed for individuals over the age of 5, and it is designed to assess both auditory and visual attention and impulse control with the aim to aid in the quantification of symptoms and diagnosis of ADHD (Sanford & Turner, 2009). The test results are reported in the form of quotient scores for various scales and sub-scales (with clinically significant scores ≤ 85), and higher scores indicate improvement upon re-assessment. The global and composite scales of Full Scale

Attention Quotient, Auditory Attention Quotient, and the Visual Attention Quotient were included in this study.

Joint time frequency analysis (JTFA). A method of digitizing the EEG signal which allows for moment-to-moment (i.e. real time) measures of EEG signal changes (Collura, 2014).

Montage. The configuration of the electrodes and software defining the reference point and electrode linkages, for the differential recording of the EEG signals (Thatcher, 2012). For example, in a linked-ears montage, the signal for each electrode site is referenced to the signal of the ear electrodes linked together. In a Laplacian montage, the signal for each electrode site is referenced to the signal of the weighted average of the surrounding electrode sites.

Neurofeedback. An oversimplified, yet accurate, definition of neurofeedback is that it is simply biofeedback with brainwaves. Generally, it is an implicit learning process (involving both operant and classical conditioning) where changes in brainwave signal/patterns, in a targeted direction, generates a reward (a pleasant tone and change in a video animation) such that the desired brainwave events occur more often (Collura, 2014; Thatcher, 2012).

Normalization. In the context of NF, refers to the progression of excessive z-scores towards the mean (i.e. $z = 0$), meaning the NF trainee's EEG is moving closer to the EEG range of *normal* (i.e. typical) individuals of his/her age (Collura, 2014). Thus, the concept of normalization is generally accepted to be when the z-scores of the QEEG move towards the mean (i.e. in the *direction* of $z = 0$).

Power spectrum. The distribution of EEG energy across the frequency bands, typically from 1 Hz to 30 Hz and frequently displayed as a line graph, histogram, or color topographic (i.e. visual representations of the numerical data) images (Collura, 2014).

Phase. The temporal relationship between two EEG signals, reflecting the speed of shared information (Collura et al., 2011).

Protocol. The settings designated in NF software, informed by a treatment plan, which determines how the NF proceeds. This establishes parameters such as metrics (e.g. absolute vs. relative power), direction of training (i.e. targeting more or less), length of session, and other decision points in the NF process (Collura, et al., 2011).

Quantitative EEG (QEEG). The numerical analysis of the EEG such that it is transformed into a range of frequencies as well as various metrics such as absolute power, relative power, power ratios, asymmetry, coherence, and phase (Collura, 2014; Thatcher, 2012). The data is typically made up of raw numbers, statistical transforms into z-scores, and/or topographic images (Collura, 2014). As a dependent variable in this study, QEEG z-scores are considered a representational measure of electrocortical function. The metrics of absolute power, relative power, and coherence were included.

Relative power. A QEEG metric representing the amount of energy, divided by the total energy, at each electrode site, for a defined frequency band. It reflects how much energy is present compared to all other frequencies (Collura, 2014).

Assumptions, Limitations, Delimitations

This section identifies the assumptions and specifies the limitations, together with the delimitations of the study. The following assumptions were present in this study:

1. It was assumed that traditional neurofeedback is deemed efficacious as discussed and demonstrated in the literature (Arns et al., 2009; Pigott et al., 2013).
2. It was assumed that the subjects are representative of the population of those who seek NF treatment for various mental health disorders; thus allowing for results to be generalized to that population (Gravetter & Wallnau, 2010).
3. It was assumed the sample is homogeneous and selected from a population that fits the normal distribution such that the sample means distribution are also likely to fit a normal distribution (Gravetter & Wallnau, 2010).
4. It was assumed that responses provided on rating scale instruments accurately reflect perceived or remembered observations, thus minimizing bias for over or under-reporting of observations (Kerlinger & Lee, 2000).

The following limitations were present in this study:

1. Research design elements. A general limitation of designs that incorporate a pretest-posttest formulation is primarily related to the passage of time between administering the pre and post assessments (Kerlinger & Lee, 2000). Factors such as history and maturation cannot be controlled for; therefore it is not possible to know whether or not they have impacted the dependent variable measures (Hunter & Schmidt, 2004). However, for this study the time between the pre and post assessment is relatively short, and can be measured in terms of weeks. Therefore, the impact of time-related confounds were anticipated to be minimal. Further limitations which also

must be recognized are a lack of comparison to a traditional NF group, and a lack of a randomized control group.

2. Small sample size. Larger sample sizes are preferred in order to allow for stronger statistical analysis and more generalizability (Gravetter & Wallnau, 2010). Given this study used pre-existing archived data, the number of samples were restricted to what was found in the files; thus there was no option to increase sample size. Though, as detailed in Chapter 3, the sample sizes for each group provided sufficient power to allow for adequate statistical analysis.

The following delimitations will be present in this study:

1. This study was delimited to the scope of the *surface* formulation of 19ZNF. Therefore, it did not include in its scope other variations of 19-channel NF models, founded in inverse solution theories, such as low-resolution brain electromagnetic tomography (LORETA) ZNF or functional magnetic resonance imaging (fMRI) tomography NF models.
2. This study was delimited to a scope of NF research data collected primarily from clinical settings, as opposed to laboratory-based experimental research.
3. The academic quality standards for this dissertation delimit the literature reviewed for this study to exclude certain non-peer-reviewed sources (i.e. NF industry newsletters).

In spite of the above stated assumptions, limitations, and delimitations, this study has potential to be of value to the scientific and neurofeedback community. Given the

data for this research comes from a real-world clinical setting, the findings of this study still contribute to advancing the scientific knowledge of 19ZNF.

Summary and Organization of the Remainder of the Study

In summary, while NF has a history spanning over 40 years, it is only now gaining acceptance as an evidence-based mental health intervention (Pigott et al., 2013). Various models of NF have been developed over the years, with one of the newest iterations including 19ZNF, which is reported to lead to improved clinical outcomes in fewer sessions than other models (Thatcher, 2013; Wigton, 2013). However, there are significant gaps in terms of peer-reviewed literature and research, such that efficacy of 19ZNF has yet to be established. This dissertation intends to fill these gaps by addressing efficacy of 19ZNF, in a clinical setting, using a comparison of pretest-posttest measures of clinical assessments and QEEG z-scores.

The following chapters include the literature review in Chapter 2 and a description of the methodology, research design, and the procedures for the study in Chapter 3. The literature review first explores the background and history of the problem, then discusses theoretical foundations and conceptual frameworks, and finally reviews the literature pertaining to the NF models relevant to this study. Of note is the necessity of a significantly expanded theoretical/conceptual section. The methodological foundations of a treatment intervention based in EEG/QEEG technology, combined with the need to explore the theoretical foundations of three different NF models (traditional, QNF, and ZNF), require more in depth coverage of the topics involved in that section.

Chapter 2: Literature Review

Introduction and Background to the Problem

The focus of this study was to explore the efficacy of 19ZNF in a clinical setting, through the use of clinical assessments and QEEG z-scores as outcome measures. Yet, a review of the literature is necessary to place this research into context of NF theory and the various models that have come before 19ZNF. This literature review consists of three sections.

The first section addresses the history and background of NF in general and specifically introduces ZNF, as well as comments on how the gap in research for 19ZNF evolved into its current form. The second section focuses on the theoretical foundations and conceptual frameworks of NF and QEEG. First, an overview of the foundations of EEG and QEEG is presented. Next, an overview of learning theory as applied to NF is discussed. Then, the theoretical frameworks supporting the different models of NF (traditional, QNF, and ZNF) are reviewed. Last, key themes of NF concepts relevant to this dissertation including applications of QNF, the development of 4ZNF, and finally the emergence of 19ZNF are examined. Also included in this section is a review of suitable outcome measures for use in ZNF research, with special attention paid to prior NF research regarding performance tests, rating scale assessments, and QEEG z-scores, as outcome measures.

Of note for this literature review is the necessity to include reviews of conference oral and poster presentations (which are subject to a peer-review acceptance process). While inclusion of these sources may be an unusual dissertation strategy, it is necessary due to the scarcity of sources in the peer-reviewed published literature regarding ZNF

models. To exclude these sources would be to limit the coverage of the available literature regarding the NF model which is the focus of this dissertation (19ZNF).

The literature for this review was surveyed through a variety of means. The researcher's personal library (from nearly fifteen years of practicing in the NF field) served as the foundation for the literature search. Then, this was expanded through online searches of various university libraries via academic databases such as Academic Search Complete, PsycINFO, PsycARTICLES, and MEDLINE, with search strings of combinations of terms such as NF, QEEG, EEG biofeedback, z-score(s). Additionally, the databases of various industry specific journals, such as the Journal of Neurotherapy, Clinical EEG and Neuroscience, as well as the Applied Psychophysiology and Biofeedback journal were queried with similar search terms. Moreover, with the specified journals, names of leading authors in the QNF and ZNF field (e.g. Koberda, Surmeli, Walker) were used for search terms.

Historical overview of EEG and QEEG. A review of NF literature reveals a common theme that the deepest roots of NF go back only as far as Hans Berger's (1929) discovery of EEG applications in humans. However, the antecedents of EEG technology can actually be traced back as far as the 1790s with the work of Luigi Galvani and the discovery of excitatory and inhibitory electrical forces in frog legs, leading to the recognition of living tissue having significant electrical properties (Bresadola, 2008; Collura, 1993). The next notable application occurred when Richard Caton (1875) was the first to discover electrical activity in the brains of monkeys, rabbits, and cats, and to make observations regarding the relationship of this activity to physiological functions (Collura, 1993). Yet for applications of EEG in humans, Berger is generally recognized

as the first to record and report on the phenomenon. Thus, it would be most correct to consider Caton as the first electroencephalographer, and Berger as the first *human* electroencephalographer (Collura, 1993). Moreover, Berger's contributions were significant as they spurred a plethora of research and technological advancements in EEG technology in the 1930s and 1940s worldwide. Of note is that Berger not only identified both alpha and beta waves, but he was also the first to recognize the EEG signal as being a mixture of various frequencies which could be quantitatively estimated, and spectrally analyzed through the use of a Fourier transform, thus paving the way for QEEG technology as well (Collura, 2014; Thatcher, 2013; Thatcher & Lubar, 2009).

Even while there was an understanding of multiple components to the EEG signal as early as the 1930s, the advent of computer technology was necessary to make possible QEEG advances (Collura, 1995); for example, the incorporation of normative databases in conjunction with QEEG analysis. Therefore, the historical landmarks of EEG developments can trace the modern start of normative database applications of QEEG back to the 1970s with the work of Matousek and Petersen (1973) as well as John (1977) (Pizzagalli, 2007; Thatcher & Lubar, 2009). However, while work exploring NF applications with QEEG technology began in the 1970s, its wider acceptance and use in the NF field was not until closer to the mid-1990s (Hughes & John, 1999; Thatcher & Lubar, 2009). Here too, advances in computer technology, whereby personal computers were able to process more data in less time, made way for advances in the clinical applications of NF.

Historical overview of NF. The historical development of neurofeedback dates back to the 1960s and early 1970s when researchers were studying the EEG activity in

both animals and humans. In these early days, Kamiya (1968, 1969) was studying how humans could modify alpha waves, and Sterman and colleagues (Sterman et al., 2010; Wyricka & Sterman, 1968) were able to demonstrate that cats could generate sensory motor rhythm, which led to the discovery that this process could make the brain more resistant to seizure activity; this eventually carried over to work in humans (Budzynski, 1999). Later, Lubar (Lubar & Shouse, 1976), expanded on Sterman's work, and began studies applying NF technology to the condition of attention disorders. This work led to an expansion of clinical applications of neurofeedback to mental health issues such as ADHD, depression and anxiety, using a training protocol generally designed to increase one frequency (low beta or beta, depending on the hemisphere) and decrease two other frequencies (theta and high beta) (S. Othmer, Othmer, & Kaiser, 1999).

Then, in the 1990s QEEG technology began gaining wider acceptance in the NF community, for the purpose of guiding the development of protocols for NF (Johnstone, & Gunkelman, 2003). The use of normative referenced databases has been an accepted practice in the medical and scientific community and the advantage it brings to neurofeedback is the allowance for the comparison of an individual to a norm-referenced population, in terms of z-scores, to identify measures of aberrant EEG activity (Thatcher & Lubar, 2009). This made possible the development of models, which focused more on the individualized and unique needs of the client rather than a one-size-fits-all model. Consequently, during the ensuing decade, the QNF model began taking hold in the NF industry. However, the primary number of channels incorporated in the amplifiers of the time was still limited to only two.

In 2006, the 4-channel – 4ZNF – technique was introduced. ZNF incorporates the application of an age matched normative database to instantaneously compute z-scores, via Joint Time Frequency Analysis (as opposed to the fast Fourier transform), making possible a dynamic mix of both real-time assessment and operant conditioning simultaneously (Collura et al., 2009; Thatcher, 2012). While the QNF of the 1990s held as a common goal movement of the z-scores in the QEEG towards the mean, the advent of ZNF brought with it the more frequent use of the term *normalizing the QEEG* or *normalization* to refer to this process. It is now generally acknowledged that the term normalization, when used to describe the process of ZNF, refers to the progression of the z-scores towards the mean (i.e. $z = 0$), meaning that the NF trainee's EEG is moving closer to the EEG range of *normal* (i.e. typical) individuals of his/her age. But by 2009 the 4ZNF model was further enhanced to include the availability of up to all 19 electrode sites in the International 10-20 system.

This surface potential 19ZNF greatly expands the number of scalp locations and measures, including the ability to train real-time z-scores using various montages such as linked-ears, averaged reference, and Laplacian, as well as simultaneous inclusion of all connectivity measures such as coherence and phase lag. This, then, makes possible the inclusion of all values from the database metrics for any given montage (as many as a total of 5700 variables) in any protocol (Collura, et al., 2009). But the advent of 19ZNF not only increases the number and types of metrics available to target, it also brought two major changes to the landscape of NF. First, it established a new model wherein the target of interest for the NF is the QEEG calculated z-scores of the various metrics (frequency/power, coherence, etc.), rather than the amplitude of particular frequency

bands (theta, beta, etc.). Second, it changed the makeup of a typical NF session. In either the conventional QNF model, or 4ZNF, the clinician will develop a protocol guided by the QEEG findings, but will generally employ the same protocol settings repeatedly for multiple NF sessions until the next assessment QEEG is scheduled. However with 19ZNF, in every session the clinician can acquire and process QEEG data, compare the pre-session data to past session data, then design an individualized z-score normalization protocol based on that day's QEEG profile, and then perform a 19ZNF session, all within an hour (Wigton, 2013). Thus, each 19ZNF session uses a protocol unique to the client's brainwave activity of that day, providing further tailoring of the NF to the individual needs of the client, on a session-by-session basis. This, then, brought a new dynamic to the normalization model of NF such that z-scores (rather than amplitude of frequencies) could be targeted, on a global basis, so as to make possible a goal of normalizing all the QEEG z-scores (when clinically appropriate) in the direction of $z = 0$.

How problem/gap of 19ZNF research evolved into current form. Over its more than 40-year history NF has frequently been criticized as lacking credible research, as evident by Loo and Barkley's (2005) critique. Nevertheless, even Loo and Makeig (2012) concede recently the research has improved. For example, Arns et al. (2009) conducted the first comprehensive meta-analysis of NF, covering 1194 subjects, concluding that it was both efficacious and specific as a treatment for ADHD, with large to medium effect sizes for inattention and impulsivity, respectively. Then, in a research review sponsored by the International Society for Neurofeedback and Research (ISNR), in what is a comprehensive review of controlled studies of NF, Pigott et al. (2013) evaluated 22 studies to conclude that NF meets the criteria of an evidence-based

treatment for ADHD. This review further documents that NF has been found to be superior to various experimental group controls, shows equivalent effectiveness to stimulant medication, and leads to sustained gains even after termination of treatment. However, as encouraging as this body of research is, it is limited in that the model covered by these studies is largely limited to one of the most traditional models of NF (theta/beta ratio NF) and only addresses a single condition of ADHD. Missing from these comprehensive reviews and meta-analyses are newer QNF models, which have been in use since the 1990s, and are frequently employed for a wider range of disorders in addition to ADHD. Yet, that is not to say that QNF is devoid of research. In fact, from 2002 to 2013 there are at least 20 studies in peer-reviewed literature covering the QNF model, yet there is great diversity in the different conditions treated in these studies, as well as a greater use of individualized, custom-designed protocols; hence making meta-analysis of this collection of research less feasible. Nonetheless, these studies do represent a body of research pointing to the efficacy of the QNF model.

Yet, when it comes to the newest models of surface ZNF, there is no such collection of research in the literature. There exist only two studies (Collura et al., 2010; Hammer et al., 2011) which evaluate sample groups of the 4ZNF model, and the Collura et al. report is mostly descriptive in nature. This, then, leaves only one experimental study. There is one dissertation on 4ZNF (Lucido, 2012), but it too is a single case study. Regarding 19ZNF, as of this writing, there are only two peer-reviewed published empirical reports specifically evaluating surface potential 19ZNF (Hallman, 2012; J. L. Koberda et al., 2012b) and those are only case study in nature.

Yet, this is not to say the peer-reviewed literature landscape is entirely devoid of any mention of surface ZNF models. Nevertheless, what does exist is mostly information *about* the technique in the form of review articles (Collura, 2008; Stoller, 2011; Thatcher, 2013; Wigton, 2013), chapters in edited books (Collura et al. 2009; Wigton, 2009), as well as numerous qualitative oral and poster conference presentations since 2008. Of note is a recent poster presentation (Wigton & Krigbaum, 2012), with a later expansion of that work (Krigbaum & Wigton, 2013), which was a multicase empirical investigation of 19ZNF; however it primarily focused on a proposed research methodology for assessing the degree of z-scores progression towards the mean. There also exist anecdotal observations in the form of case reports in non-peer-reviewed publications and internet website postings. Yet, while anecdotal observations and information from review and case study reports are helpful for initial appraisals of a new model, quantitative statistical analysis is needed to validate theories born of early qualitative evaluations, to counter a lack of acceptance from the wider neuroscience community.

Much of the focus of discussions of 19ZNF is on the potential for good clinical outcomes in fewer sessions than traditional NF (J. L. Koberda et al., 2012a; Rutter, 2011; Thatcher, 2013; Wigton, 2009; Wigton, 2013). Though, before the question of number of sessions is examined, first there should be an establishment of the efficacy of this emerging model; because empirical studies evaluating the efficacy of 19ZNF are absent from the literature. This dissertation was intended to fill this gap of knowledge, by analyzing the efficacy of 19ZNF in a clinical setting.

Theoretical Foundations

Foundations of EEG and QEEG. Hughes and John (1999) discussed a decade-long history, inclusive of over 500 EEG and QEEG related reports, the findings of which indicate that cortical homeostatic systems underlie the regulation of the EEG power spectrum, that there is a stable characteristic in healthy humans (both for age and cross-culturally), and that the EEG/QEEG measures are sensitive to psychiatric disorders. These factors made possible the application of Gaussian-derived normative data to the QEEG metrics such that these measures are independent of ethnic or cultural factors, which allow objective brain function assessment in humans of any background, origin, or age. As a result, Hughes and John assert when using artifact-free QEEG data, the probability of false positive findings are below that which would be expected by chance at a *p* value of .0025. Thus, changes in the QEEG values would not be expected to occur by chance, nor is there a likelihood of a regression to the mean of QEEG derived z-scores because EEG measures, and the corresponding QEEG values, are not random. Since the work of Hughes and John, well over a decade ago, there have been numerous studies published in the literature further demonstrating the reliability and validity of QEEGs (Cannon et al., 2012; Corsi-Cabrera, Galindo-Vilchis, del-Río-Portilla, Arce, & Ramos-Loyo, 2007; Hammond, 2010; Thatcher, 2012; Thatcher & Lubar, 2009).

Learning theory as applied to NF. As has been stated, NF is also frequently referred to as EEG biofeedback, and biofeedback has been defined simply as the process whereby an individual learns how to change physiological activity (AAPB, 2011). As Demos (2005) asserted, biofeedback is a two-way model such that 1) the physiologic activity of interest is recorded, and 2) reinforcement is provided each time the activity

occurs; as a result, voluntary control of the targeted physiologic activity becomes possible. On the surface this is a basic descriptor of operant conditioning. As a result, a common practice in the literature is for NF to be referred to only as an operant conditioning technique. However, the theoretical frameworks of NF are more correctly framed as encompassing both classical and operant conditioning mechanisms (Collura, 2014; Sherlin, Arns, Lubar, Heinrich, Kerson, Strehl, & Sterman., 2011; Thatcher, 2012; M. Thompson & Thompson, 2003). Operant conditioning – as first conceptualized by Edward Thorndike (1911) with the Law of Effect, which holds that satisfying rewards strengthens behavior, and as further advanced by B. F. Skinner (1953) – has as its primary principle when an event is reinforced/rewarded it is likely to reoccur (Hergenhahn, 2009); and for Skinner the reinforcer is anything that has contingency to a response. It is important to note that operant conditioning relates to the learning of volitionally controlled responses, motivation is necessary, and rewards need to be desired or meaningful (M. Thompson & Thompson, 2003).

In contrast, classical conditioning, established by Ivan Pavlov (1928), differs in that it deals with learning of reflexive or autonomic nervous responses. The primary mechanism is based in the associative principles of contiguity and frequency such that the presence of a dog's food, which naturally elicits a salivation reflex, when paired (contiguity) with a bell, repeatedly (frequency), will lead to the dog salivating upon the presentation of only the bell (Hergenhahn, 2009). Thus, the pairing of two previously unpaired events results in automatic learning in the form of classical conditioning. Yet, it is important to note that while operant conditioning involves volitionally oriented behavior modification, NF is a learning process which occurs largely outside of

conscious awareness; in essence, an implicit learning process (Collura, 2014). As applied to NF, the change in the EEG, as reflected in brainwave frequencies, patterns, or z-scores, is the behavior which is modified as a result of the combined classical and operant conditioning occurring in the NF session (Collura, 2014).

In this context then, successful NF involves a motivated trainee experiencing the repeated pairing of meaningful auditory and/or visual reward signals, when the recorded brainwaves fall in a targeted range. The reward signal is typically in the form of an auditory tone (beep, chime, music) in combination with an animated visual display (simple game-like displays or movies), which when aesthetically pleasant to the trainee enhances and promotes the process. Some have noted the importance of additional learning theory components such as shaping (Collura, 2014; Sherlin et al., 2011; M. Thompson & Thompson, 2003), anticipation of future rewards (Thatcher, 2012), and secondary reinforcers (Sherlin et al., 2011; M. Thompson & Thompson, 2003) to further inform NF to varying degrees. These variations as applied to NF have served to generate a range of NF models over the years; however the basic foundations of classical/operant conditioning remain constant in all the models.

Traditional/amplitude-based models of NF. In NF, when the EEG is divided into different frequency bands (alpha, beta, etc) the amplitude measures *how much* of that frequency is present within the total EEG spectrum recording. The basic goal of amplitude NF treatment models is to either increase or decrease the amplitude of a particular frequency. These models are the longest-standing conceptualization of NF techniques and for that reason, for the purposes herein, the term *traditional* will be used to refer to these models of NF. The earliest traditional model of NF started with Kamiya's

(1968) discovery in the early 1960s that human alpha waves could be increased and trained to occur for increased periods of time. Next, Sterman and Fiar (1972) followed up on this work by expanding the training Sterman had been conducting with cats to include humans, with the first known case of resolving a seizure disorder in a person using NF. In this model the goal was to increase the beta frequency of 12-15 Hz, also referred to as sensorimotor rhythm (SMR), along the sensorimotor cortex of the brain. Others then expanded on this model. For example, Lubar believed the model Sterman developed would be applicable to children with attention disorders (Robbins, 2000). After a year-long academic fellowship with Sterman, he moved on to develop his own model which incorporated decreasing the theta frequency in addition to increasing beta (Robbins, 2000). Lubar and Shouse (1976) reported on the first use of this approach, which was the foundation for what would become one of the most commonly reported and researched protocols (for use with attention disorders) in the literature since the early 1990s; that of the theta/beta ratio model.

Another example of a traditional NF model with roots to Sterman's efforts is the Othmer model (S. Othmer, Othmer, & Kaiser, 1999), employing a combination of increasing beta (either 12-15 Hz or 15-18Hz) together with decreasing theta (4-7 Hz), and a higher beta band (22-30 Hz); again with electrode placements primarily along the sensorimotor cortex locations of the scalp. In the years since its introduction, there have been different modifications and variations of the Othmer approach (S. F. Othmer & Othmer, 2007). Nevertheless, consistent with traditional NF, this model makes use of targeting the amplitudes of frequency bands in particular directions (i.e. make more or less of targeted frequencies).

While some built models based in the original findings of Sterman, others expanded on Kamiya's work, by developing models which targeted the increase of alpha and/or theta frequencies (in parietal brain regions) to enhance relaxation and creative states (Budzynski, 1999). Peniston and Kullkosky (1990, 1991) developed applications of these approaches, which led to treatment models for alcoholism and posttraumatic stress disorders. Yet still others, such as Baehr, Rosenfeld, and Baehr (1997), established protocols targeted to balance alpha in the frontal regions as a treatment for depression.

While each of the above models targeted different frequencies with a variety of protocols, consistent was a focus on changing the amount of the brainwave of interest; the desired outcome is either greater or lesser amplitude of a target frequency. Moreover, pre-treatment assessment of EEG activity to inform NF protocols is limited to nonexistent in the majority of these models, with a typical one-size-fits-all approach. While selecting the particular NF model for a treatment approach (i.e. theta-beta ratio versus alpha-theta training) is informed by the presenting symptoms of each case, personalizing a NF protocol to address the individual brainwave patterns of the client is not the focus of these approaches.

QNF model of NF. A key focus of QNF is precisely tailoring the NF protocol, based on the individual EEG baseline and symptom status of the client, as determined by the QEEG, in conjunction with clinical history and presenting symptoms (Arns et al., 2012). The primary premise of this approach is that localized cortical dysfunctions, or dysfunctional connectivity between localized cortical areas, correspond with a variety of mental disorders and presenting symptoms (Coben & Myers, 2010; Collura, 2010; Walker, 2010a). When the EEG record of an individual is then compared to a normative

database representing a sample of healthy individuals, the resulting outlier data (deviations of z-scores from the mean) help link clinical symptoms to brain dysregulation (Thatcher, 2013). For example, when an excess of higher beta frequencies are found, the typical associated symptoms include irritability, anxiety, and a lowered frustration/stress tolerance (Walker, 2010a).

The conceptual framework of the stability of QEEG, as noted above, applies to QNF in that a stable EEG is not expected to change without any intervention, thus the changes seen as a result of QNF is not occurring by chance, but due to the operant conditioning of the brainwaves as a result of the NF process (Thatcher, 2012). Therefore, in the example of excess beta frequencies, when the symptoms of anxiety and irritability are resolved after QNF, and the post QEEG shows the beta frequencies to be reduced (closer to the mean), it is assumed the improvement in symptoms is due to the change in the QEEG; thus representing improved electrocortical functioning (Arns et al., 2012; Walker, 2010a). The term for this process, which has arisen secondary to QNF, is generally referred to as *normalization of the QEEG*, or simply *normalization* (Collura, 2008; Surmeli & Ertem, 2009; Walker, 2010a). Consequently, the concept of normalization is generally accepted to be when the z-scores of the QEEG move towards the mean (i.e. $z = 0$).

It is also important to note that the QNF model, with its reliance on the QEEG to guide the NF protocol, embraces the heterogeneity of QEEG patterns as discussed by Hammond (2010). In understanding that a particular clinical symptom presentation may be related to varied deviations in the QEEG, it quickly becomes apparent that each NF protocol needs to be personalized to the client; as well as monitored and modified for

maximum treatment effect (Surmeli et al., 2012). This, then, results in different electrophysiological presentations being treated differently, even if the overarching diagnosis is the same. This clinical approach is supported through multiple reports in the literature discussing how training the deviant z-scores towards the mean (i.e. normalize the QEEG) in QNF results in the greatest clinical benefit (Arns et al., 2012; Breteler et al., 2010; Collura, 2008; Orgim & Kestad, 2013; Surmeli et al., 2013; Surmeli & Ertem, 2009, 2010; Walker, 2009. 2010a, 2011, 2012a).

However, while the personalization of NF protocols aids in greater specificity in client treatment, it creates methodological challenges for researching QEEG based NF models; which will be discussed further below. When boiling down the elements of study to a lowest common denominator, overall normalization of the QEEG is the only common point of measurement. Therefore a reasonable tool, as a measure of change in the QEEG, would be a value reflecting the change of targeted z-scores for a particular metric.

In summary then, in the normalization model of QNF, when the QEEG data show excessive deviations of z-scores, and those deviations correspond to the clinical picture, the NF protocol is targeted to train the amplitude of the frequency in the direction of the mean (i.e. create more or less energy within a specified frequency band). In other words, if the QEEG indicates an excess of a beta frequency (i.e. high z-scores), and the presenting symptoms are expected with that pattern (i.e. anxiety), the protocol would be designed to decrease the amplitude of that beta frequency. Conversely, if the QEEG indicates a deficit of an alpha frequency, with corresponding symptoms, the protocol would be designed to increase the amplitude of the alpha frequency. The QNF model

then, is simply traditional amplitude based NF using the QEEG to guide the protocol development for the NF sessions.

ZNF model of NF. The ZNF model leverages the statistical underpinnings of a normal distribution, where a value converted to a z-score is a measure of the distance from the mean of a population, such that the mean represents a range considered to be normal (or typical) (Collura, 2014). With ZNF the real-time QEEG metrics are incorporated into the NF session using a joint time frequency analysis (rather than fast Fourier transform) to produce instantaneous z-scores, which allows for real-time QEEG assessment to be paired with operant conditioning (Collura, 2014; Thatcher, 2013). Therefore, where the QNF model has amplitude (as guided by the QEEG) as its targeted metric, in its most basic form, the ZNF model targets the calculated real-time z-scores. Yet, that being said, it is important to note that the z-scores can be considered a meta-component of EEG metrics (i.e. amplitude or connectivity) and ultimately, even when z-scores are targeted, the underlying EEG components are still being trained.

Nevertheless, directly targeting z-scores results in a different dynamic in the NF training protocol. The goal is no longer to simply make more or less frequency amplitude, but for the targeted excessive z-score metrics (whether high or low) to move towards the mean, that is to normalize. Thus, there is a greater focus on the construct of normalization. A second change is the inclusion of many more metrics to target. ZNF makes available simultaneously, for up to ten frequency bands, both absolute and relative power, ratios between frequencies (i.e. theta/beta ratio or alpha/beta ratio), as well as the inclusion of connectivity metrics such as asymmetry, coherence, or phase lag, all as active training metrics. Therefore, when applied to 4ZNF, the maximum number of

metrics to train is 248 (Collura, 2014) and, within the scope of the 19ZNF the maximum number of metrics is 5700 (Collura, et al., 2009). These changes make the entire range of all QEEG metrics, or a subset of selected metrics, available for targeting with ZNF models. Moreover, the increased number of metrics targeted by 19ZNF may allow for an increase in regulation and synchronization of neural activity simply by the greater number of training variables. Nonetheless, one consistent theme remains aligned with the QNF model, in that the decision to target normalization of QEEG metrics is determined by the presenting clinical symptoms; thus when QEEG deviations correspond to presenting symptoms, normalization is a reasonable treatment goal.

In asking if the 19ZNF improves attention, behavior, executive function, or electrocortical function, the research questions for this study add to what is known regarding whether operant conditioning with 19ZNF, produces clinical results that are comparable to those reported in the literature for traditional or QNF models. Moreover, this study also evaluates questions regarding 19ZNF and normalization of QEEG metrics. This research fits within the overarching NF model with a specific focus on evaluating efficacy of the ZNF model. As has been demonstrated in the literature, traditional NF is well researched (Arns et al., 2009; Pigott et al., 2013), and as will be discussed in the next section, the QNF model is well addressed in the literature. Conversely, as will be seen, the ZNF models (4ZNF and 19ZNF) are still minimally represented in the literature. Therefore, this study addresses an area which calls for further research.

Review of the Literature – Key Themes

QNF in the literature. Beginning with QNF models in reviewing the NF literature is applicable in that the QNF model laid the ground-work for the ZNF models

that followed. Both QNF and ZNF models hold the generalized goal of normalizing the QEEG, and for that reason, QNF is chosen as the first key theme in reviewing NF in the literature. With few exceptions, literature presented on the QNF model comes from research conducted in clinical settings. As a result, given the ethical constraints of conducting research in clinical settings (e.g. asking clients to accept sham or placebo conditions) (Gevensleben et al., 2012) few are blinded and/or randomized-controlled studies.

Arns et al. (2012) conducted a well-designed open-label study of 21 ADHD subjects using the QNF model, incorporating pre-post outcome measures and QEEG data. The purpose was to investigate if the personalized medicine approach of QNF was more efficacious (as defined by effect size) for ADHD than the traditional theta/beta or slow cortical potential models, as reported in his meta-analysis three years earlier (Arns, et al., 2009). The outcome measures incorporated were a self-report scale based on the Diagnostic and Statistical Manual-IV list of symptoms and the Beck Depression Inventory. The findings of this study were statistically significant improvements ($p \leq .003$) in both the attention (ATT) and hyperactivity (HI) subtypes of ADHD symptoms as well as depression symptoms. In this study, the mean number of sessions was 33.6, and the effect size was 1.8 for the ATT subtype, and 1.2 for the HI subtype; this was a substantial increase over the traditional model effect sizes of 1.0 (ATT) and 0.7 (HI) respectively. This suggests the QNF model is more efficacious (i.e. effect size of clinical improvements) than the older traditional theta/beta or slow cortical potential models. Furthermore, in this study, non-z-score EEG microvolt data was reported for only nine frontal and central region electrode sites, and three frequency bands, on a pre-post basis.

In addition to that the protocols employed are described as a selection of one of five standard protocols, with QEEG informed modifications. The limitations of this study were few but include a lack of a control group, a fairly small sample size, and that some outcome measures were collected on only a sub-group of participants (thus reducing net sample size). Moreover the pre-post QEEG data analysis was limited.

J. L. Koberda, Hillier, et al., (2012) reported on the use of QNF in a clinical setting of a neurology private practice. All 25 participants were treated with at least 20 sessions of a single-channel traditional NF protocol, which was guided by QEEG data and symptoms, with a goal to improve symptoms and normalize the QEEG. Clinical improvement was measured by subjective reports from the participants in the categories of not sure (n = 4), mild if any (n = 1), mild improvement (n = 3), improved/improvement (n = 13), much improved (n = 2), and major improvement (n = 2); with a total of 84% (n = 21) reporting some degree of improvement. QEEG change was reported as a clinical subjective estimation (based on visual inspection of the QEEG topographic images) of change in the targeted frequencies, in the categories of no major change/no improvement (n = 6), mild improvement (n = 9), improvement (n = 8), or marked improvement (n = 1), and one participant not interested in post-QEEG; with a total of 75% (n = 18) showing estimation of improvement in the QEEG. Of note with this study was the heterogeneous collection of symptoms treated which included ADD/ADHD, anxiety, autism spectrum, behavior symptoms, cognitive symptoms, depression, fibromyalgia, headaches, major traumatic brain injury, pain, seizures, stroke, and tremor, in varying degrees of comorbidity per case. However, the primary limitation of this study was the loosely

defined subjective estimations of improvement for both clinical symptoms and QEEG outcomes.

In their randomized control study, Breteler et al. (2010) evaluated QNF as an additional treatment with a linguistic education program. From the total sample of 19, ten participants were in the NF group and nine were in the control group. Individual NF protocols were based on QEEG results and four rules, with a generally (though not strictly adhered to) 1.5 z-score cutoff; which resulted in the use of eight personalized protocols. Improvement was determined by results of outcome measures of various reading and spelling tests, as well as computerized neuropsychological tests. Paired *t t*ests were applied for analysis of the difference values between the pre and post scores. The reported findings showed the NF group improved spelling scores with a very large Cohen's *d* effect size of 3; however no improvement in reading or neuropsychological scores. QEEG data was reported, in terms of pre-post z-scores, on an individual basis (i.e. per each case) for a limited number of targeted sites, frequencies, and coherence pairs; with most showing statistically significant normalization.

In a retrospective study using archived clinical case files, Huang-Storms, Bodenhamer-Davis, Davis, and Dunn (2006) evaluated the efficacy of QNF for 20 adopted children with a history of abuse who also had behavioral, emotional, social, and cognitive problems. The children all received 30 sessions of NF (from a private practice setting) with QNF protocols, which were individualized based on the QEEG profiles. Data from the files of 20 subjects were collected to include pre and post scores for outcome measures from a behavioral rating scale (Child Behavior Checklist; CBCL), and a computerized performance test (Test of Variables of Attention; TOVA). The findings

for the CBCL were statistically significant ($p < .05$) for most scales and the TOVA findings were statistically significant ($p < .05$) for three scales, thus demonstrating QNF efficacy for the subjects in this study. There was no quantified QEEG reported; only observations of general trends in the pretreatment QEEG findings, such as excess slow waves in frontal and/or central areas.

Two researchers are most notable for several published studies evaluating the QNF model, that being Walker and then Surmeli and colleagues. Each has a particular consistent style in structuring their studies; and both have reported on the use of QNF with a wide variety of clinical conditions. Therefore their works will be reviewed in a grouping format. Walker has reported on mild closed head injury (Walker, Norman, & Weber, 2002), anxiety associated with posttraumatic stress (Walker, 2009), migraine headaches (Walker, 2011), enuresis (Walker, 2012a), dysgraphia (Walker, 2012b), and anger control disorder (Walker, 2013). His QNF protocol development centers on tailoring the protocol to the individual clinical QEEG data, with some restrictions of either increasing or decreasing the amplitude of certain frequency ranges. For example, the protocols for the anger outburst study restricted the target range to decrease only excess z-scores of beta frequencies, combined with decreasing excess z-scores of 1-10 Hz frequencies. For the migraine and anxiety/posttraumatic stress studies both were based on individual excess z-score values found in the beta frequencies in a range of 21-30 Hz (to decrease) with an addition of increasing 10 Hz. For all studies the electrode sites selected were ones where the deviant z-scores in the targeted range were found. In the mild closed head injury article, the protocol was different because the study was meant to evaluate coherence training with a stated goal to normalize coherence z-scores. Thus, the most

deviant coherence pair was selected first (for five sessions each) and, then progressed to lesser deviant pairs until the symptoms resolved or until 40 sessions were completed. None of Walker's reports declare a particular research design; still all involve pretest-posttest comparisons of various outcome measures.

The outcome measures that Walker typically employs are primarily Likert or percentage-based self-reports, except in the anger control disorder study where the DeFoore Anger Scale self-report instrument was used to track the number of anger outbursts. However, while all protocols are personalized, and based on QEEG findings, there are no quantified pre-post QEEG data used as an outcome measure, and none are reported in his studies. Overall the findings of all of Walker's studies show improvements in the targeted clinical conditions. In the mild closed head injury study, with an $n = 26$, 84% of the participants reported greater than 50% improvement in symptoms. For the anxiety/post-traumatic stress article, with an $n = 19$, all improved on a Likert scale (1 - 10; 10 being *worst*) from an average rating of 6 before NF treatment to an average rating of 1 after NF treatment. With the migraine study, where 46 NF participants were compared to 25 patients who chose to remain on medication, 54% had complete remission of headaches, 39% had a greater than 50% reduction, and 4% experienced less than 50% reduction in migraines, all in the NF group, while in the medication group, 84% had no change in migraines and only 8% had a greater than 50% reduction in headaches. In three of his more recent studies, for the enuresis ($n = 11$), dysgraphia ($n = 24$), and anger control research ($n = 46$), Walker reported all findings for all participants (in all three studies) showed statistically significant improvement at $p < .001$.

Surmeli and colleagues reported on Down syndrome (Surmeli & Ertem, 2007), personality disorders (Surmeli & Ertem, 2009), mental retardation (Surmeli & Ertem, 2010), obsessive compulsive disorder (Surmeli & Ertem, 2011), and schizophrenia (Surmeli et al., 2012). Notable in this collection of work are conditions previously not known to respond to NF, such as personality disorders, mental retardation, Down syndrome, and schizophrenia. All of these studies report the QNF protocol as being individualized, as informed by a combination of the QEEG findings and clinical judgment; with an overall goal to normalize the QEEG patterns. Notable for most of Surmeli et al. studies are a high number of sessions reported for the cases; ranging from an average of 50 to an average of 120 sessions. No particular research design is declared in the Surmeli et al. studies, but here too, comparisons of pretest-posttest outcome measures are reported.

The outcome measures in the studies mentioned above generally make use of clinical assessment instruments designed to measure the symptoms targeted for the QNF treatment. For example, the schizophrenia study employed the Positive and Negative Syndrome Scale (PANSS), and for the obsessive compulsive disorder research they incorporated the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS). For many studies, the computerized performance Test of Variable Attention (TOVA) was used. Yet, as with Walker's work, in spite of all protocols being individually QEEG-guided, QEEG data is not used or reported as an outcome measure; only observations of general trends of the changes in QEEGs are discussed. However, the targeted clinical symptoms, as measured by the clinical assessments, were reported as having statistically significant improvement in all studies. For the personality disorder study, with an $n = 13$, twelve were significantly

improved on all outcome measures; with the Symptom Assessment 45 Questionnaire at $p = .002$, the Minnesota Multiphasic Personality Inventory (MMPI) Psychopathy scale at $p = .000$, and the TOVA at $p < .05$ on the visual and auditory impulsivity scales. With the article reporting the study with mentally retarded participants, including an $n = 23$, for 19 there was improvement on the Wechsler Intelligence Scale for Children-Revised (Verbal scale, $p = .034$; Performance scale, $p = .000$; Total scale, $p = .000$) and the TOVA (Auditory and Visual Omission scale, $p < .02$; Auditory and Visual Commission scale, $p < .03$; Auditory and Visual Response Time Variability scale, $p < .03$). In the Down syndrome study, while the outcome measure was not a commercialized assessment, they did develop a questionnaire formulated to evaluate symptoms associated with Down syndrome. The findings were that all subjects in the study ($n = 7$) showed improvement at $p < .02$ on all questionnaire scales. With QNF for obsessive compulsive disorder, with an $n = 36$, 33 showed improvement on the Y-BOCS (Obsession subscale, Compulsion subscale, and Total score all $p < .01$). Finally, in the schizophrenia study, with an $n = 51$, 47 out of 48 patients who completed pre and post PANSS improved on all scales at $p < .01$. Moreover of the 33 who were able to complete the MMPI, findings showed significant improvements ($p < .01$) on the scales of Schizophrenia, Paranoia, Psychopathic Deviation, and Depression.

This review of QNF research fits within this dissertation topic as examples of how prior studies with QEEG data have been addressed in the literature. As can be seen, studies evaluating QNF are typically found in clinical settings, with a wide variety of clinical symptoms and/or mental health diagnoses, and frequently have relatively small sample sizes. Moreover the NF protocols employed typically are tailored to the

individual, informed by QEEG, with a goal to normalize the QEEG. The overwhelming majority of clinical QNF research employs retrospective pre-post comparison research designs and the outcome measures used are tied to the symptoms of investigation. Yet few, if any, report pre-post QEEG metrics, and only one (Arns et al., 2012) incorporated statistical analysis of QEEG metrics as an outcome measure (and that was to a limited degree). Therefore, in the QNF literature, it has become an accepted practice to define efficacy in terms of measuring symptom improvement with various clinical assessments (both commercially and informally developed). Nevertheless, clearly there is a gap in the reporting of group QEEG z-score mean data in the present QNF research.

4ZNF in the literature. Given that 4ZNF is the forerunner to 19ZNF, this topic is explored to provide historical context on both its development and its coverage in the literature. While there are numerous studies in the literature for QNF, when it comes to ZNF studies, such is not the case. However, for the 4ZNF model there are four representations of 4ZNF clinical results in the literature.

In a first poster presentation on the topic, Wigton (2008) presented a single case study where 4ZNF was used with an adult to address a diagnostic history of ADHD, Bipolar disorder, and anxiety symptoms. The primary pre-post outcome measure was the IVA. Also included were topographic images of pre and post QEEG assessments. After 25 sessions of 4ZNF, in addition to multiple subjective reports of symptom improvement from the participant, the scaled scores for the IVA showed marked improvement. The full scale Response Control scale improved from 29 to 94, and the full scale Attention scale from 0 to 96. The QEEG findings (as reported by visual presentation of QEEG topographic images) showed improvements in terms of normalization in the QEEG, most

noticeably in the left frontal delta and theta frequencies, as well as coherence and phase lag normalization. However, a limitation of this study was a lack of statistical analysis of pre-post QEEG data and the use of only one clinical assessment for outcome measures.

Collura et al. (2010) was the first peer-review publication addressing 4ZNF although its organization was a loosely structured collection of clinical reports from six clinicians covering 24 successful cases. Nonetheless, for a model with little scientific evidence, it does stand as the only representation in the literature of a multiple-clinician report of clinical results with 4ZNF. All cases reported clinical improvement, with no abreactions, and the average number of sessions for all cases presented were 21.1. The limitations of this case study are the lack of a structured methodology, no statistical analysis, and limited pre-post outcome measures and/or QEEG data.

The study conducted by Hammer et al. (2011) represents, to-date, the only quantitative analysis of 4ZNF. Its strength is a sound methodology with a randomized, parallel group, single-blind design, together with QEEG z-scores as an outcome measure. Though, the setting for this research was not in a clinical setting, but rather a university psychophysiology laboratory wherein participants were recruited specifically for the study. The purpose was to both explore 4ZNF as a new NF model, and to evaluate the efficacy of two different 4ZNF protocols for insomnia. The primary findings suggest that 4ZNF may be a beneficial treatment for insomnia. While this study had very small group sample sizes (n = 5 and n = 3) all insomnia related outcome measures resulted in pre-post treatment improvement in symptoms, and normal (or near normal) sleep was achieved by all participants. Moreover, at follow-up 6 to 9 months after treatment, over half sustained the treatment response. The findings of this study included QEEG measures showing

statistically significant electrocortical change occurred for the delta frequency ($p < .001$) and beta frequency ($p < .01$), but not high beta ($p < .11$). However, a limitation is that the reported findings only included three frequencies, and the absolute and relative power z-scores were combined in the analysis; therefore a more discrete picture of overall QEEG normalization was not available. Further limitations of this study were the small sample size and the lack of control group. Yet this study does stand alone, being a peer-reviewed publication, as an example of a quantitative methodology for measuring normalization of QEEG z-scores with the binomial test of significance, with the 4ZNF model.

A dissertation conducted by Lucido (2012) was a single case study to evaluate the use of 4ZNF for an adult with Autism spectrum condition (ASC). This study used a multiple baseline design, with five rounds of assessment data gathered before the 4ZNF sessions, and a round of assessments at five incremental points during/after the NF treatment. The outcome measures employed were the Neuropsych Questionnaire, the CNS Vital Signs computerized neurocognitive assessment, and the Test of Nonverbal Intelligence. While QEEG data was gathered and purported as an outcome measure, only limited pre-post colorized topographic images were provided as a means to demonstrate generalized changes in QEEG metrics. The results were that, with only one exception (cognitive processing speed), all symptoms assessed with the outcome measures improved. These included ASC symptoms, executive function, depression, anxiety, mood stability, attention, and intelligence. To the study's credit, this was a well-designed, well-controlled case study; however still a representation of a single case, nonetheless.

Overall, the 4ZNF model is poorly represented in the NF literature. However, there are still themes relevant to this dissertation. Of the studies reported, most are from

clinical settings. Moreover, clinical assessments, as outcome measures, are used in all studies. A particular stand out, though, is the Hammer et al. (2011) research, wherein statistical analysis of QEEG metrics was used as an outcome measure.

19ZNF in the literature. With 19ZNF being the focus of this study, reviewing what literature is available is necessary. Yet, there is an even greater dearth of published literature for 19ZNF than 4ZNF. Therefore a review of conference oral and poster presentations is necessary to sufficiently address what is known regarding 19ZNF. Moreover, the literature reviewed herein is restricted to evaluative and/or case study research reports regarding clinical applications of 19ZNF (rather than technical reviews of 19ZNF).

In a first published clinical review of 19ZNF, Wigton (2009) reported initial findings in which substantial QEEG normalization and clinical improvement was achieved in as little as three sessions. While research into this technique was clearly needed, the degree of success achieved in just a few sessions was a novel finding for previously known NF models. Later in a conference presentation, Wigton (2010a) reported on a series of case reviews that employed the Laplacian montage with 19ZNF. There were 10 cases which included conditions such as anger issues, anxiety, ADHD, and impaired cognition. The findings were that 19ZNF led to clinical improvements and QEEG normalization, in less than 10 sessions, in seven out of the 10 cases. In this presentation outcome measures included the IVA, the DSMD, and Likert scale reports. A year later Rutter (2011) described, in a conference presentation, her use of 19ZNF and how she was able to see initial indications of QEEG normalization in as little as five sessions.

In their conference oral presentation, J. L. Koberda, et al. (2012a) reported on a comparison between 25 cases using traditional 1-channel NF and a mixed pool of 15 cases using either surface 19ZNF or LORETA ZNF. However, it is not clear how many were 19ZNF and how many were LORETA ZNF cases. In this presentation the clinical symptoms addressed in the 15 cases was varied and included anxiety, headaches, chronic pain, cognitive and behavioral disorders, as well as focal neurological disorders. The essential finding of this presentation was that both the traditional single-channel NF and the 19ZNF/LORETA ZNF lead to improvement in clinical symptoms and improvements in QEEG measures, but the 19ZNF/LORETA ZNF did so in fewer sessions. The traditional NF group showed subjective self-report improvements of 84% and an improvement of 75% of QEEG improvements, whereas the 19ZNF/LORETA ZNF group showed 95% subjective improvement and 62.5% improvement in QEEG measures. However an operationalized definition of these improvements was not clearly described or quantified; nor were there any follow-up data reported. Nevertheless, the number of sessions for the traditional NF was at least 20, whereas the number for the 19ZNF/LORETA ZNF group was an average of nine sessions.

Hallman (2012) presents a qualitative style clinical review of a single case study, of a child with fetal alcohol syndrome. The purpose of the article was to describe the case wherein 80 sessions of 19ZNF resulted in unexpectedly remarkable symptom and behavior improvements. Moreover, the topographic images of pre-post QEEG data also showed almost complete normalization; still there was no quantified measurement or statistical analysis of QEEG data. There also were only subjective parental reports and no outcome measures to quantify degree of symptom improvement.

J. L. Koberda et al. (2012b) also conducted a single case study, of a 23 year-old male, for the purpose of reporting clinical outcomes using two types of 19ZNF (surface and LORETA). After only 15 sessions, improvements in a cognitive assessment outcome measure were achieved, still there were no inferential statistical analysis reported for the pre-post outcome measures. Moreover, the use of two distinctly different 19ZNF modalities (surface and LORETA ZNF) makes it hard to know if one better accounted for the improvement over the other. Finally, while improvements in QEEG data were reported, again no inferential statistical analyses of these improvements were presented.

Krigbaum and Wigton (2013) present findings for 10 cases with 19ZNF. This study is notable in that it introduced a proposed methodology for statistically demonstrating z-score progression towards the mean (i.e. $z = 0$), and an approach for plotting individual learning curves as a result of 19ZNF. Additionally, cases in the study included outcome measures such as the IVA, DSMD, BRIEF and Likert scale (reported on a supplementary basis, with only an indication of improvement or not), and all outcome measures showed improvement at case completion. Repeated measures analysis of variance (rANOVA) and paired *t t*ests supported all three research questions such that the z-scores progressed towards the mean (rANOVA absolute power, $p < .001$; relative power, $p < .04$; coherence, $p < .001$); the post z-scores were closer to the mean than the pre z-scores (paired *t t*est absolute power, $p < .007$; relative power, $p < .05$; coherence, $p < .03$); and clinical improvement was reported in all cases. However, no follow-up data was reported.

Clearly, the research evaluating 19ZNF is in its infancy and there is a great need for scientifically sound investigations. More so, the research needs to move beyond

clinical reviews and case studies. As is incorporated in QNF research, use of clinical assessments as outcome measures are important elements; additionally, finding ways to include QEEG metrics as outcome measures would benefit 19ZNF research.

Outcome measures for ZNF research. This topic is included to explore outcome measures that are suitable for ZNF research. A good deal of NF research occurs in clinical settings, where assessment instruments are employed as part of the case workup. As such, the use of those same measures after treatment is a natural fit for what are frequently pretest-posttest research frameworks. Other than informal self-reports (i.e. Likert scales) two types of popular outcome measures found in the NF literature are rating-scale type assessments and computerized performance tests. Moreover, commonly found in NF studies is the use of multiple outcome measures. Further, while the use of EEG metrics as outcome measures of electrocortical change are infrequently incorporated in NF research, there are a few reports in the literature which will be reviewed.

***Computerized performance tests*.** Computerized performance tests are common outcome measures in NF research, usually as a means to evaluate attention-related symptoms associated with ADHD. One of those instruments is the IVA. While the IVA was designed as a diagnostic aid for ADHD, the manual provides usage indications to include assessing self-control and attention problems related to other disorders such as depression, anxiety, head injuries, dementia, and other medical problems (Sanford & Turner, 2009). Several NF studies have incorporated the IVA as an outcome measure to assess attention related symptoms.

In their study to evaluate NF in a nonclinical group of college students' cognitive abilities, Fritson, Wadkins, Gerdes, and Hof (2008) used the IVA as one of their outcome

measures; each group (experimental and control) had an $n = 16$. The stated objective was to determine effects of NF on attention, impulsivity, mood, intellectual functioning, emotional intelligence, and general self-efficacy. The IVA was one of several outcome measures and was included to assess response control (i.e. impulsivity) and attention. The researchers reported results in terms of the means and standard deviations of pre-post values of eight of the primary scales of the instrument. The statistical analysis performed were multivariate analysis of variance (MANOVA) between the control and experimental groups.

In evaluating the utility of the Tower of London test, as a suitable assessment instrument for clients with Asperger's who undergo NF, Knezevic, Thompson, and Thompson (2010) employed the IVA as one of the outcome measures. They included six scales of the IVA (Auditory and Visual Prudence, Auditory and Visual Vigilance, and Auditory and Visual Speed) to assess the efficacy of NF, and evaluate the measure of impulse control as compared to the Tower of London test. The number of subjects reported for the IVA varied for the different scales used from a low of $n = 6$ to a high of $n = 12$, because they only included for analysis cases where pre-test scores needed to improve. The researchers reported the means and standard deviations of the pre-post values of the included scales, and performed paired *t* tests for statistical analysis.

Steiner, Sheldrick, Gotthelf, and Perrin (2011) conducted a randomized controlled study with 41 children, comparing NF to a standardized computer attention training program and used four outcome measures including the IVA. However, they only included for analysis the two most broadly defined full-scale components of Response Control and Attention, and only reported on an $n = 6$ for the NF group, and an $n = 10$ for

the computerized training group. Repeated measures ANOVAs were performed to analyze the pre-post outcome measures in this study.

Rating scales. Rating scale instruments are one of the most common assessment tools found in NF literature for measuring clinical outcomes. Rating scales are instruments which require rated objects to be assigned to categories or numerical continua, by the rater or observer, based on their perception or remembrance of the behavior being rated (Kerlinger & Lee, 2000). Rating scales frequently employed in NF literature include the BRIEF, the Conner's' Rating Scale-revised (CRS-R), the Behavior Assessment Scale for Children (BASC), and the Child Behavior Checklist (CBCL). The following are examples from the literature of their use in NF studies.

In a randomized study, Orgim and Kestad (2013) compared NF to medication for a heterogeneous ADHD group with various comorbidities; each group had an $n = 16$, and the NF group was administered 30 NF sessions. The outcome measures included the rating scales of CRS-R and BRIEF. They conducted analysis of covariance (ANCOVA) statistical tests, using baseline measurement (Time-1) as the covariate; and they analyzed group differences at Time-2 for selected scaled scores.

The study of Huang-Storms et al. (2006) provided an example of the use of rating scales, in a retrospective clinical study, in the form of the CBCL together with a computerized performance test. The total number of valid CBCLs reported on was an $n = 18$, and all aforementioned scales were included in the analysis. The statistics employed were two-tailed paired *t t*est analysis.

Drechsler et al. (2007) conducted a study with an experimental design to assess the efficacy of slow cortical potential NF with ADHD using multiple outcome measures;

where the experimental group had an $n = 17$ and the control group had an $n = 13$. Here they employed two rating scales: The CPS-R and the BRIEF. Moreover, they only included the composite or global scales from these instruments and performed repeated measures MANOVAs for analysis.

In a randomized control study, Steiner et al. (2011) compared traditional NF to computerized attention training to a waitlist control group; the group sizes were $n = 9$, $n = 11$, and $n = 15$. In this study, they used three rating scales: the CRS-R, the BASC, and the BRIEF. Here too, they included selected scales from the assessments for analysis. The statistics applied were rANOVAs, in an effort to detect if the experimental conditions resulted in greater effects for the post NF assessment over the control group.

QEEG z-scores. As has been stated, with the QNF studies, by far, the vast majority did not use pre-post EEG metrics or z-scores as an outcome measure. Though, equally so, few traditional NF studies included EEG values as an outcome measure. Yet, in one study purported to evaluate EEG effects of NF, Gevensleben et al. (2009) reported values, as grouped together for nine regions across the scalp, and four frequency bands. The averages of the microvolt values (raw, non z-score EEG values) were computed for each region and frequency band, and post values minus pre values were used as a measure of change. Since this was a study for traditional/amplitude NF, no z-score metrics were used. Further, there were no goals of normalization in the NF protocols.

Two QNF studies do stand out for reporting, to some degree, pre-post EEG metrics as part of the research. With Arns et al. (2012), non z-score pre-post EEG microvolt data was analyzed, but for only nine sites, exclusive to frontal and central areas, and for just three power frequencies. The group data was averaged, and presented

in a graph, for each site and frequency combination. Statistically significant pre-post differences were noted for this data. The second QNF study (Breteler et al., 2010), did report some pre-post z-scores information, but it was lacking in depth. The QEEG data were reported for a limited number of sites and frequencies, as well as coherence pairs, presumably as identified from the personalized training protocols.

Hammer et al. (2011) presented a unique offering in performing the binomial test of significance to evaluate z-scores as an outcome measure of normalization. While the results did show a statistically significant number of z-scores normalized after 4ZNF, the findings were for only three frequencies (delta, beta, and high beta), and combined values for absolute and relative power. Moreover, this methodology is limited in that it only provides a yes/no level of analysis for normalization, not a discrete measure of change towards the mean. Nonetheless, it is a useful offering in an effort to present a measure of normalization of the QEEG in response to 4ZNF.

One reason for the lack of reporting of z-scores as outcome measures may be due to the nature of z-scores encompassing both positive and negative values, which, when averaged, tend to cancel out a magnitude of effect. This was noted in Ramezani's (2008) dissertation, which was a study comparing pre and post z-scores of coherence and phase lag as a result of traditional NF. He noted that mean comparisons of z-scores, with both positive and negative values being cancelled in the averaging process, had the potential of masking true differences. In an effort to account for this, he chose to transform the values by computing the absolute value of the z-score. He then used a score of $z \geq 1.0$ as inclusion criteria for analysis. This approach allowed for statistical analysis, (i.e.

averaging, ANOVAs, *t* tests) to be performed on the resulting z-scores transformed to absolute values.

Krigbaum and Wigton (2013) presented a methodological approach to account for positive and negative z-scores, by splitting the positive from negative z-scores, outside of a cut-off score of ± z = 1.0, to calculate what is termed Sites of Interest (SoI). The averaged SoI values were then plotted to display a learning curve for each participant, and statistical analysis (i.e. *t t*ests and rANOVAs) performed on the mean SoI z-score values. While this methodology fits well for a single-subject design, and in quantifying the progression of the z-scores towards the mean, its limitation lies in that (in the form presented) it is not well suited for comparisons of group mean QEEG data. For example, the split of positive and negative z-scores does not provide a single overall measure of change for the z-scores. However, there is room to build on this research to develop a methodology for comparing group data of QEEG z-scores.

Therefore, while few NF studies include EEG or QEEG z-score metrics as outcome measures, when they do, frequently they only analyze selective components (i.e. selected sites and/or frequencies). As a result, to date, no proposed methodology for quantifying overall normalization has been published. Averaging non-transformed z-scores is less than optimal due to the cancelling factor of the positive and negative values; and the binomial test of significance provides only limited categorical analysis of the data, without a measure of distance from the mean. The Krigbaum and Wigton (2013) study appeared the closest to providing a model for measuring overall normalization of the QEEG at this time. Still, building on this approach, by taking the absolute value of the

z-scores, to provide a single value as a measure of the distance from the mean, could prove advantageous.

In summary, common themes in the literature present suitable outcome measures for NF research to consist of computerized performance tests, rating scale instruments, and QEEG metrics. Examples such as the IVA, the BRIEF, and z-scores were discussed. These findings are relevant to this research in that the same or similar instruments were used for the present study.

Summary

In reviewing the 40 year history of NF, a discussion of the historical context of EEG, QEEG, and NF was presented. NF is grounded in learning theory and through the years various models, such as traditional NF, QNF, ZNF, have emerged. While 19ZNF is one of the newest NF models, it does not enjoy a demonstration of efficacy by evidence-based research, which exists for the traditional models. In fact, there are significant gaps in the literature in that no scientifically rigorous studies of 19ZNF have been found. This study aims to address this empirical gap by analyzing the question of efficacy of 19ZNF in a clinical setting, thus contributing to the field in terms of beginning to fill this empirical gap. Thus this study aims to contribute to the body of scholarly knowledge regarding 19ZNF.

Prior QNF and ZNF research is commonly found in clinical settings. These research studies typically employ pretest-posttest designs using relatively small sample sizes, while incorporating clinical assessment instruments and occasionally QEEG metrics, as outcome measures. Moreover, NF protocols are generally individually tailored, based on QEEG findings, with a goal to normalize the QEEG; and

heterogeneous collections of conditions included in studies is frequently found. These traditions were followed for this study, in both design and outcome measures, in evaluating 19ZNF. In utilizing QEEG z-scores as an outcome measure, prior research methods (SoI and taking absolute values of z-scores) were expanded on to establish a measure of distance from the mean, for statistical analysis of group data. The ZNF theory, grounded in the use of real-time z-scores with a goal of normalizing the QEEG, such that the z-scores move towards the mean ($z = 0$), underlies the 19ZNF approach; which was the focus of investigation in this pretest-posttest comparison research.

A detailed review and description of the methodology for this research is presented in the following chapter. To be included is an overview of the study, as well as further discussion of data collection and analysis methods. Additionally, the instrumentation, together with reliability and validity issues, will be discussed as it applies to the study. Limitations will also be reviewed.

Chapter 3: Methodology

Introduction

Over the years, new models of NF have been developed, and one of the most current is 19ZNF. To-date, case study and anecdotal clinical reports within the field indicate this new 19ZNF approach is an improvement over traditional NF models (J. L. Koberda et al., 2012a; Wigton, 2013). Still, the efficacy of this new model has not yet been established from empirical studies.

This research is different from other 19ZNF studies. It is a quantitative analysis of pre-post outcome measures, with group data from a clinical setting, and thus, it is a beginning in establishing empirical evidence regarding 19ZNF. The purpose of this retrospective one-group pretest-posttest research was to compare the difference between pre and post clinical assessments and QEEG z-scores data, before and after 19ZNF sessions, from archived data of a private neurofeedback practice in the Southwest region of the United States.

The remainder of this chapter reviews the problem statement and research questions, discusses the methodology and research design, and also describes the population and sample selection. Next, the instrumentation is presented together with a discussion of the associated validity and reliability. Then, data collection and data analysis is covered. Finally, a discussion of ethical considerations and the study limitations are presented.

Statement of the Problem

It is not known, by way of statistical evaluation of either clinical assessments or QEEG z-scores, if 19ZNF is an effective NF technique. This is an important problem

because 19ZNF is a new NF model currently in use by a growing number of practitioners, yet scientific research investigating its efficacy is lacking. Anecdotal reports are insufficient as a basis for determining treatment efficacy and uncontrolled case studies are scientifically weak (La Vaque et al., 2002). Therefore, scientifically sound evidence of efficacy for 19ZNF is needed.

Research Questions and Hypotheses

For this research, the independent variable was the 19ZNF and the dependent variables were clinical outcomes, as measured by the scaled scores from three clinical assessments (IVA, DSMD, BRIEF) and z-scores from QEEG data. Given the retrospective nature of this study, the approach for data collection was gathering archived de-identified data, from closed case files, of a NF private practice. The process consisted of collecting the necessary data elements (i.e. subject demographics, assessment scales scores, and z-scores) into spreadsheets, for further analysis by statistical software (SPSS). As will be discussed in detail in the research design section below, this study employed a one-group pretest-posttest design. This was the best design for the proposed research because the goal was to compare the means of the outcome measures at two different time points (before and after 19ZNF) (Kerlinger & Lee, 2000).

As will be detailed in the instrumentation section, and briefly reviewed below, the clinical assessments are generally designed to measure symptom severity of attention, behavior, and executive functioning; and the z-scores are a representational measure of electrocortical function. The clinical assessments are commercially available instruments, widely used in the mental health field for measuring symptom severity. The QEEG data

has been collected with a commercially available QEEG software package, which has been in general use in the neurofeedback field since 2002.

The instrument to measure attention was the IVA continuous performance test. This is a computerized test designed to assess both auditory and visual attention and impulse control symptoms associated with ADHD (Sanford & Turner, 2009). The associated research question and hypothesis was:

R1a. Does 19ZNF improve attention as measured by the IVA assessment?

H_a1a: The post scores will be higher than the pre scores for the IVA assessment.

$H_0$1a: The post scores will be lower than, or not significantly different from, the pre scores of the IVA assessment.

The instrument to measure behavior was the DSMD. This is a behavioral rating scale, completed by parents, designed to assess behavior problems and psychopathology in children and adolescents (Cooper, 2001). The associated research question and hypothesis was:

R1b. Does 19ZNF improve behavior as measured by the DSMD assessment?

H_a1b: The post scores will be lower than the pre scores for the DSMD assessment.

$H_0$1b: The post scores will be higher than, or not significantly different from, the pre scores of the DSMD assessment.

The instrument to measure executive function was the BRIEF. This is a rating scale, completed by parents, or self-rated in adults, design to measure observations of executive function skills in everyday environments (Gioia, Isquith, Guy, & Kenworthy,

2000; Roth, Isquith, & Gioia, 2005). The associated research question and hypothesis was:

R1c. Does 19ZNF improve executive function as measured by the BRIEF assessment?

H_a1c: The post scores will be lower than the pre scores for the BRIEF assessment.

$H_0$1c: The post scores will be higher than, or not significantly different from, the pre scores of the BRIEF assessment.

The instrument to measure the QEEG z-scores, which is a representational measure of electrocortical function, was the QEEG assessments collected using the Neuroguide software. This is software designed to provide statistical analysis of the quantified EEG metrics, such that z-scores are calculated to allow a comparison to a normative database (Thatcher, 2012). The associated research question and hypothesis was:

R2. Does 19ZNF improve electrocortical function as measured by QEEG z-scores such that the post z-scores are closer to the mean than pre z-scores?

H_a2: The post z-scores will be closer to the mean than the pre z-scores.

$H_0$2: The post z-scores will be farther from the mean, or not significantly different from, the pre z-scores.

Research Methodology

The field of clinical psychophysiology makes use of quantifiable variables and the associated research should include specific independent variables, as well as dependent variables, which relate to treatment response (i.e. clinical assessments) and the measured

physiological component (i.e. EEG metrics) (La Vaque et al., 2002). Yet, many NF studies do not use the EEG metrics as a measure of the cortical component of psychophysiologic function (Arns et al., 2009), but rather provide reports, which are more qualitative in nature to discuss NF related QEEG changes. Moreover, NF research needs to include quantitative methodologies, using QEEG data as an outcome measure, to learn more about the psychophysiological basis of NF (Gevensleben, 2009). Therefore, a quantitative methodology was selected, as opposed to qualitative, to address this need.

Currently, the available 19ZNF studies are in the form of qualitative research (Hallman, 2012; J. L. Koberda et al., 2012a). This literature entails presenting data from single case studies in the form of unstructured subjective reports of symptom improvement, as well as graphical images of before and after QEEG findings, where the improvement is represented by a change in color on the picture (without statistical analysis of data). However, for this dissertation, the goal was to explore statistical relationships between the variables under investigation; thus calling for a quantitative approach. The strength of quantitative methodologies, including quasi-experimental research, is that they provide sufficient information, regarding the relationship, and the level of significance, for the investigation variables, to enable the study of the effects of the independent variable upon the dependent variable (Carr, 1994). Therefore employing a quantitative method is intended to leverage this strength in the evaluation of 19ZNF.

Research Design

This quasi-experimental research used a retrospective, one-group pretest-posttest design. When the goal of research is to measure a modification to a behavior pattern, or internal process that is stable and likely unchangeable on its own, the one-group pretest-

posttest design is appropriate (Hunter & Schmidt, 2004; Kerlinger, 1986). This type of design answers the research questions by comparing the collected dependent variable pretest measures to the posttest values for each subject; thus comparing the members of the group to themselves, rather than to a control or comparison group (Kerlinger & Lee, 2000). Consequently, the group members become their own control; thus controlling for and thereby reducing the potential for extraneous variation due to individual-to-individual differences (Kerlinger & Lee, 2000). Moreover, the size of the treatment effect can be estimated by analyzing the difference between the pretest to the posttest measures (Reichardt, 2009). The rationale for this being a retrospective study is because the data available for analysis is from pre-existing archived records, which frequently provides a rich source of readily accessible data (Gearing et al., 2006). Therefore, the chosen design for this investigation is the best to evaluate the pre-post outcome measures from a clinical setting, as well as the identified research questions for this study.

As previously stated, the independent variable was the 19ZNF and the dependent variables were the data from the three clinical assessments and QEEG files; as such, the specific instruments used to collect the data were the IVA, DSMD, and BRIEF psychometric tests, as well as the QEEG software. A sample group was formed for each dependent variable outcome measure so as to form four groups for analysis. Therefore, using a one-group pretest-posttest design with these identified groups is fitting.

Population and Sample Selection

When individuals seek NF services they must choose among a variety of NF models. Yet the dearth of scientific literature regarding 19ZNF limits the information available for that decision process. The identified population for this research was made

up of those seeking NF services (both adults and children), or those who accessed NF services. These individuals may have had an array of symptoms, which adversely affect their daily functioning, most commonly in the areas of attention, behavior, and executive function; they may also have been previously diagnosed related mental health disorders.

From the total population (those seeking, or already have, accessed NF services), this particular study population was identified as all prior clients of the NF private practice which provided the retrospective data. Given the retrospective nature of this research, there was no active recruitment of subjects; thus sample selection was determined by inclusion criteria from available pre-existing cases. The study sample, then, were the cases which met the inclusion criteria of being a 19ZNF case, having both a pre and post QEEG assessment, as well as either an IVA, or a DSMD, or a BRIEF assessment, for both pre and post conditions. Moreover, given the sample consisted only of pre-existing de-identified data, as will be further detailed below (Data Collection section), there was no need for an informed consent process. For this research, the total aggregate sample size was 21 subjects, which was then divided into three additional outcome measures groups (IVA, DSMD, or BRIEF). The sample size for the IVA group was 10, the DSMD group was 14, the BRIEF group was 12, and all 21 subjects had QEEG data.

In a meta-analysis evaluating traditional NF, for ADHD, not using QEEG–targeted specificity in the NF protocols, Arns et al. (2009) reported an average (averaged for attention and hyperactivity symptoms) *Hedge's d* effect size of 0.85 (0.3 as small, 0.5 as medium, and 0.8 as large). In a more recent NF study where the treatment was more personalized and targeted with QNF, Arns et al. (2012) reported the average *Hedge's d*

effect size was found to be nearly double to 1.45 for the combined symptoms of attention and hyperactivity. Arns et al. (2012) suggested these findings indicate the personalization of treatment protocols, afforded by QNF, improves clinical outcomes. Given that 19ZNF also incorporates personalized QEEG-informed treatment protocols, it is reasonable to expect equivocal effect sizes with 19ZNF. Thus, in determining a needed sample size using the *G*Power3* software (Faul, Erdfelder, Lang & Buchner, 2007), for the reasons cited by Arns et al. (2012), it would be reasonable to use a predicted effect size in the range of 1.0 to 1.5. Using the more conservative effect size value of 1.0, with a one-tail analysis, alpha level of .05, and a power level of 0.80, for repeated measures *t* tests, the calculated needed minimum sample size is eight. Therefore, groups with a sample size of 10 or more are sufficient for the data analysis to be performed in this study.

Instrumentation

The type of archived data used was from the following instruments: One computerized performance test (IVA), two rating scales (DSMD and BRIEF), and QEEG z-scores (Neuroguide software). All clinical assessments are commercially available validated instruments, having a history of common use in the mental health industry. The QEEG software is also commercially available, and since 2002 has been used internationally by NF clinicians, in university research settings, and military/veteran institutions (Besenyei, et al. 2012; Thatcher, North, & Biver, 2005). All instruments were completed as part of the pre and post assessment routines during the previously completed NF treatment process. All treatments were provided by the researcher who is a state Licensed Professional Counselor, a board certified Neurofeedback Therapist, and a

certified QEEG Diplomate. Descriptions of each of the instruments are provided next, with a discussion of validity and reliability in separate following sections.

IVA. As reported by Sanford and Turner (2009), the IVA is a 13-minute computerized test, with 500 responding or inhibiting trials, normed for ages six to adult, designed to assess both auditory and visual attention and impulse control; with the aim to aid in the quantification of symptoms and diagnosis of ADHD. Yet, the manual provides usage indications to include assessing attention and self-control problems related to other disorders, such as depression, anxiety, head injuries, dementia, and other medical problems. The test taker is given standardized instructions, from a computer digitized voice file, that they will see or hear the numbers *1* or *2*, and to click the mouse when they see or hear the number *1*, and to refrain from clicking if they see or hear the number *2*. There are two global full-scale composite scores of Full Scale Response Control Quotient, and Full Scale Attention Quotient. Each full scale is broken into auditory and visual scales. Auditory and visual primary scales for Response Control include Prudence (impulse control), Consistency (response reliability), and Stamina (sustained attention over time). Auditory and visual subscales for Attention include Vigilance (inattention), Focus (mental processing variability), and Speed (reaction time). The test results are reported in the form of quotient scores such that a score of ≤ 85 is indicative of clinical significance. As a performance test, the IVA is completed directly by the subject.

DSMD. The DSMD is a behavior rating scale designed to assess behavior problems and psychopathology in children and adolescents; the child form (ages 5 to 12) and adolescent forms (ages 13 to 18) have 110 items which describe problem behaviors, with a 65% overlap between the two forms (Cooper, 2001). The rater can be either a

parent or teacher, with separate norms for each; in this research, only parent ratings are used. Both versions have (1) a composite Externalizing scale made up of Conduct and Attention (child)/Delinquency (adolescents), (2) a composite Internalizing scale made up of Anxiety and Depression, (3) a composite Critical Pathology scale made up of Autism and Acute Problems, and (4) a global Total scale (Peterson, 2001). The instrument scores are expressed in T scores, with scores ≥ 60 indicating clinical significance, and can be completed in about 15 minutes.

BRIEF / BRIEF-A. The BRIEF is a rating scale, with 86 items, designed to sample observations of children's (ages 5 to 18) executive function skills in everyday natural settings, with forms suitable for completion by parents and teachers (Donders, 2002). For this study only the parent form was available. This instrument is intended to assess behavioral, emotional, and metacognitive skills, which broadly encompass executive skills, rather than measure behavior problems or psychopathology (Donders, 2002). The BRIEF-A is the adult version (ages 18 to 90), self-report form, with 75 items, which is designed to assess the views of one's own executive function skills (self-regulation) in their everyday environment (Gioia et al., 2000). Both instruments have an overall summary scale of Global Executive Composite (GEC), which is comprised of two primary sub-scales of Behavioral Regulation Index (BRI) and Metacognition Index (MI). The BRI is made up of clinical scales of Inhibit, Shift, and Emotional Control for both the adult and child versions, with the BRIEF-A adding a scale of Self-Monitor to the behavior regulatory clinical scales category. The MI, for both the BRIEF and BRIEF-A, is made up of five clinical scales of Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor. Both assessments take approximately 15 minutes

to complete; and scores are expressed in terms of T scores, with scores ≥ 65 indicating clinical significance (Gioia et al., 2000; Roth et al., 2005).

Neuroguide and QEEG acquisition. The QEEG data was acquired and processed with the Neuroguide software. This software is designed to collect conventional EEG data, and then allow for simultaneous visual inspection of the raw EEG waveforms together with statistical analysis of the quantified EEG metrics (Thatcher, 2012). Software modules allow the EEG data to be compared to a lifespan normative database. The database has been normed, for both eyes open and eyes closed conditions, with 625 individuals from ages of 2 months to 82 years of age, with the included subjects being screened for normalcy (normal intelligence, lack of pathology or mental health disorders) through history, interviews, neuropsychological testing and other evaluations (Thatcher, Walker, Biver, North, & Curtin, 2003). The amplifier used for the EEG acquisition was the Brainmaster-Discovery 24E (Brainmaster Technologies, Inc, Bedford, OH), with an A/D conversion of 24 bits resolution, a sampling rate of 256 Hz, and input impedance of 1000GOhms. Impedance is the obstruction of flow of electrical current when measuring non-direct current signals (Farley & Connolly, 2005).

EEG data was acquired and processed as has been described by Krigbaum and Wigton (2013), using accepted standards of QEEG acquisition methods, thus ensuring quality recordings. An electrode cap (Electro-Cap Inc; Eaton, OH) was used to place the 19 electrodes according to the International 10-20 System referenced to linked ears, with Electro-Cap brand electro-conductive gel. Electrode impedances were adjusted to be below 10k ohm for all electrodes and balanced. The digital format of the EEG recording was with a low-pass filter of 50 Hz and a high-pass filter of 0.5 Hz. The pre and post

EEG recordings were acquired with eyes open in a waking-relaxed state, sitting in an upright relaxed position. The instructions given were to remain still, inhibit muscle activity from forehead, neck, and jaws, as well as eye movements and blinks. Screening of EEG was conducted carefully to exclude technical and biological artifacts. The *EEG Selection* method (Thatcher, 2012) was used to eliminate artifacts prior to submitting the EEG to a fast Fourier transformation (FFT) procedure. The remaining edited EEG consisted of an average of 1 minute of data (30 2s epochs), thus ensuring a representative sample of data verified by the split-half and test-retest values being ≥ .90. The digitally filtered frequency bands, for surface potential metrics of absolute power, relative power, and coherence, were as follows: Delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), alpha1 (8-10 Hz), alpha2 (10-12 Hz), beta (12-25 Hz), beta1 (12-15 Hz), beta2 (15-18 Hz), beta3 (18-25 Hz), and high beta (25-30 Hz).

Validity

The concept of test validity refers to degree to which it accurately measures that which it proposes to measure, and also how well it measures the target in question (Anastasi & Urbina 1997). Thus, the emphasis is on the accuracy of the measure with regard to the aspect of *what* is to be measured. Aspects of validity of the outcome measures for this study will next be addressed.

IVA. A concurrent and diagnostic validity study was conducted by Nova Southeastern University and BrainTrain, Incorporated. The findings suggested the overall accuracy, when compared to diagnoses of ADHD provided by physician/psychologists, to be statistically significant ($p < .0001$). Moreover, the sensitivity (true positives) was

reported to be 92%, specificity (true negatives) as 90%, and positive and negative predictive power as 89% and 93% respectively (Sanford & Turner, 2009).

DSMD. Peterson (2001) reported content validity for the DSMD to be good, with a strong congruence with the Diagnostic Statistical Manual-IV criteria regarding the behaviors examined. Moreover, the DSMD scales have a diagnostic potential to identify normal versus hospitalized children/adolescents with an accuracy range of 70% to 90% (Cooper, 2001). In a study to examine concurrent validity with the BASC and the CBCL, Smith and Reddy (2002) found the DSMD to demonstrate strong concurrent validity with scales, which were conceptually similar. For example, between the DSMD and the CBCL, the correlations were .81 for the Externalizing scale, .83 for the Internalizing scale, and.86 for Total scale (Smith & Reddy, 2002). This is important, given that many NF studies have previously used the BASC and CBCL; thus demonstrating the DSMD to be similar to other rating scales, as a behavior measure, used with prior NF studies.

BRIEF / BRIEF-A. Content validity for the BRIEF was determined by seeking agreement between multiple pediatric neuropsychologists and the test authors for fit of each test item. The items retained in the clinical scales have item-total correlations that range from .43 to .73 (Gioia et al., 2000). Content validity for the BRIEF-A was conducted in a similar manner by seeking agreement among multiple neuropsychologists experienced with executive function issues in clinical practice. Of the retained items for the clinical scales the agreement ranged from .38 to .98 (Roth et al., 2005).

Neuroguide QEEG database. As described by Thatcher et al. (2003), the validation procedure for the Neuroguide QEEG database included a leave one out Gaussian (normal distribution) cross-validation process, whereby the data for each

subject in the database was removed and then compared to that same database. This is important because the database, which is being compared to needs to fit the normal curve to ensure unbiased error estimates. In a normal distribution cross-validation with a perfect fit, it would be expected that 2.3% of the comparison sample would fall outside of +2 standard deviations (SD) and again at -2 SD, and that 0.13% at +3 SD and again at –3 SD. Therefore, percentages which approximate these values can be deemed as validating the normal distribution. The cross-validation process for the Neuroguide database revealed an overall percentage (of all metrics) at +2 SD to be 2.58%, and for –2 SD to be 1.98%; then for +3 SD to be 0.18%, and for -3 SD to be 0.14%. Moreover, the kurtosis and skewness of the database, if fitting the normal distribution, would be within a few percentage points of zero. Thatcher, Walker et al. reported the validation process found the Neuroguide database to meet the criteria for skewness with an overall percentage of 0.17%, and for kurtosis with an overall percentage of 2.91%.

Reliability

Reliability is an important aspect in determining if one can trust that a particular assessment will give a comparatively similar measure if it is given at another time. As such, reliability reflects score consistency and predicts how much variation one can expect from one administration of the test to the next (Anastasi & Urbina, 1997). Thus, reliability allows an estimate of the error of measurement for the instrument. Aspects of reliability of the outcome measures for this study will next be addressed.

IVA. A test-retest study was conducted by Nova Southeastern University and BrainTrain, Incorporated, with a testing interval of 1 to 4 weeks. The results showed statistically significant ($p < .01$) reliability coefficients ranging from .37 to .75 (Attention

scales: .66 to .75; Response Control scales: .37 to .41). The findings of this study are further reported to support the IVA as being a stable measure of performance while also being robust against learning or practice effects, such that changes in test scores over time can reliably be attributed to environmental or treatment effects (Sanford & Turner, 2009).

DSMD. The test-retest reliability was measured for the DSMD and is reported to range from .80 to .90 for the scales, with an interval of a 24-hour period (Peterson, 2001).

BRIEF/BRIEF-A. The test-retest reliability was measured for both the clinical and normative samples of the BRIEF, which was reported to be .81 for the normative sample and .79 for the clinical sample, with an average interval of two to three weeks; whereas the reliability for the BRI, MI, and GEC was ≥ .80 for both the clinical and normative samples (Gioia et al., 2000). For the BRIEF-A the test-retest reliability, over an average interval of four weeks, for the clinical scales was reported to range from .82 to .93; with the reliability for the BRI, MI and GEC being > .92 (Roth et al., 2005).

Neuroguide QEEG software. Recently, a study was conducted to evaluate the reliability of the FFT metrics of the Neuroguide software. Cannon et al. (2012) found the Neuroguide test-retest reliability, at a 30-day interval, to be ≥ .77 for absolute and relative power, and coherence. A further measure of reliability, with the individual EEG records in Neuroguide, are a test-retest and split-half measure which is calculated when the artifacts are removed, which when ≥ .90 provide a representative sample of the overall EEG record (Thatcher, 2012). The edited EEG records for this study were edited such that both the split-half and test-retest measures were on average ≥ .90.

Data Collection Procedures

The sample consisted of a convenience sample from reviewed closed cases, of clients from a private neurofeedback practice, who were administered the clinical assessments and QEEGs before and after 19ZNF treatment. Regarding the retrospective data used in this study, those clients were informed that after their treatment was completed and their case closed, non-identifying data could be used for quality assurance and/or future research purposes; they were all given the opportunity to opt-out. To be considered an available 19ZNF case, the clinical symptoms presented during the intake assessment corresponded with the z-score deviations of the QEEG findings, such that a treatment goal of overall QEEG normalization was clinically appropriate. While the 19ZNF protocol developed for each case was individually tailored to the clinical and QEEG findings, and possibly modified at each session to correspond with the baseline QEEG data of that day, the same treatment goal always applied; that of overall QEEG normalization. Therefore, the underlying 19ZNF protocol of overall QEEG normalization was consistent for all cases. The hardware platform was the Brainmaster Discovery 24E amplifier, and the software platform was either the Brainmaster Discovery or Neuroguide NF-1 19ZNF software. The 19ZNF sessions used the Brainmaster Flashgame visual NF displays (i.e. simple non-movie animations); and the reward percentages were approximately 30% to 50% (i.e. 20 to 30 rewards-per-minute).

As depicted in Figure 1.1, from the available 19ZNF cases, an initial group was formed for which pre-post QEEG assessments existed, and for which *either* the IVA, DSMD, *or* BRIEF pre-post assessment data were also available ($n = 21$). From this collection, three additional groups were formed. One group was created for the IVA data

(n = 10), a second group for the DSMD data (n = 14), and a third group for the BRIEF data (n = 12).

The data collected for this study were from pre-existing documents/files and recorded by the investigator in a manner such that the subjects cannot be identified. Therefore, in accordance with 45 CFR 46.101(b) and 46.101(b)(4), this research was exempt from the requirements of the Protection of Human Subjects 45 CFR part 46 (2009) regulation. Consequently, the university Institutional Review Board (IRB) determined this study to be exempt from IRB review, under exemption category 7.4 (see Appendix B). As such, IRB-approved informed consent for use of the de-identified data for this research was not necessary. All data for this study were previously obtained during the course of subjects' NF treatment. While the data came from records that already exist *prior* to the start of the study, there was a form of data collection by pulling *de-identified* information from a review of the archived records of the private practice. Upon IRB approval, the information was gathered and de-identified in a format such that it was impossible to identify the subjects. For example, copies/scans were made of the assessment scoring sheets, but names and/or birthdates (or any other identifying information) were redacted, and only a sequential case number was assigned and written on documents associated with that case. The pre and post scaled score data, from the copied assessment forms, were entered into a spreadsheet to facilitate data analysis. For the QEEG data, with the Neuroguide software, the report generation feature was used to save the z-scores into tab delimited text files, which were then saved as Microsoft Excel worksheet files, thus preparing the data for further analysis.

The redacted paper forms of the collected data set are stored in a secured manner (i.e. under lock/key) and separate from the clinical source files (which also provides for physical backup of data). For data that was entered into spreadsheets and the statistical software package, those digital files are stored on an external flash drive separate from any installed computer hard-drive; and, when not in use, will be kept with the paper data files in the same secured manner. The data are stored in a secured manner, with hard-copy (i.e. paper) data as a form of permanent backup, separate from the archived source files, and will be maintained for the required 3 years after the completion of the study. At the end of the 3 years paper files will be shredded and electronic media digitally erased. A further subject identity protection were that findings reported were only descriptive group data, and no individual case was described or discussed; thus preventing any possible inadvertent identification of persons.

Data Analysis Procedures

In general, the research questions asked if 19ZNF improved attention, behavior, executive function, and electrocortical function, as measured by the clinical assessments of the IVA, DSMD, BRIEF, and QEEG z-scores. All alternative hypotheses were similar in that the IVA hypothesis predicted the post scores would be higher than the pre scores, the DSMD and BRIEF hypotheses predicted the post scores would be lower than the pre scores, with the z-score hypothesis predicted post z-scores to be closer to the mean than pre z-scores. The null hypotheses all predicted no significant difference, or a difference opposite the direction of improvement. The level of significance for this study was alpha = .05.

As previously described in the above section, the scaled scores from the clinical assessments and QEEG z-scores were collected from archived clinical files and organized by data entry into spreadsheets for analysis in SPSS v21 software. Columns for relevant data categories (demographics, pre scores, post scores, difference scores), and identified relevant scales (composite and global), were created to facilitate entry of data into the fields of the spreadsheets. The data analysis started with performing descriptive statistics on each of the sample groups; the means for the pre, post, and difference scores were also calculated. The specific scales that were analyzed from each clinical assessment are described next, and are followed by details related to the z-score data analysis.

IVA. The IVA assessment has two primary categories of scales, Response Control and Attention. The research question associated with this variable focused on improvement of attention. Thus, in order to maintain alignment with the research question, only the overall scales specific to attention were analyzed. Therefore scores from the Full Scale Attention Quotient, Auditory Attention Quotient, and the Visual Attention Quotient, were collected and analyzed. Additionally, these scales have higher reliability measures than Response Control scales.

DSMD. The DSMD has two composite scales more specific to generalized behavior, that being the Externalizing Composite and Internalizing Composite scales, as well as a Total scale. These three scales correlate strongly (.81, .83, .86, respectively) with similarly named scales from the CBCL, which is an instrument commonly used as an outcome measure of behavior in NF studies. Thus this strategy maintained alignment with the associated research question (improvement of behavior) for this variable.

BRIEF / BRIEF-A. While all scales on the BRIEF instruments capture elements of executive function, in order to maintain alignment with the analysis of the other instruments (i.e. analyzing generalized composite/global scales), only the composite scales of Behavior Regulation Index and Metacognition Index, as well as the Global Executive Composite scale were analyzed. Moreover, both the BRIEF and BRIEF-A contain these composite/global scales, thus maintaining consistency in the child and adult assessment measures. Therefore these scales maintained alignment with the associate research question (improvement of executive function) for this variable.

QEEG z-scores. The QEEG z-scores are a representational measure of electrocortical function, such that z-scores which are closer to the mean represent improved functioning; thus maintaining alignment with the research question associated with this variable. The z-score data were calculated for the QEEG metrics of absolute power, relative power, and coherence; the same procedure was followed for each metric. First the z-scores were converted into a spreadsheet format. Next, the values were transformed to the absolute value. Then, the pre z-scores which were ≥ 1.0 were highlighted as being the targeted (by site and frequency) z-scores. Those targeted z-scores were averaged to create a single value, representing an overall distance from the mean for that metric, for that case. Next, the same targeted z-scores for the corresponding post values (i.e. same site and frequency) were identified and averaged. This allowed the pre and post averaged targeted z-score values to be compared, as a measure of change, such that a lower post value (compared to the pre value) would be closer to the mean.

Statistical analysis. Given that each of the variables forms a separate analysis group, the proposed data analysis aligned with the one-group pretest-posttest design. The

paired (within-subjects/repeated measures) *t* test was appropriate (assuming the difference scores to be normally distributed) for this quantitative research, with continuous variables, because it was based on the difference scores (between pre and post) for measures taken for each person in a single sample, while allowing for sufficient statistical power with smaller sample sizes (Gravetter & Wallnau, 2010). Effect size was computed for discussion of practical results, and compared to that previously reported from prior studies in the literature.

The statistical analysis was conducted with the SPSS v21 statistical package. For all hypotheses, the plan was for paired *t* tests on the pre/post difference scores, for the means of the selected scales and z-scores, for each outcome measure. The data from the spreadsheet columns, for the pre and post values (for the scales of each outcome measure) was transferred into SPSS. Next, the SPSS command sequence selected was *Analysis>Compare Means>Paired Samples T Test*. The pre values were identified as *Variable1* and the post values identified as *Variable2*, and the *Confidence Interval Percentage* will be set at 95%. Finally, Hedge's *d* effect sizes were calculated with the Metawin 2.1 software.

Ethical Considerations

There were no ethical problems for this dissertation primarily because it was determined to be exempt from the requirements of the Protection of Human Subjects 45 CFR part 46 (2009) regulation. Consequently, IRB-approved informed consent for research was not necessary. As described above, all data was pre-existing prior to the start of the study and recorded such that there was no potential for revealing the identity of any person included. The researcher owned the private practice data

therefore no data use or site authorization was needed. The data was stored in a secured manner, with hard-copy (i.e. paper) data as a form of permanent backup, separate from the archived source files, and will be maintained this way for the required 3 years after the completion of the study. At the end of the 3 years, paper files will be shredded and electronic media digitally erased.

Limitations

There were three primary limitations to this study; that of research design elements, sample size, and the question of efficacy. Moreover, it is important to examine potential sources for bias in any research. Thus, this aspect will also be discussed.

Most criticisms of pretest-posttest designs, which imply they are inadequate due to threats to internal validity, can be traced back to Campbell and Stanley (1963). However, as pointed out by Hunter and Schmidt (2004), the identified limiting elements (history, maturation, instrumentation, testing, and regression) were only presented by Campbell and Stanley as *potential* threats, which may or may not adversely impact a study. Moreover, in studies of psychological factors, where the intent is intervention evaluation, the behavior targeted by the treatment (i.e. the DV) is typically quite difficult to change without some intervention; thus the Campbell and Stanley potential validity threats were ruled out (Hunter & Schmidt, 2004).

Nonetheless, a general limitation of designs, which incorporate a pretest-posttest framework is primarily related to the passage of time between administering the pre and post assessments (Kerlinger & Lee, 2000). Factors such as history (concurrent events external to the study scope) and maturation (internal growth factors occurring regardless of interventions) cannot be controlled for. Therefore, it is not possible to know whether or

not they have impacted the DV measures (Hunter & Schmidt, 2004). Yet, when the time between testing points is short, the impact of extraneous variation is lessened (Kerlinger & Lee, 2000; Reichardt, 2009). In this study, the time between the pre and post assessment was relatively short, measured in terms of weeks. Therefore, the impact of time-related confounds were considered to be minimal. Also, identified as a potential validity threat is the phenomenon of a regression to the mean, where high or low scores are, by chance, found to be closer to the mean when retested. However, there is an inverse relationship between the degree of statistical regression and an instrument's reliability (Kirk, 2009); such that instruments with higher reliability have less variability in the measurement error. Given the reliability of the instruments in this study are relatively high, the estimate of the error of measurement is comparatively low. Thus, potential validity threats related to regression effects were minimal.

Larger sample sizes are preferred in order to allow for stronger statistical analysis and more generalizability (Gravetter & Wallnau, 2010). Given this study used pre-existing archived data, the number of samples was restricted to what was found in the files; thus there was no option to increase sample size. Though, as discussed, the sample sizes for each group had sufficient power to allow for adequate statistical analysis.

In order to fully address the question of efficacy, additional studies involving both follow-up data and control group comparison data are necessary. This is especially true in answering whether 19ZNF is superior to other QEEG-based approaches. Therefore, limitations of this study, which also must be recognized, are a lack of comparison to a traditional NF group and a lack of a randomized control group. Nevertheless, given the

data for this research comes from a real-world clinical setting, the findings of this study can still contribute to advancing the scientific knowledge of 19ZNF.

Finally, in examining potential sources of bias, in a retrospective study where the data comes from the archived treatment cases of the researcher, a question could be asked regarding how the researcher can account for the potential. Given the data was pre-existing in closed cases, and could only be reported, the numerical information could not be changed nor manipulated. In other words, the data existed in a set form, and the statistical analysis conveys the message. Moreover, by de-identifying the data such that every subject was reduced to merely a case number, the researcher even became blind to the identity to the subjects within the study. Further, there was no qualitative data in this study for the researcher to interpret. For these reasons, it is believed the potential for bias was minimized in this study.

Summary

In summary, the methodology for this retrospective pretest-posttest comparative research was presented. As was reviewed, the independent variable was the 19ZNF and the dependent variables were the data from the three clinical assessments and QEEG z-scores; the instruments were the IVA, DMSD, BRIEF psychometric tests, and QEEG software. The population was described as those who seek NF services, with the study sample being the pre-existing data available meeting the inclusion criteria, such that four groups were formed; one group for each outcome measure. A discussion was presented regarding data aspects germane to a retrospective study, such as how the data was pre-existing and only de-identified information was collected. Consequently this study qualified as exempt from the requirements of the Protection of Human Subjects 45 CFR

part 46 (2009) regulation, and IRB-approved informed consent was not necessary. Finally, limitations were reviewed, and what are typically identified as potential weaknesses in pretest-posttest designs (Campbell & Stanley, 1963), were minimally impactful because intervention targeted behaviors frequently do not change without effective intervention (Hunter & Schmidt, 2004), there was a short pre-post time interval (Reichardt, 2009), and the instruments employed in this study have relatively high reliability measures (Kirk, 2009).

In this study, all research questions were similar and the paired *t* test was an appropriate statistic to compare the means of the different data groups. Moreover, effect size was computed and compared to prior studies. In the following chapter, the process of the data analysis, as well as results, will be discussed.

Chapter 4: Data Analysis and Results

Introduction

Addressing efficacy of 19ZNF is important because it was not known, by way of statistical evaluation of either clinical assessments or QEEG z-scores, if 19ZNF is an effective NF technique. Therefore, the purpose of this quantitative research was to evaluate 19ZNF, in a clinical setting, using a retrospective one-group pretest-posttest research design. Generally, the research questions asked if 19ZNF improves attention, behavior, executive function, and electrocortical function, as measured by the outcome measures of the IVA, DSMD, BRIEF, and QEEG z-scores. All alternative hypotheses were similar in that the IVA hypothesis predicted the post scores would be higher than the pre scores, the DSMD and BRIEF hypotheses predicted the post scores would be lower than the pre scores, and the QEEG hypothesis predicted post z-scores would be closer to the mean than pre z-scores. The null hypotheses all predicted no significant difference, or a difference opposite the direction of improvement.

This chapter first presents the descriptive data of each of the groups for the IVA, DSMD, BRIEF, and QEEG z-score data. Then, the steps taken for data analysis are described. Finally, results of the data analysis are presented.

Descriptive Data

The QEEG group represented the inclusion of all subjects for the study, from which the other groups were formed; therefore, it is described first. Then, the groups for the IVA, DSMD, and BRIEF are described. Table 4.1 summarizes the descriptive information as discussed in the following sections. It is important to note, while the clinical assessment groups were diverse diagnostically, when viewed by clinical

complaints, in terms of the neuropsychological constructs of attention, behavior, or executive function, the subjects collectively formed well-defined groups for which the assessment instruments are designed to measure.

QEEG group. The total sample size for this group was 21; there was no reported experience of 19ZNF prior to coming to this practice. The subjects ranged in age from 7 to 63 years, with a mean age of 21.19 years (*SD* = 18.12); including 15 children and six adults, 10 males and 11 females. Seventeen of the subjects were White, two were Asian, and two were Latino; while five were categorized as low socioeconomic status (SES), 14 as medium SES, and two as high SES. The make-up of the diagnosis[2] and/or presenting conditions included mostly a combination of ADHD-Inattentive presentation (ADHD-I) and ADHD-Combined presentation (ADHD-C) (ADHD-I = 4, ADHD-C = 7); yet, there were three subjects with ADHD-C comorbid with another disorder (ADHD-C/Unspecified Anxiety Disorder, ADHD-C/Autism Spectrum Disorder, ADHD-C/Unspecified Learning Disorder). Finally, the other diagnoses included one comorbid Unspecified Anxiety/Unspecified Depressive Disorder, one Autism Spectrum Disorder, one Unspecified Bipolar Disorder, one Reactive Attachment Disorder, one comorbid Obsessive-Compulsive Disorder/issues with executive function, and two subjects with presenting issues of difficulty with executive functioning. A total of 16 subjects had no medication usage, two subjects were on medication, two subjects started on medication

[2]Given the retrospective nature of the data, all initial diagnoses were made in accordance with the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM-lV-TR*; American Psychiatric Association, 2000). However, all diagnoses criteria were confirmed with, and are reported in accordance with the *DSM-5* (American Psychiatric Association, 2013) taxonomy.

but had ceased medication by the post assessment, and one subject had a reduction of medication by one-third at the time of post assessment. The number of sessions from pre assessment to post assessment ranged from three to 20, with a mean of 10.90 (*SD* = 3.88). The targeted session frequency was once per week. The number of weeks for treatment (pre to post assessment) ranged from two to 22, with a mean of 11.76 (*SD* = 5.19). Finally, the number of weeks from pre assessment to post assessment ranged from two to 43, with a mean of 15.10 (*SD* = 10.03). The descriptive data for this group is summarized in Table 4.1.

IVA group. The total sample size for this group was 10. The subjects ranged in age from 7 to 63 years, with a mean age of 26.80 years (*SD* = 19.84); including five children and five adults, five males and five females. Nine of the subjects were White, and one was Latino; while three were categorized as low SES, five as medium SES, and two as high SES. The make-up of the diagnoses and/or presenting conditions included mostly a combination of ADHD, with three ADHD-I and four ADHD-C; yet, there were two subjects with ADHD-C comorbid with another disorder (ADHD-C/ Unspecified Anxiety Disorder, ADHD-C/ Unspecified Learning Disorder). Finally, the other diagnoses included one subject with presenting issues of difficulty with executive functioning. A total of eight subjects had no medication usage, one subject was on medication, and one subject started on medication but had ceased medication by the post assessment. The number of sessions from pre assessment to post assessment ranged from three to 15, with a mean of 9.70 (*SD* = 3.92). The targeted session frequency was once per week. The number of weeks for treatment (pre to post assessment) ranged from two to 15, with a mean of 9.40 (*SD* = 4.40). Finally, the number of weeks from pre

assessment to post assessment ranged from two to 43, with a mean of 13.20 (*SD* = 11.11). The descriptive data for this group is summarized in Table 4.1.

DSMD group. The total sample size for this group was 14. The subjects ranged in age from 7 to 17 years, with a mean age of 10.86 years (*SD* = 2.91); including 14 children and no adults, seven males and seven females. Ten of the subjects were White, two were Asian, and two were Latino; while three were categorized as low SES, nine as medium SES, and two as high SES. The make-up of the diagnoses and/or presenting conditions included a combination of ADHD, with two ADHD-I and five ADHD-C; yet, there were two subjects with ADHD-C comorbid with another disorder (ADHD-C/Autism Spectrum Disorder, ADHD-C/Unspecified Learning Disorder). Finally, the other diagnoses included one comorbid Unspecified Anxiety/Unspecified Depressive Disorder, one Autism Spectrum Disorder, one Unspecified Bipolar Disorder, one Reactive Attachment Disorder, and one subject with presenting issues of difficulty with executive functioning. A total of 11 subjects had no medication usage, one subject was on medication, one subject started on medication but had ceased medication by the post assessment, and one subject had a reduction of medication by one-third at the time of post assessment. The number of sessions from pre assessment to post assessment ranged from three to 20, with a mean of 11.43 (*SD* = 4.13). The targeted session frequency was once per week. The number of weeks for treatment (pre to post assessment) ranged from three to 22, with a mean of 12.57 (*SD* = 5.60). Finally, the number of weeks from pre assessment to post assessment ranged from six to 37, with a mean of 15.36 (*SD* = 8.63). The descriptive data for this group is summarized in Table 4.1.

BRIEF group. The total sample size for this group was 12. The subjects ranged in age from 7 to 63 years, with a mean age of 20.25 years (*SD* = 19.97); including 10 children and two adults, six males and six females. Eleven of the subjects were White, and one was Latino; while two were categorized as low SES, nine as medium SES, and one as high SES. The make-up of the diagnoses and/or presenting conditions included a combination of ADHD, with two ADHD-I and two ADHD-C; yet, there were two subjects with ADHD-C comorbid with another disorder (ADHD-C/Autism Spectrum Disorder, ADHD-C/ Unspecified Learning Disorder). Finally, the other diagnoses included one comorbid Unspecified Anxiety/Unspecified Depressive Disorder, one Autism Spectrum Disorder, one Reactive Attachment Disorder, one comorbid Obsessive-Compulsive Disorder and issues with executive function, and two subjects with presenting issues of difficulty with executive functioning. A total of 10 subjects had no medication usage, one subject was on medication, and one subject started on medication but had ceased medication by the post assessment. The number of sessions from pre assessment to post assessment ranged from three to 20, with a mean of 11.83 (*SD* = 2.69). The targeted session frequency was once per week. The number of weeks for treatment (pre to post assessment) ranged from approximately three to 22, with a mean of 13.50 (*SD* = 3.97). Finally, the number of weeks from pre assessment to post assessment ranged from six to 37, with a mean of 16.17 (*SD* = 8.44). The descriptive data for this group is summarized in Table 4.1.

Table 4.1

Descriptive Data for All Groups

Category	QEEG Group (*n* = 21)	IVA Group (*n* = 10)	DSMD Group (*n* = 14)	BRIEF Group (*n* =12)
Age *M (SD)*	21.19 (18.12)	26.80 (19.84)	10.86 (2.91)	20.25 (19.97)
Children	15	5	14	10
Adults	6	5	0	2
Gender				
Male	10	5	7	6
Female	11	5	7	6
Ethnicity				
White	17	9	10	11
Asian	2	0	2	1
Latino	2	1	2	0
Socioeconomic Status				
Low	5	3	3	2
Medium	14	5	9	9
High	2	2	2	1
Diagnosis or Condition				
ADHD-Inattentive	4	3	2	2
ADHD-Combined	7	4	5	2
ADHD-C/Anxiety	1	1	0	0
ADHD-C/ASD	1	0	1	1
ADHD-C/LD	1	1	1	1
Anxiety/Depression	1	0	1	1
ASD	1	0	1	1
Bipolar	1	0	1	0
Executive Function	2	1	1	2
OCD/Exec Function	1	0	0	1
RAD	1	0	1	1
Medication				
No	16	8	11	10
Yes	2	1	1	1
Yes to off	2	1	1	1
Yes to reduced	1	0	1	0
# Sessions pre-to-post *M (SD)*	10.90 (3.88)	9.70 (3.92)	11.43 (4.13)	11.83 (2.69)
# Weeks for treatment *M (SD)*	11.76 (5.19)	9.40 (4.40)	12.57 (5.60)	13.50 (3.97)
# Weeks pre to post assessment *M (SD)*	15.10 (10.03)	13.20 (11.11)	15.36 (8.63)	16.17 (8.44)

Note. ADHD: Attention Deficit Hyperactivity Disorder; ASD: Autism Spectrum Disorder; LD: Learning Disorder; RAD: Reactive Attachment Disorder; OCD: Obsessive-Compulsive Disorder.

Data set limitations. Given the data for this study was derived from a retrospective collection of information from existing files, there are some inherent limitations. Yet, while these limitations are unavoidable, taking data from real-world records gives an opportunity to evaluate an intervention using realistic information typically found in a clinical setting. Such was the case with this study. One example is with regard to medication usage of the subjects. In a true experimental setting, having no medication usage in all subjects would be ideal; however, NF clinicians routinely have clients who seek out NF while still taking medications. The frequency of cases involving medication use in this study, with an overall five out of 21 for the QEEG group, two out of 10 in the IVA group, three out of 14 for the DSMD group, and two out of 12 for the BRIEF group, was a fairly accurate representation of the overall population that has been seen in this practice for close to 15 years. Therefore, while an argument could be made that a data set with no medication usage may provide for more credible results; in reality, the data in this study made the results more generalizable to the population of those who actually seek NF services.

The other example in this study, impacted by a fixed data set, was regarding the number of weeks from pre assessment to post assessment. The apparent great variability in number of weeks was accounted for by two outlier cases on the high end (i.e. 37 and 43 weeks), and one outlier case on the low end (i.e. 2 weeks). If these outliers were excluded, the range of number of weeks would have been from six to 26 for the QEEG and DSMD groups, from seven to 15 for the IVA group, and from seven to 26 for the BRIEF group. This then, better explains why the group means averages of the number of weeks from pre assessment to post assessment (as seen in table 4.1) are 15 weeks for the

QEEG and DSMD groups, 13 weeks for the IVA group, and 16 weeks for the BRIEF group. Also of note, it is important to realize that the number of weeks from pre to post assessment does *not* represent the number of sessions; it is only the time elapsed between assessment points. To clarify this aspect, the number of sessions pre-to-post treatment, as well as the number of weeks for the treatment, are reported to better illustrate the timeframes and sessions performed during the 19ZNF.

Data Analysis Procedures

The data analysis procedures were conducted with no deviation from what was described in the previous Methodology chapter. Given this study consisted solely of data collection and analysis, the greatest source of error was data collection and data entry errors, which would negatively impact this research with inaccurate results. This was mitigated by being careful in the data handling, as well as double checking the data collection and entry. An additional check of data processing was accomplished by separately repeating the data collection and analysis steps two separate times, thus providing a thorough accuracy check of the values collected and analyzed.

Prior to analysis, using SPSS v. 21, the data were reviewed and there were no outliers or missing data found. For each data group (IVA, DSMD, BRIEF, and QEEG) an SPSS file was set up and variables such as *Case Number*, *Pre* and *Post* variables for each scale were established. Next, data was transferred from the data collection spreadsheet to the appropriate SPSS columns. Then, *Difference* variables were created, for each scale, using the SPSS command sequence of *Transform>Compute Variable>Create Difference* to calculate the difference score between the pre and post scale values.

Normality checks. Prior to running parametric *t* tests, for repeated measures data, checking to ensure the difference scores meet the necessary assumption of normality is an important step to demonstrate the validity and reliability of the data analysis and inference (Gravetter & Wallnau, 2010). There are various techniques for checking normality. Graphical methods include histograms or Q-Q plots, while numerical methods include skewness/kurtosis coefficients or formal normality tests. As shown in Appendix C, the Q-Q plots for all the difference scores analyzed provide visual evidence of the difference scores meeting the assumption of normality. However, only formal normality tests provide conclusive evidence, with specific cut-off values (i.e. *p* values), that the requirement for a normal distribution has been met (Razali & Wah, 2011). In a study comparing four formal normality tests (i.e. Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors, and Anderson-Darling), Razali and Wah (2011) found the Shapiro-Wilk test to be the most powerful for all sample sizes and distribution types. Therefore, the Shapiro-Wilk test was also used to check the difference scores for normality. This was accomplished using the SPSS command sequence of *Analyze>Descriptive Statistics>Explore.* The Shapiro-Wilk computations for all scales, in all groups, resulted in $p > .05$ (ranging from $p = .084$ to $p = .980$); thus ensuring the difference scores met the normality assumption. Meeting this assumption provides confidence that the statistical analysis yields reliable and valid results (Razali & Wah, 2011). Therefore, the Shapiro-Wilk testing indicates the validity and reliability of the interpretation of the data as well as the inference of the data in this study was demonstrated. A breakdown of the difference scores Shapiro-Wilk *p* values are provided in Table 4.2.

Table4.2

Shapiro-Wilk Results for Difference Scores

Groups **Scales**	**Shapiro-Wilk** ***p* Values**
IVA (*n* = 10)	
Audio Attention	.429
Visual Attention	.314
Full Scale Attention	.980
DSMD (*n* = 14)	
Externalizing	.771
Internalizing	.336
Total	.582
BRIEF (*n* = 12)	
BRI	.178
MI	.934
GEC	.084
QEEG (*n* = 21)	
Absolute Power	.930
Relative Power	.778
Coherence	.437

Paired *t* tests. To compute the paired *t* tests, for the three scales in each group, the SPSS command sequence executed was *Analyze>Compare Means>Paired Samples T Test*. For each scale, the *Pre* variable was moved to the *Variable 1* position and the *Post* variable was moved to the *Variable 2* position. Given the directional nature of all hypotheses, it was necessary to divide the SPSS-computed 2-tailed *p* value by two in order to derive the 1-tailed *p* value. Finally, the Hedges' *d* effect sizes were calculated using the MetaCalc module of the Metawin 2.1 software.

The analysis for each of the psychometric assessment groups maintained alignment with the associated research questions by including the specified scales; those which most closely associate with the constructs of interest. These included the Attention scales for the IVA group, the Internalizing, Externalizing, and Total scales for the DSMD group, and the composite indices of Behavior Regulation, Metacognition, and Global Executive for the BRIEF group. For the QEEG group, analyzing whether the post z-

scores are closer to the mean maintains alignment with the z-score research question. Moreover, the paired *t* test analyses, where the means of the pre values are compared to the means of the post values, was appropriate for the one-group pretest-posttest design of this study.

Results

For all of the research questions in this study, the group means direction of change was first determined; then, the paired *t* test was performed to compare the means of the pre and post scores. Finally, the Hedges' *d* effect size (H*d*) was calculated. No outliers were found in the group means data analyzed. Line graphs showing the pretest and posttest scores, for each individual subject, are shown in Appendix D to provide a detailed picture of individual assessments.

Research question 1a: IVA group. Does 19ZNF improve attention as measured by the IVA assessment?

H_a1a: The post scores will be higher than the pre scores for the IVA assessment.

$H_0$1a: The post scores will be lower than, or not significantly different from, the pre scores of the IVA assessment.

For this research question, the scales of Auditory Attention, Visual Attention, and Full Scale were evaluated; with the threshold for clinical significance being ≤ 85. The mean post scores were higher than the pre scores for all scales; thus the change was in the predicted direction. The mean of the Auditory Attention scale pre scores was 86.50 (*SD* = 14.11), 95% CI [76.40, 96.60], and the mean of the post scores was 106.20 (*SD* = 10.76), [98.50, 113.90]. The mean of the Visual Attention scale pre scores was 83.60 (*SD* =

19.37), [69.74, 97.46], and the mean of the post scores was 103.70 (*SD* = 13.21), [94.25, 113.15]. The mean of the Full Scale pre scores was 83.40 (*SD* = 18.23), [70.36, 96.44], and the mean of the post scores was 105.60 (*SD* = 12.25), [96.84, 114.36]. Moreover, the mean pre scores for all three scales were at or below the cutoff threshold indicating clinical significance; and the mean post scores for all three scales were above the clinical cutoff threshold. The one-tailed *t* test results showed the pre and post scores differed significantly; with the Auditory Attention scale $t(9) = -4.29$, $p = .001$, H$d = 1.84$; the Visual Attention scale $t(9) = -3.00$, $p = .008$, H$d = 1.29$; and the Full Scale $t(9) = -3.78$, $p = .002$, H$d = 1.62$. Therefore, the null hypothesis was rejected in favor of the alternative hypothesis, as the post scores were higher than the pre scores for the IVA assessment; thus suggesting improvement in attention. See Figure 4.1 for a graphical representation of the pre and post scale scores.

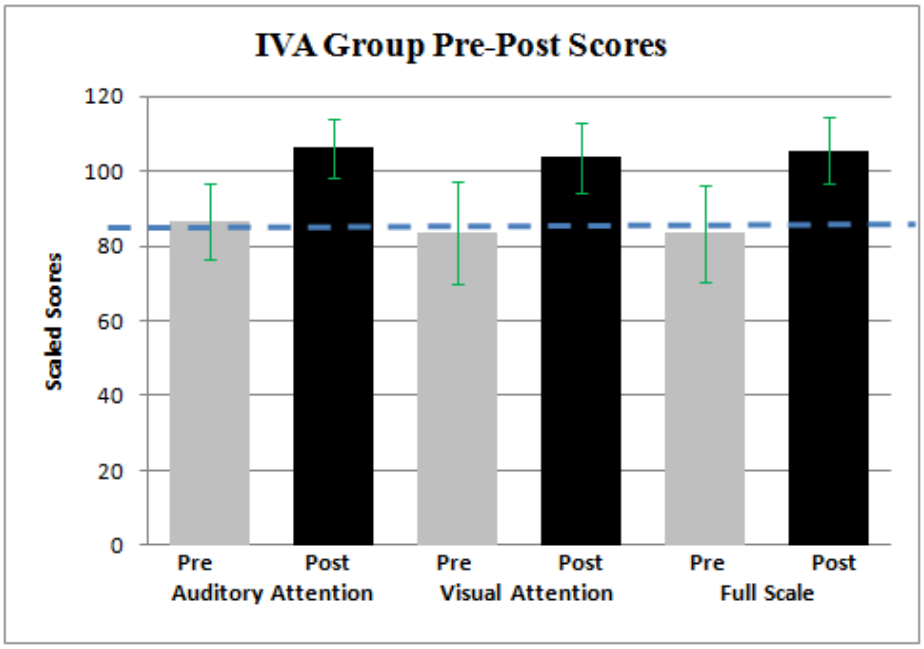


Figure 4.1. Mean IVA group standard scores before and after 19ZNF sessions. The dotted line indicates threshold for clinical significance; values at or below the line suggest clinically relevant symptoms. Post values above the line suggest improvements in attention. All post scores are statistically significant at $p \leq .008$.

Research question 1b: DSMD group. Does 19ZNF improve behavior as measured by the DSMD assessment?

H_a1b: The post scores will be lower than the pre scores for the DSMD assessment.

$H_0$1b: The post scores will be higher than, or not significantly different from, the pre scores of the DSMD assessment.

For this research question, the scales of Externalizing, Internalizing, and Total were evaluated; with the threshold for clinical significance being ≥ 60. The mean post scores were lower than the pre scores for all scales; thus the change was in the predicted direction. The mean of the Externalizing scale pre scores was 68.21 (*SD* = 15.49), 95% CI [59.27, 77.16], and the mean of the post scores was 57.71 (*SD* = 12.78), [50.28, 65.14]. The mean of the Internalizing scale pre scores was 66.21 (*SD* = 9.82), [60.55, 71.88], and the mean of the post scores was 57.29 (*SD* = 9.85), [51.60, 62.97]. The mean of the Total scale pre scores was 65.00 (*SD* = 10.58), [58.89, 71.11], and the mean of the post scores was 55.64 (*SD* = 10.76), [49.43, 61.86]. Moreover, the mean pre scores for all three scales were above the cutoff threshold indicating clinical significance, and the mean post scores for all three scales were below the clinical cutoff threshold. The one-tailed *t* test results showed the pre and post scores differed significantly; with the Externalizing scale $t(13) = 4.97$, $p = .000$, H$d = 1.83$; the Internalizing scale $t(13) = 6.43$, $p = .000$, H$d = 2.36$; and the Total scale $t(13) = 9.36$, $p = .000$, H$d = 3.42$. Therefore, the null hypothesis was rejected in favor of the alternative hypothesis, as the post scores were

lower than the pre scores for the DSMD assessment; thus suggesting improvement in behavior. See Figure 4.2 for a graphical representation of the pre and post scale scores.

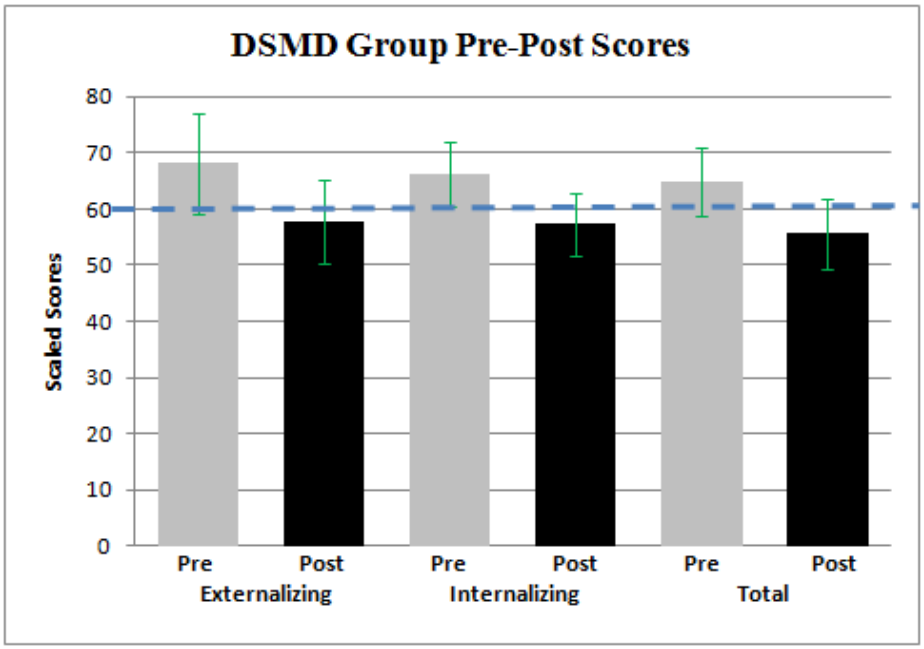


Figure 4.2. Mean DSMD group standard scores before and after 19ZNF sessions. The dotted line indicates threshold for clinical significance; values at or above the line suggest clinically relevant symptoms. Post values below the line suggest improvements in behavior. All post scores are statistically significant at $p = .000$.

Research question 1c: BRIEF group. Does 19ZNF improve executive function as measured by the BRIEF assessment?

H_a1c: The post scores will be lower than the pre scores for the BRIEF assessment.

H_01c: The post scores will be higher than, or not significantly different from, the pre scores of the BRIEF assessment.

For this research question, the scales of BRI, MI, and GEC were evaluated; with the threshold for clinical significance being ≥ 65. The mean post scores were lower than the pre scores for all scales; thus the change was in the predicted direction. The mean of the BRI scale pre scores was 71.00 (*SD* = 11.40), 95% CI [63.77, 78.23], and the mean of the post scores was 60.17 (*SD* = 10.27), [53.64, 66.69]. The mean of the MI scale pre scores was 76.08 (*SD* = 8.24), [70.85, 81.32], and the mean of the post scores was 65.67 (*SD* = 10.36), [59.08, 72.25]. The mean of the GEC scale pre scores was 75.75 (*SD* = 9.33), [69.82, 81.68], and the mean of the post scores was 64.50 (*SD* = 9.91), [58.20, 70.80]. Moreover, the mean pre scores for all three scales were above the cutoff threshold indicating clinical significance, and the mean post scores for all three scales were below the clinical cutoff threshold. The one-tailed *t* test results showed the pre and post scores differed significantly; with the BRI scale $t(11) = 4.37$, $p = .001$, H$d = 1.72$; the MI scale $t(11) = 4.39$, $p = .001$, H$d = 1.73$; and the GEC scale $t(11) = 4.66$, $p = .000$, H$d = 1.84$. Therefore, the null hypothesis was rejected in favor of the alternative hypothesis, as the post scores were lower than the pre scores for the BRIEF assessment; thus suggesting improvement in executive function. See Figure 4.3 for a graphical representation of the pre and post scale scores.

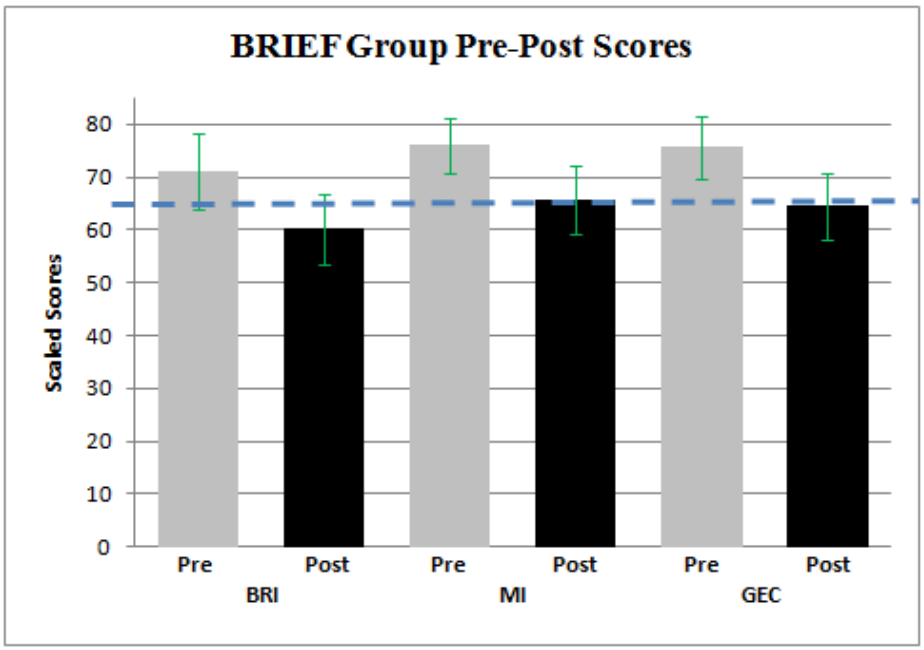


Figure 4.3. Mean BRIEF group standard scores before and after 19ZNF sessions. The dotted line indicates threshold for clinical significance; values at or above the line suggest clinically relevant symptoms. Post values below the line suggest improvements in executive function. All post scores are statistically significant at $p \leq .001$.

Research question 2: QEEG group. Does 19ZNF improve electrocortical function as measured by QEEG z-scores such that the post z-scores are closer to the mean than pre z-scores?

H_a2: The post z-scores will be closer to the mean than the pre z-scores.

H_02: The post z-scores will be farther from the mean, or not significantly different from, the pre z-scores.

For this research question, the QEEG metrics of Absolute power, Relative power, and Coherence were evaluated; with the targeted transformed z-score threshold value being $z \geq 1.00$. The mean post z-scores were lower than the pre z-scores for all metrics; thus the change was in the predicted direction and the z-scores were closer to the mean. The mean of the Absolute power pre z-scores was 1.46 (*SD* = 0.28), 95% CI [1.33, 1.59],

and the mean of the post scores was 1.03 (*SD* = 0.37), [0.87, 1.20]. The mean of the Relative power pre z-scores was 1.51 (*SD* = 0.22), [1.41, 1.61], and the mean of the post scores was 1.13 (*SD* = 0.35), [0.97, 1.29]. The mean of the Coherence pre z-scores was 1.46 (*SD* = 0.14), [1.40, 1.53], and the mean of the post scores was 0.96 (*SD* = 0.32), [0.82, 1.11]. Moreover, the mean pre scores for all metrics were above 1.00, and the mean post scores for all metrics approached or were below 1.00. The one-tailed *t* test results showed the pre and post scores differed significantly; with the Absolute power $t(20) = 7.73$, $p = .000$, H$d = 2.29$; the Relative power $t(20) = 5.22$, $p = .000$, H$d = 1.76$; and the Coherence $t(20) = 6.55$, $p = .000$, H$d = 1.88$. Therefore, the null hypothesis was rejected in favor of the alternative hypothesis, as the post z-scores were closer to the mean than the pre z-scores; thus suggesting improvement in electrocortical functioning. See Figure 4.4 for a graphical representation of the pre and post scale scores.

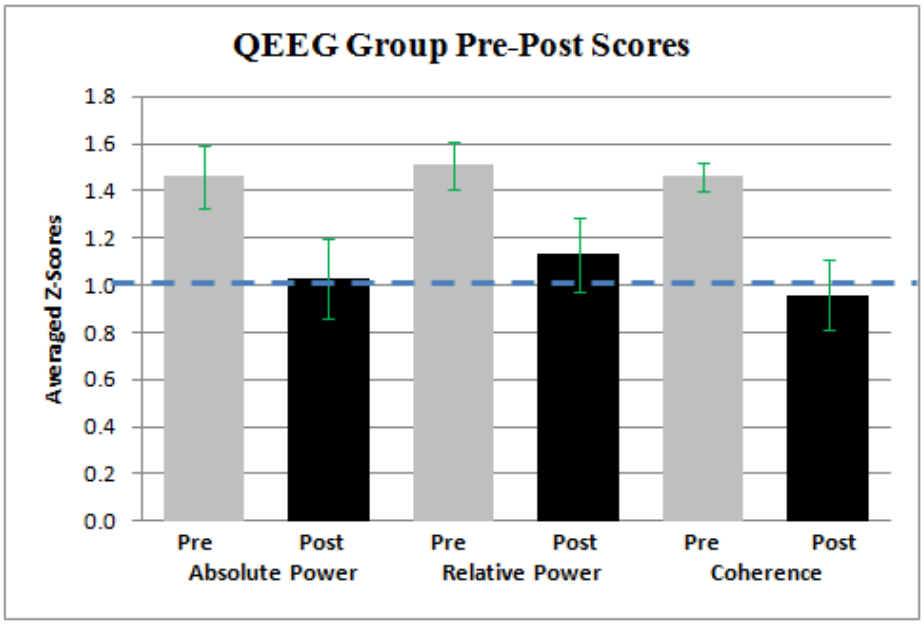


Figure 4.4. Mean QEEG group targeted z-scores before and after 19ZNF sessions. The dotted line indicates threshold for inclusion as targeted z-scores; values above the line suggest electrocortical dysfunction. Post values at or below the line suggest improvements in electrocortical function. All post scores are statistically significant at $p = .000$.

Summary

The research questions for this study asked if the independent variable of 19ZNF improved attention, behavior, executive function, and electrocortical function. The dependent variables to test the hypotheses included the scaled scores from the IVA, DSMD, and BRIEF clinical assessments and QEEG z-scores. The difference scores were normally distributed, thus supporting the use of one-tailed *t* tests to compare the pre to the post scores for each of the dependent variables.

For all pre-post comparisons, the direction of change in the scores was in the predicted direction for all hypotheses. Moreover, for all the outcome measures, the averaged scores were beyond the clinically significant threshold before 19ZNF and changed to no longer being so after 19ZNF. Finally, for all research questions, the null hypothesis was rejected, in favor of the conclusion that 19ZNF improved attention, behavior, executive function, and electrocortical function (respective to each hypothesis). All differences were statistically significant, with results ranging from $p = .000$ to $p = .008$; and H*d* values ranging from 1.29 to 3.42. Table 4.3 provides a cumulative summary of the results of these findings for all groups.

In the chapter that follows, a discussion of these findings will be presented. Conclusions and interpretations regarding the contributions of this research will be offered. Furthermore, a review of the implications (practical, theoretical, and future) of this research, and recommendations for future research and practice will be provided.

Table 4.3

Summary of Results – All Groups

Groups **Scales**	**PRE Scores** ***M (SD)***	**POST Scores** ***M (SD)***	***t(df)***	***p***	***Hedges' d***
IVA					
Audio Attention	86.50 (14.11)	106.20 (10.76)	-4.29 (9)	.001	1.84
Visual Attention	83.60 (19.37)	103.70 (13.21)	-3.00 (9)	.008	1.29
Full Scale Attention	83.40 (18.23)	105.60 (12.25)	-3.78 (9)	.002	1.62
DSMD					
Externalizing	68.21 (15.49)	57.71 (12.87)	4.97 (13)	.000	1.83
Internalizing	66.21 (9.82)	57.29 (9.85)	6.43 (13)	.000	2.36
Total	65.00 (10.58)	55.64 (10.76)	9.36 (13)	.000	3.42
BRIEF					
BRI	71.00 (11.40)	60.17 (10.27)	4.37 (11)	.001	1.72
MI	76.08 (8.24)	65.67 (10.36)	4.39 (11)	.001	1.73
GEC	75.75 (9.33)	64.50 (9.91)	4.66 (11)	.000	1.84
QEEG Z-Scores					
Absolute Power	1.46 (0.28)	1.03 (0.37)	7.73 (20)	.000	2.29
Relative Power	1.51 (0.22)	1.13 (0.35)	5.22 (20)	.000	1.76
Coherence	1.46 (0.14)	0.96 (0.32)	6.55 (20)	.000	1.88

Chapter 5: Summary, Conclusions, and Recommendations

Introduction

The primary problem this research sought to address was how it was not known, by way of statistical evaluation of either clinical assessments or QEEG z-scores, if 19ZNF was an effective NF technique. This problem, manifest as a lack of literature, leaves clinicians and prospective NF clients alike without research-based evidence to evaluate 19ZNF. Currently, mostly qualitative case-study reports have been found in the literature. Thus, this study has importance in its aim to fill this empirical gap.

As has been discussed, NF is gaining recognition as an evidence-based intervention grounded in learning theory. Among the different models developed over the last 40 years, 19ZNF is one of the newest. Yet, while 19ZNF is reported to lead to improved clinical outcomes in fewer sessions than traditional NF, and a growing number of clinicians are adding this model to their practice, the peer-review literature is lacking regarding the efficacy of the model. This study was different in its use of group means data to directly compare pre and post outcome measure variables, to include QEEG data, to begin an evaluation of efficacy of 19ZNF. The use of a quasi-experimental design in this research, which has not been typical in prior 19ZNF evaluations, provides baseline research for investigating the efficacy of 19ZNF. The use of quantitative methods, with group means data, contributes to the base of knowledge regarding 19ZNF by providing statistical analysis, which allows for greater generalization over qualitative and/or case study investigations.

Summary of the Study

This chapter aims to first present a summary and conclusions of the study. Next, practical, theoretical, and future implications will be reviewed. Finally, future research and practice recommendations will be discussed.

This retrospective pretest-posttest study investigated if 19ZNF improved attention, behavior, executive function, and electrocortical functioning. To that end, the research questions asked if 19ZNF improved: Attention as measured by the IVA, behavior as measured by the DSMD, executive function as measured by the BRIEF, and electrocortical function as measured by QEEG z-scores. Paired *t* tests were performed to compare the means of four outcome measures; which included three clinical assessments (IVA, DSMD, and BRIEF) and QEEG z-scores. Each of the clinical assessments framed a sample group such that the efficacy of 19ZNF was evaluated, as it relates to the particular neuropsychological constructs of attention ($n = 10$), behavior ($n = 14$), executive function ($n = 12$), and additionally as related to electrocortical functioning ($n = 21$). The focus of the IVA sample group was attention, and the scales specific to attention were the Auditory Attention, Visual Attention, and Full Scale. The focus of the DSMD sample group was behavior, and the scales specific to behavior were the Externalizing, Internalizing and Total. The focus of the BRIEF sample group was executive function, and the composite scales included were BRI, MI, and GEC. The focus of the QEEG sample group was electrocortical function, and the QEEG metrics included were Absolute power, Relative power, and Coherence.

Overall, the makeup of the sample was a diagnostically diverse mixture of adults and children, with most diagnoses related to ADHD. The sample consisted of more

children (QEEG = 15, IVA = 5, DSMD = 14, BRIEF = 10) than adults (QEEG = 6, IVA = 5, DSMD = 0, BRIEF = 2). Other sample characteristics consistent across all groups are they were evenly divided with respect to gender, were primarily ethnically white, and were mostly medium SES.

In Chapter 1, an orienting framework of the study was presented to include the problem statement and study purpose, as well as the methodology rationale and nature of the research design. In Chapter 2, a review of the literature was presented. The history and background of ZNF was first addressed; then, the theoretical foundations and conceptual frameworks of NF and QEEG were presented. Theoretical frameworks supporting the models of traditional NF, QNF, and ZNF were then reviewed, as were key NF themes related to applications of QNF and the emergence of 19ZNF. Moreover, outcome measures suitable for ZNF research were discussed. The focus of Chapter 3was the methodology of the study and Chapter 4 presented research findings and results.

Summary of Findings and Conclusion

Operant conditioning is the theoretical foundation of NF, with demonstrated efficacy in improving brain functioning and clinical symptoms, through the resulting electrocortical changes. However, whether this also holds true for the new 19ZNF model has been an outstanding question. As discussed in Chapter 1, and again in Chapter 3, the aim of this study was to provide the beginnings of an evidence-based foundation for the efficacy of 19ZNF. The focus was to evaluate if 19ZNF would result in improved clinical symptoms and electrocortical function as measured by the identified outcome measures. In general, the findings of this study were that attention, behavior, executive function, and electrocortical function all improved after approximately ten 19ZNF sessions; with

the number of sessions ranging from an average of 9.70 to 11.83 sessions across the four groups. This study also supported the clinical reports of Thatcher (2013) and Wigton (2013) that 19ZNF results in improvement in clinical symptoms in fewer sessions than the 40+ sessions typical in traditional NF. Also notable, is that the frequency of the sessions was an average of once per week, rather than the two to three times per week as is typical of traditional NF or QNF. Each finding will next be reviewed separately, to further discuss the significance of this study as related to the identified constructs of attention, behavior, executive function, and electrocortical function.

Research question 1a: IVA group. Does 19ZNF improve attention as measured by the IVA assessment? In answering this research question, as seen in Table 4.3, the post scores were higher than the pre scores for the IVA, thus lending support for attention being improved. Although this group was made up of subjects with varying diagnoses (though mostly associated with ADHD), as a collective group, they all initially exhibited symptoms of attention dysfunction; as all the group means Attention scales scores fell at or below the clinically significant range (Auditory Attention = 86.50, Visual Attention = 83.60, and Full Scale = 83.40). As was expected, 19ZNF resulted in a positive clinical outcome of improved attention, as the subjects' performance on the posttest assessment significantly improved. After 19ZNF, all the included group means Attention scales were no longer in the clinically significant range (Auditory Attention = 106.2, Visual Attention = 103.70, and Full Scale = 105.60). The effect sizes for the three scales (1.84, 1.29, and 1.62, respectively) are all considered very large. Therefore, the results of this research question were both clinically and statistically significant.

Given that no prior 19ZNF studies were found which analyzed IVA data as an outcome measure, no direct comparison to prior research is possible. Moreover, there were no QNF studies found incorporating the IVA as an outcome measure. In looking at traditional NF studies, while Knezevic, et al. (2010) incorporated the IVA in their study, they did not use the any of the composite Attention scales. The Fritson et al. (2008) study is not a relevant comparison as they used a sample of non-clinical college students. Finally, in the research of Steiner, et al. (2011), the only comparable scale used was the Attention Full scale; yet, with an $n = 6$, while the post scores were in the desired direction, the pre-post difference scores were not statistically significant.

Research question 1b: DSMD group. Does 19ZNF improve behavior as measured by the DSMD assessment? In answering this research question, as seen in Table 4.3, the post scores were lower than the pre scores for the DSMD, thus lending support for behavior being improved. Although this group was made up of subjects with varying diagnoses, as a collective group, they all initially exhibited symptoms of behavioral issues; as all the included group means scales scores fell at or above the clinically significant range (Externalizing = 68.21, Internalizing = 66.21, and Total = 65.00). As was expected, 19ZNF resulted in a positive clinical outcome of improved behavior, as the subjects' scores on the posttest assessment significantly improved. After 19ZNF, all the included group means scales were no longer in the clinically significant range (Externalizing = 57.71, Internalizing = 57.29, and Total = 55.64). The effect sizes for the three scales (1.83, 2.36, and 3.42, respectively) are all interpreted as being very large; and are the largest effect sizes in this study. Therefore, the results of this research question were both clinically and statistically significant.

To date, no prior NF studies (ZNF, QNF, or traditional NF) have conducted outcome measure analysis with the DSMD; as such, there are no relevant existing studies with which to directly contrast or compare. However, the DSMD scales of Externalizing, Internalizing, and Total correlate well to the similarly named scales of the CBCL. Huang-Storms et al. (2006) used the CBCL as an outcome measure in their retrospective pretest posttest study evaluating traditional NF. All post scores were in the desired direction and difference scores for all scales were statistically significant ($p < .01$) with medium to large effect sizes (Externalizing, Cohen's $d = .94$; Internalizing, Cohen's $d = .59$; Total, Cohen's $d = .78$).

Research question 1c: BRIEF group. Does 19ZNF improve executive function as measured by the BRIEF assessment? In answering this research question, as seen in Table 4.3, the post scores were lower than the pre scores for the BRIEF, thus lending support for executive function being improved. Although this group was made up of subjects with varying diagnoses, as a collective group, they all initially exhibited symptoms of compromised executive function; with all the included group means scales scores falling at or above the clinically significant range (BRI = 71.00, MI = 76.08, and GEC = 75.75). As was expected, 19ZNF resulted in a positive clinical outcome of improved executive function, as the subjects' scores on the posttest assessment significantly improved. After 19ZNF, all the included group means scales were no longer in the clinically significant range (BRI = 60.17, MI = 65.67, and GEC = 64.50). The effect sizes for the three scales (1.72, 1.73, and 1.84, respectively) are all interpreted as being very large. Therefore, the results of this research question were both clinically and statistically significant.

Here too, no prior 19ZNF studies were found which conducted outcome measure analysis with the BRIEF; thus, no direct comparison to prior research is possible. However, the Orgim and Kestad (2013) study, which compared QNF to medication for ADHD, included the BRI and MI scales of the BRIEF among various outcome measures. For the BRIEF scales analyzed, while the post scores were in the desired direction for both groups, the difference between NF and medication groups were not significant. The Drechsler et al. (2007) study compared SCP NF to group therapy and incorporated the BRI and MI scales of the BRIEF as two of their outcome measures. Their findings indicated a statistically significant ($p = .004$) improvement for NF, more than group therapy, on the MI scale of the BRIEF; whereas there were no significant differences for NF versus group therapy for the BRI scale. Finally, Steiner et al. (2011) incorporated the GEC scale of the BRIEF as one of many outcome measures in comparing traditional NF to computerize attention training to a waitlist control. For all groups, for the primary parent and co-parent ratings, all post scores moved in the desired direction, however, only the computerized attention training resulted in a significant difference ($p < .05$) for the GEC scale.

Research question 2: QEEG group. Does 19ZNF improve electrocortical function as measured by QEEG z-scores, such that the post z-scores are closer to the mean than pre z-scores? In answering this research question, as seen in Table 4.3, the post z-scores were closer to the mean than the pre z-scores, thus lending support for electrocortical function being improved. Although this group was made up of subjects with varying diagnoses, as a collective group, they all exhibited electrocortical dysregulation; with all the targeted z-scores group means falling above the z-score

threshold (Absolute power = 1.46, Relative power = 1.51, and Coherence = 1.46). As was expected, 19ZNF resulted in a positive clinical outcome of improved electrocortical function, as the subjects' averaged targeted z-scores on the posttest assessment significantly improved. After 19ZNF, the targeted z-score group means for Absolute power (1.03) and Coherence (0.96) were at or below the z-score threshold, with the Relative power (1.13) approaching the threshold. The effect sizes for the three scales (Absolute power = 2.29, Relative power = 1.76, and Coherence = 1.88) are all interpreted as being very large. Therefore, the results of this research question were both clinically and statistically significant. Moreover, these findings suggested that, as a group, the subjects' QEEG z-scores normalized as a result of 19ZNF; and perhaps more importantly, the normalization was accompanied by clinical symptom improvement.

As has been stated, few NF studies make use of QEEG metrics as outcome measures. More so, as of this writing, no prior NF studies (ZNF, QNF, or traditional NF) have been found incorporating a measure of overall QEEG normalization. Thus, there are no relevant existing studies with which to contrast or compare.

Conclusions. The literature reviewed for this study found both traditional NF and QNF studies consistently employed retrospective pretest-posttest designs. This research was consistent with those prior works. Significant differences were found between the pre and post scores, thus indicating positive clinical outcomes. However, this research was also innovative in that it made use of QEEG metrics, as outcome measures, to provide an overall measure of the distance from the mean, for determining overall normalization of the z-scores. Here too, the pre to post score differences were significant

for all metrics indicating normalization of the QEEG z-scores, thus indicating improved electrocortical function.

Arns et al. (2009, 2012) have discussed effect sizes in studies evaluating NF to treat ADHD. For traditional NF models, H*d* effect sizes were 0.7 and 1.0 for hyperactive and attention symptoms, respectively; yet for the QNF models, H*d* effect sizes were 1.2 and 1.8 (hyperactive and attention symptoms, respectively). In this research, H*d* effect sizes ranged from 1.29 to 3.42, with an average of 1.97. Therefore, the effect sizes for this study were similar, or greater, than what has been reported for QNF and traditional NF models. Moreover, if NF efficacy is defined in terms of large effect sizes when comparing pre-post outcome measure data (Arns et al., 2012), then the effect sizes of this study support 19ZNF as being effective.

Therefore, as was proposed in Chapter 1, it is reasonable to conclude that the theory of operant conditioning, upon which NF is founded, can be expanded to include 19ZNF. It is also reasonable to conclude that, in the context of this study, the findings supported the efficacy of 19ZNF in improving attention, behavior, executive function, and electrocortical function. Thus, this research addressed the literature gap and begins to lend credence to the position that 19ZNF could be considered an evidence-based intervention. Further, this study demonstrated that QEEG z-scores data can be used for group comparison studies, in a way not previously developed; thus, this study has the potential for cultivating future QEEG-based research.

Implications

The objective of this research was a comparison of outcome measures before and after 19ZNF to evaluate the efficacy of this NF intervention. In reviewing the theoretical

framework discussed in the literature review, certain elements are pertinent to the findings of this research. Hughes and John (1999) demonstrated EEG/QEEG measures to be sensitive to psychiatric disorders. The QNF model (which informs the 19ZNF model) is founded on the premise that electrocortical dysfunctions correspond with clinical symptoms and mental disorders (Coben & Myers, 2010; Collura, 2010; Walker, 2010a), such that clinical symptoms can be linked to brain dysregulation (Thatcher, 2013). Further, when NF results in symptom resolution, together with QEEG normalizing, this represents an improvement in electrocortical functioning (Arns et al., 2012; Walker, 2010a). Therefore, the findings of this study (with the 19ZNF protocol of QEEG normalization) were consistent with the multiple reports in the literature suggesting QEEG normalization protocols bring about clinical benefits (Arns et al., 2012; Breteler et al., 2010; Collura, 2008; Orgim & Kestad, 2013; Surmeli et al., 2013; Surmeli & Ertem, 2009, 2010; Walker, 2009. 2010a, 2011, 2012a).

Theoretical implications. QEEG normalization is a theoretical construct which has grown in popularity with the advent of the QNF model; as has the use of individually tailored QEEG-based protocols to bring about that normalization. Additionally, clinical reports have suggested 19ZNF may exhibit better performance than traditional NF. These findings supported 19ZNF as a NF modality which can bring about both QEEG normalization and symptom improvement. More so, it can do so quite efficiently, as evidenced by the results of this study occurring on average of within 10 sessions, at a target frequency of once per week.

As discussed in the literature review, the greater specificity that QEEG-based methods allowed in treatment also creates methodological challenges due to the need to

account for both positive and negative z-scores. This study's method of transforming the z-scores to the absolute value, then tracking pre to post changes of the targeted z-scores, presented an innovative methodology for measuring overall normalization of the QEEG. If further validated, this approach has the potential to open new avenues for QEEG-based research, both within the NF community as well as the broader neuroscience fields.

The implications of this study, as related to cognition and instruction, are twofold. First, the findings suggested 19ZNF improves the attention and executive function components of cognition. Second, when cognition improves, more mental resources are made available for an individual to better engage in instructional processes. The findings of this study also suggested that 19ZNF can improve behavior. In group educational settings, when disruptive behavior improves, distractions to other learners are reduced and the effectiveness of the instructional environment can be enhanced. Therefore, this study lent support to 19ZNF as benefiting both cognition and instruction.

Practical implications. This research begins to address the literature gap regarding evidence-based findings of 19ZNF. Thus, this study can provide NF clients and clinicians with information regarding its efficacy in improving attention, behavior, executive function, and electrocortical function. Furthermore, it suggests that 19ZNF may address the need for 40+ sessions for success with NF. If 19ZNF is shown to be an evidence-based intervention which requires fewer sessions than tradition NF or QNF, clients will benefit through the associated cost savings. Also of note, while not a specific focus in this research, is that the 19ZNF in this study occurred at a frequency of only once per week, rather than the two to three times per week as other models. These

aspects, taken together, may potentially serve to reduce resistance of third-party payers to include NF as covered services.

Future implications. Future implications of this study depend on future research. This study only provided the beginning steps of forming an evidence-based framework for 19ZNF. As will be discussed below, much remains to be investigated and evaluated through further research. However, that being said, this study has the potential of widening the acceptance of 19ZNF, as well as opening new frontiers for QEEG-based research.

Strengths of this study include being a first quantitative analysis of group means data from 19ZNF, of which, as of this writing, none has been found. Thus, this research contributed in taking the empirical evaluation of 19ZNF beyond clinical reports and case study presentations. Moreover, data coming from a real-world clinical setting suggests clinicians employing this new model may have similar results. Given the pretest-posttest design, and the group means averaged time between pre and post assessments ranged from 13 to 16 weeks (see table 4.1), the previously identified limitation of potential maturational or history effects likely had minimal impact on the findings. This, then, increased the credibility of the conclusions. However, remaining weaknesses, inherent in retrospective studies in clinical settings, included limitations already discussed, such as small sample size, lack of a separate control groups (lack of randomization), or comparison to traditional and QNF models. Therefore, recommendations for further research are next provided.

Recommendations

As discussed in the Limitations section of Chapter 3, the question of efficacy cannot be fully explored without further research. More so in investigating 19ZNF as being superior to other NF approaches. Therefore, specific recommendations for further research are presented. Additionally, recommendations for practice will also be reviewed.

Recommendations for future research. As has been discussed, this study was only a beginning step toward proving 19ZNF as efficacious; thus the recommendations herein serve to propose next steps in forwarding this line of research. A notable significance of this study, in advancing scientific knowledge, was that it filled the gap of a lack of quantitative studies evaluating 19ZNF. However, the gap is large and more research is needed. Therefore, all the following recommendations would be best implemented through the use of quantitative methodologies, in order to apply evidence-based strategies and statistical analysis to evaluate outcomes of 19ZNF as a treatment intervention.

A single study is insufficient to fully validate the efficacy of any treatment intervention. Thus, replication of this study would add to the scientific integrity of the results; however, doing so with larger sample sizes would, of course, be recommended. Next, follow-up studies are a needed area of focus. While 19ZNF may be effective in the short-term, the question of whether the benefits hold over time is still outstanding. With 19ZNF being new among other approaches, ones backed by more research, direct comparisons to the traditional or QNF models are needed; particularly with randomized assignments. Additional suggestions for randomized control group research are for comparisons to waitlist groups. However, randomized controlled methods are less

feasible in clinical settings; and as such, these studies will likely require university and/or grant-supported research settings (more conducive to true experimental designs) to complete. Other comparison research should also explore comparisons of 19ZNF using surface montages (as with this study) to 19ZNF using inverse-solution montages (e.g. LORETA).

As has been discussed, few NF studies employ QEEG metrics as a direct outcome measure; and even fewer do so in analyzing group means data. Therefore, an additional notable significance of this study, in advancing scientific knowledge, is the novel development of a measure of overall QEEG normalization, by tracking the pre-post values of the targeted transformed z-scores. Here too, though, replication and further validation is needed. Also recommended is an investigation of whether $z \pm 1.00$ is an optimal threshold value to determine targeted z-scores.

Recommendations for practice. Both NF clinicians and prospective clients will benefit from reviewing this study. Researchers will also find this study of interest in furthering what is known about NF, and/or using QEEG metrics as outcome measures in NF or other QEEG-based investigations. For clinicians employing 19ZNF, who do not already do so, incorporating the regular use of pre and post outcome measures, and gathering pre-session baseline QEEG data, is important to furthering what is known about 19ZNF. Currently, 19ZNF is in its infancy, and likely will face resistance in the scientific community, much the same as traditional NF has until only recent years. The settings where conventional experimental work occurs (i.e. grant-funded and/or university laboratories) may be less likely to embrace research with newer 19ZNF models, in favor of traditional NF models; at least in the short term. As a result, the

clinical setting is currently the primary source of data to evaluate 19ZNF. Therefore, performing quality pre-post assessments, and then moving forward in research with that data, will be necessary to advance the acceptance of 19ZNF by the wider scientific community.

As was discussed in Chapter 1, this study has the potential of opening doors to future QEEG-based research, in demonstrating that z-scores from QEEG data can be used for group comparison studies, in a way not previously developed. In moving forward with this line of research, this study proposed a method for using QEEG metrics for measuring the degree of normalization. Therefore, incorporating QEEG data as outcome measures is a practical reality for NF researchers. Thus, practice recommendations are for including these metrics in future research.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Publishing.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Pearson.

Arns, M., de Ridder**,** S., Strehl, U., Breteler, M., & Coenen. A. (2009). Efficacy of neurofeedback treatment in ADHD: The effects on inattention, impulsivity and hyperactivity: A meta-analysis. *Clinical EEG and Neuroscience, 40*(3), 180-189.

Arns, M., Drinkenburg, W., & Kenemans, J. L. (2012). The effects of QEEG-informed neurofeedback in ADHD: An open-label pilot study. *Applied Psychophysiology and Biofeedback*, *37*, 171-180. doi:10.1007/s10484-012-9191-4

Association for Applied Psychophysiology and Biofeedback (AAPB). (2011). *What is biofeedback?* Association for Applied Psychophysiology and Biofeedback, [website home page], retrieved from: www.aapb.org

Baehr, E., Rosenfeld, J. P., & Baehr, R. (1997). The clinical use of an alpha asymmetry protocol in the neurofeedback treatment of depression: Two case studies. *Journal of Neurotherapy, 2*(3), 10-23.

Berger H. (1929) Über das elektroenkephalogramm des menschen [About the elektroenkephalogram of humans]. *Archiv für Psychiatrie und NerveMrankheilen, 87*, 527-570.

Besenyei, M. , Varga, E., Fekete, I., Puskás, S., Hollódy, K., Fogarasi, A., …Clemens, B. (2012). EEG background activity is abnormal in the temporal and inferior parietal cortex in benign rolandic epilepsy of childhood: A LORETA study. *Epilepsy Research, 98*(1), 44-49. doi:10.1016/j.eplepsyres.2011.08.013

Brandeis, D. (2011). Neurofeedback training in ADHD: More news on specificity. *Clinical Neurophysiology, 122*, 856-857. doi:10.1016/j.clinph.2010.08.011

Bresadola, M. (2008). Animal electricity at the end of the eighteenth century: The many facets of a great scientific controversy. *Journal of the History of the Neurosciences*, *17*(1), 8-32. doi:10.1080/09647040600764787

Breteler, M. H. M., Arns, M., Peters, S., Giepmans, I., & Verhoeven, L. (2010). Improvements in spelling after QEEG-based neurofeedback in dyslexia: A randomized controlled treatment study. *Applied Psychophysiology and Biofeedback*, *35*, 5–11. doi:10.1007/s10484-009-9105-2

Budzynski, T. H. (1999). From EEG to neurofeedback. In J. R. Evans & A. Abarbanel (Eds.), *Introduction to quantitative EEG and neurofeedback* (pp. 65-79). San Diego, CA: Academic Press.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally.

Cannon, R. L., Baldwin, D. R., Shaw, T. L., Diloreto, D. J., Phillips, S. M., Scruggs, A. M., & Riehl, T. C. (2012). Reliability of quantitative EEG (qEEG) measures and LORETA current source density at 30 days. *Neuroscience Letters, 518*, 27-31. doi:10.1016/j.neulet.2012.04.035

Carr, L. T. (1994). The strengths and weaknesses of quantitative and qualitative research: What method for nursing? *Journal of Advanced Nursing, 20*, 716-721.

Caton, R. (1875). The electric currents of the brain. *British Medical Journal, 2*, 278.

Coben, R. & Myers, T. E. (2010). The relative efficacy of connectivity guided and symptom based EEG biofeedback for autistic disorders. *Applied Psychophysiology and Biofeedback, 35*, 13–23. doi:10.1007/s10484-009-9102-5

Collura, T. F. (1993). History and evolution of electroencephalographic instruments and techniques. *Journal of Clinical Neurophysiology, 10*(4), 476-504.

Collura, T. F. (1995). History and evolution of computerized electroencephalography *Journal of Clinical Neurophysiology, 12*(3), 214-229.

Collura, T. F. (2008). Neuronal dynamics in relation to normative electroencephalography assessment and training. *Biofeedback, 36*(4), 134-139.

Collura, T. F. (2010). Conclusion: QEEG-guided neurofeedback in context and in practice. *Applied Psychophysiology and Biofeedback, 35*, 37-38. doi: 10.1007/s10484-009-9108-z

Collura, T. F. (2014). *Technical foundations of neurofeedback*. New York, NY: Routledge.

Collura, T. F., Guan, J. G., Tarrant, J., Bailey, J., & Starr, F. (2010). EEG biofeedback case studies using live z-score training and a normative database. *Journal of Neurotherapy*, *14*(1), 22-46. doi:10.1080/10874200903543963

Collura, T. F., Kaiser, D., Lubar, J., & Evans, J. (2011). *Biofeedback Glossary: Dictionary of Biofeedback / Neurofeedback Terms.* [HTML webpage].

Association for Applied Psychophysiology and Biofeedback (AAPB). Retrieved from: http://www.aapb.org/i4a/pages/index.cfm?pageid=3462

Collura, T. F., Thatcher, R. W., Smith, M. L., Lambos, W. A., & Stark, C. R. (2009). EEG Biofeedback training using live z-scores and a normative database. In T. H. Budzynski, H. K. Budzynski, J. R. Evans & A. Abarbanel (Eds.), *Introduction to quantitative EEG and neurofeedback: Advanced theory and applications* (2nd ed., pp. 29-59). Burlington, MA: Elsevier.

Cooper, C. (2001). Review of the Devereeux Scales of Mental Disorders. In B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* (pp. 408-410). Lincoln, NE: Buros Institute of Mental Measurements.

Corsi-Cabrera, M. M., Galindo-Vilchis, L. L., del-Río-Portilla, Y. Y., Arce, C. C., & Ramos-Loyo, J. J. (2007). Within-subject reliability and inter-session stability of EEG power and coherent activity in women evaluated monthly over nine months. *Clinical Neurophysiology*, *118*(1), 9-21. doi:10.1016/j.clinph.2006.08.013

Demos, J. N. (2005). *Getting started with neurofeedback*. New York, NY: W. W. Norton.

Donders, J. (2002). The Behavior Rating Inventory of Executive Function: Introduction. *Child Neuropsychology, 8*(4), 229-230.

Drechsler, R., Straub, M., Doehnert, M., Heinrich, H., Steinhausen, H., & Brandeis, D. (2007). Controlled evaluation of a neurofeedback training of slow cortical potentials in children with attention deficit/hyperactivity disorder. *Behavioral and Brain Functions, 3*(35), 1-13. doi:10.1186/1744-9081-3-35

Farley, J. W., & Connolly, J. J. (2005). Low impedances: How important are they with digital recordings? *American Journal of Electroneurodiagnostic Technology, 45*(2), 139-44.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fritson, K. K., Wadkins, T. A., Gerdes, P., & Hof, D. (2008). The impact of neurotherapy on college students' cognitive abilities and emotions. *Journal of Neurotherapy, 11*(4), 1-9. doi:10.1080/10874200802143998

Gearing, R. E., Mian, I. A., Barber, J., & Ickowicz, A. (2006). A methodology for conducting retrospective chart review research in child and adolescent psychiatry. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 15*(3), 126-134.

Gevensleben, H., Holl, B., Albrecht, B., Schlamp, D., Kratz, O., Studer, P.,Wangler, S,… Heinrich, H. (2009). Distinct EEG effects related to neurofeedback training in children with ADHD: A randomized controlled trial. *International Journal of Psychophysiology, 74*, 149–157. doi:10.1016/j.ijpsycho.2009.08.005

Gevensleben, H., Rothenberger, A., Moll, G. H., & Heinrich, H. (2012). Neurofeedback in children with ADHD: Validation and challenges. *Expert Reviews of Neurotherapeutics, 12*(4), 447–460. doi:10.1586/ERN.12.22

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *BRIEF: Behavior Ration Inventory of Executive Function: Professional manual*. Lutz, FL: Psychological Assessment Resources PAR.

Gravetter, F. J., & Wallnau, L. B. (2010). *Statistics for the behavioral sciences* (9th ed.). Belmont, CA: Wadsworth Cengage Learning.

Hallman, D. W. (2012). 19-Channel neurofeedback in an adolescent with FASD. *Journal of Neurotherapy, 16*(2), 150-154. doi:10.1080/10874208.2012.677646

Hammer, B., Colbert, A., Brown, K., & Ilioi, E. (2011). Neurofeedback for insomnia: A pilot study of z-score SMR and individualized protocols. *Applied Psychophysiology and Biofeedback*, *36*(4), 251-64. doi:10.1007/s10484-011-9165-y

Hammond, C. D. (2010). The need for individualization in neurofeedback: Heterogeneity in QEEG patterns associated with diagnoses and symptoms. *Applied Psychophysiology and Biofeedback, 35*(31–36). doi:10.1007/s10484-009-9106-1

Hergenhahn, B. R. (2009). *An introduction to the history of psychology,* (6th ed.). Belmont, CA: Wadsworth.

Huang-Storms, L., Bodenhamer-Davis, E., Davis, R., & Dunn, J. (2006). QEEG-guided neurofeedback for children with histories of abuse and neglect: Neurodevelopmental rationale and pilot study. *Journal of Neurotherapy*, *10*(4), 3-16. doi:10.1300/J184v10n04_02

Hughes, J. R., & John, E. R. (1999). Conventional and quantitative electroencephalography in psychiatry. *The Journal of Neuropsychiatry and Clinical Neurosciences, 11*(2), 190-208.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of Meta-Analysis.* (2nd ed.). Thousand Oaks, CA: SAGE Publications, Inc.
doi: http://dx.doi.org.library.gcu.edu:2048/10.4135/9781412985031

John, E. R. (1977). Neurometrics: Quantitative electrophysiological analyses. In E. R. John & R. W. Thatcher (Eds.), *Functional neuroscience,* (Vol. II). New Jersey: L. Erlbaum Assoc.

Johnstone, J., & Gunkelman, J. (2003). Use of databases in QEEG evaluation. *Journal of Neurotherapy*, *7*(3), 31-52. doi:10.1300/J184v07n03_02

Kamiya, J. (1968). Conscious control of brainwaves. *Psychology Today, 1*, 56-60.

Kamiya, J. (1969). Operant control of the EEG alpha rhythm and some of its reported effects on consciousness. In C. T. Tart (Ed.), *Altered states of consciousness* (p. 489-501). New York: Wiley.

Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). Orlando, FL: Holt, Reinhardt, & Winston.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research* (4th ed.). United States: Wadsworth Thomson Learning.

Kirk, R. (2009). Experimental design. In R. E. Millsap, & A. Maydeu-Olivares (Eds.), *The Sage handbook of quantitative methods in psychology.* (pp. 23-46). London: Sage Publications Ltd.
doi: http://dx.doi.org.library.gcu.edu:2048/10.4135/9780857020994.n2

Knezevic, B., Thompson, L., & Thompson, M. (2010). Pilot project to ascertain the utility of Tower of London test to assess outcomes of neurofeedback in clients

with Asperger's Syndrome, *Journal of Neurotherapy, 14*(1), 3-19. doi:10.1080/10874200903543922

Koberda, J. L., Hillier, D. S., Jones, B., Moses, A., & Koberda, L. (2012). Application of neurofeedback in general neurology practice. *Journal of Neurotherapy, 16*(3), 231-234. doi:10.1080/10874208.2012.705770

Koberda, J. L., Moses, A., Koberda, P., & Koberda, L. (2012a, September). *Comparison of the effectiveness of z-score surface/LORETA 19-electrodes neurofeedback to standard 1-electrode neurofeedback.* Oral Presentation at the 20th Annual Conference of the International Society for Neurofeedback and Research, Orlando, FL.

Koberda, J. L., Moses, A., Koberda, L., & Koberda, P. (2012b). Cognitive enhancement using 19-electrode z-score neurofeedback. *Journal of Neurotherapy, 16*(3), 224-230. doi:10.1080/10874208.2012.705770

Krigbaum, G., & Wigton, N. L. (2013). A proposed methodology of analysis for monitoring treatment progression with 19-channel z-score neurofeedback (19ZNF) in a single-subject design. Manuscript submitted for publication.

La Vaque, T. J., Hammond, D. C., Trudeau, D., Monastra, V., Perry, J., Lehrer, P., … Sherman, R. (2002). Template for developing guidelines for the evaluation of the clinical efficacy of psychophysiological interventions. *Applied Psychophysiology & Biofeedback, 27*(4), 273–281.

Lofthouse, N., Arnold, L. E., Hersch, S., Hurt, E., & DeBeus, R. (2012). A review of neurofeedback treatment for pediatric ADHD. *Journal of Attention Disorders, 16*, 351-372. doi:10.1177/1087054711427530

Loo, S. K., & Barkley, R. A. (2005). Clinical utility of EEG in attention deficit hyperactivity disorder**.** *Applied Neuropsychology, 12*(2), 64–76.

Loo, S. K., & Makeig, S. (2012). Clinical utility of EEG in attention deficit/ hyperactivity disorder: A research update. *Neurotherapeutics, 9*(3), pp 569-587. doi:10.1007/s13311-012-0131-z

Lubar, J. F., & Shouse, M. N. (1976). EEG and behavioral changes in a hyperactive child concurrent with training of the sensorimotor rhythm (SMR): A preliminary report. *Biofeedback & Self-Regulation, 1*, 293–306.

Lucido, M. L. (2012). *Effects of neurofeedback on neuropsychological functioning in an adult with autism* (Doctoral dissertation). Available from ProQuest Central database (UMI No. 3499929).

Machado, S., Portella, C. E., Silva, J. G., Velasques, B., Terra, P., Vorkapic, C. F., …Ribeiro, P. (2007). Changes in quantitative EEG absolute power during the task of catching an object in free fall. *Arquivos de neuro-psiquiatria, 65*(3-A):633-636.

Matousek, M. and Petersen, I. (1973). Automatic evaluation of background activity by means of age dependent EEG quotients. *EEG & Clinical Neurophysiology, 35*, 603–612.

Niv, S. (2013). Clinical efficacy and potential mechanisms of neurofeedback. *Personality and Individual Differences, 54*, 676-686.

Orgim, G. & Kestad, K. A. (2013). Effects of neurofeedback versus stimulant medication in attention-deficit/hyperactivity disorder: A randomized pilot study. *Journal of Child and Adolescent Psychopharmacology, 23*(7), 448-57. doi:10.1089/cap.2012.0090

Othmer, S., Othmer, S. F., & Kaiser, D. A. (1999). EEG biofeedback: An emerging model for its global efficacy. In J. R. Evans & A. Abarbanel (Eds.), *Introduction to quantitative EEG and neurofeedback* (pp. 243-310). San Diego, CA: Academic Press.

Othmer, S. F., & Othmer, S. (2007). Interhemispheric EEG training: Clinical experience and conceptual models. In J. R. Evans (Ed.), *Handbook of neurofeedback,* (pp.109-136). New York: Hawthorn Medical Press.

Pavlov, I. (1928). Natural science and the brain (W. H. Gantt, Trans.). In, *Lectures on conditioned reflexes: Twenty-five years of objective study of the higher nervous activity (behaviour) of animals* (pp. 120-130). New York: Liverwright Publishing. doi:10.1037/11081-010

Peniston, E. G., & Kullkosky, P. J. (1990). Alcoholic personality and alpha-theta brainwave training. *Medical Psychotherapy, 3*, 37-55.

Peniston, E. G., & Kullkosky, P. J. (1991). Alpha-theta brainwave neuro-feedback therapy for Vietnam veterans with combat-related post-traumatic-stress disorder. *Medical Psychotherapy, 4*, 47-60.

Peterson, C. A. (2001). Review of the Devereux Scales of Mental Disorders. In B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* (pp. 408–412). Lincoln, NE: Buros Institute of Mental Measurements.

Pigott, H. E., De Biase, L., Bodenhamer-Davis, E., & Davis, R. E. (2013). The evidence-base for neurofeedback as a reimbursable healthcare service to treat attention deficit/hyperactivity disorder. Retrieved from International Society for Neurofeedback and Research website: http://www.isnr.org/uploads/nfb-adhd.pdf

Pizzagalli, D. A. (2007). Electroencephalography and high-density electrophysiological source localization. In J. T. Cacioppo, L. G. Assinaru, G. G. Berntson (Eds.), *Handbook of psychophysiology,* (3rd ed.) (pp. 56-84), Cambridge, England: Cambridge University Press.

Protection of Human Subjects, 45, C.F.R. pt. 46 (2009).

Ramezani, A. (2008). *The effects of sequential versus referential montage neurofeedback amplitude training on QEEG Measures of phase and coherence.* (Doctoral dissertation). Available from ProQuest Central database (UMI No. 3352127).

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2(*1), 21-33.

Reichardt, C. (2009). Quasi-experimental design. In Roger E. Millsap, & A. Maydeu-Olivares (Eds.), *The Sage handbook of quantitative methods in psychology*. (pp. 46-72). London: Sage Publications Ltd.
doi: http://dx.doi.org.library.gcu.edu:2048/10.4135/9780857020994.n3

Robbins, J. (2000). *A Symphony in the brain.* New York, NY: Grove Press.

Rosenberg, M. S., Adams, D. C., & Gurevitch, J. (2000**).** *MetaWin: Statistical software for meta-analysis, version 2.0*. Sunderland, MA: Sinauer Associates.

Roth, R. M., Isquith, P. K., & Gioia, G. A. (2005). *BRIEF: Behavior Ration Inventory of Executive Function – Adult version: Professional manual*. Lutz, FL: PAR Inc.

Rutter, P. (2011, September). *Potential clinical applications for 19 channel live z-score training using Percent ZOK and ZPlus protocols.* Oral Presentation at the 19th

Annual Conference of the International Society for Neurofeedback and Research, Carefree, AZ.

Sanford, J. A., & Turner, A. (2009). *Integrated Visual and Auditory Continuous Performance Test Administration manual*. North Chesterfield, VA: BrainTrain Inc.

Sherlin, L. H., Arns, M., Lubar, J., Heinrich, H., Kerson, C., Strehl, U., & Sterman, M. B. (2011). Neurofeedback and basic learning theory: Implications for research and practice, *Journal of Neurotherapy, 15*(4), 292-304. doi: 10.1080/10874208.2011.623089

Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.

Smith, S. R., & Reddy, L. A. (2002). The concurrent validity of the Devereux Scales of Mental Disorders. *Journal of Psychoeducational Assessment, 20*, 112-127. doi: 10.1177/073428290202000201

Steiner, N. J., Sheldrick, R. C., Gotthelf, D., & Perrin, E. C. (2011). Computer-based attention training in the schools for children with attention deficit/hyperactivity disorder: a preliminary trial. *Clinical Pediatrics, 50*, 615-622.

Sterman, M. B., & Friar, L. (1972). Suppression of seizures in an epileptic following sensorimotor EEG feedback training. *Electroencephalography and Clinical Neurophysiology*, *33(*1), 89-95.

Sterman, M. B., LoPresti, R. W., & Fairchild, M. D. (2010). Electroencephalographic and behavioral studies of monomethyl hydrazine toxicity in the cat. *Journal of Neurotherapy, 14*(4), 293-300. doi:10.1080/10874208.2010.523367

Stoller, L. (2011). Z-Score training, combinatorics, and phase transitions, *Journal of Neurotherapy, 15*(1), 35-53. doi:10.1080/10874208.2010.545758

Sürmeli, T., & Ertem, A. (2007). EEG neurofeedback treatment of patients with Down syndrome, *Journal of Neurotherapy, 11*(1), 63-68. doi:10.1300/J184v11n01_07

Surmeli, T., & Ertem, A. (2009). QEEG guided neurofeedback therapy in personality disorders: 13 case studies. *Clinical EEG and Neuroscience*, *40*, 5-10.

Surmeli, T., & Ertem, A. (2010). Post WISC-R and TOVA improvement with QEEG guided neurofeedback training in mentally retarded: A clinical case series of behavioral problems. *Clinical EEG and Neuroscience, 41*(1), 32-41.

Surmeli, T., & Ertem, A. (2011). Obsessive compulsive disorder and the efficacy of qEEG-Guided neurofeedback treatment: A case series. *Clinical EEG and Neuroscience, 42*(3), 195-201.

Surmeli, T., Ertem, A., Eralp, E., & Kos, I. H. (2012). Schizophrenia and the efficacy of qEEG guided neurofeedback treatment: A clinical case series. *Clinical EEG and Neuroscience, 43*, 133-144. doi:10.1177/1550059411429531

Thatcher, R. W. (2012). *Handbook of quantitative electroencephalography and EEG biofeedback*. St. Petersburg, Florida: Anipublishing.

Thatcher, R. W. (2013). Latest developments in live z-score training: Symptom check list, phase reset, and LORETA z-score biofeedback. *Journal of Neurotherapy, 17*(1), 69-87. doi:10.1080/10874208.2013.759032

Thatcher, R. W., & Lubar, J. E. (2009). History of the scientific standards of QEEG normative databases. In T. H. Budzynski, H. K. Budzynski, J. R. Evans & A.

Abarbanel (Eds.), *Introduction to quantitative EEG and neurofeedback: Advanced theory and applications* (2nd ed., pp. 29-59). Burlington, MA: Elsevier.

Thatcher, R. W., North, D., & Biver, C. (2005). Evaluation and validity of a LORETA normative EEG database. *Clinical EEG and Neuroscience*, *36*(2), 116–122.

Thatcher, R. W., Walker, R. A., Biver, C., North, D., & Curtin, R. (2003). Quantitative EEG normative databases: Validation and clinical correlation. *Journal of Neurotherapy*, *7*(3-4**)**, 87–122. doi:10.1300/J184v07n03_05

Thompson, M., & Thompson, L. (2003). *The neurofeedback book*. Wheat Ridge, CO: Association of Applied Psychophysiology and Biofeedback.

Thorndike, E. L. (1911). *Animal intelligence*. New York: Macmillan.

Walker, J. E. (2009). Anxiety associated with post-traumatic stress disorder: The role of quantitative electroencephalograph in diagnosis and in guiding neurofeedback training to remediate the anxiety. *Biofeedback, 37*, 67-70.

Walker, J. E. (2010a). Recent advances in quantitative EEG as an aid to diagnosis and as a guide to neurofeedback training for cortical hypofunctions, hyperfunctions, disconnections, and hyperconnections: Improving efficacy in complicated neurological and psychological disorders. *Applied Psychophysiology and Biofeedback, 35*, 25–27. doi:10.1007/s10484-009-9107-0

Walker, J. E. (2010b). Using QEEG-guided neurofeedback for epilepsy versus standardized protocols: Enhanced effectiveness? *Applied Psychophysiology and Biofeedback, 35*, 29–30. doi:10.1007/s10484-009-9123-0

Walker, J. E. (2011). QEEG-guided neurofeedback for recurrent migraine headaches. *Clinical EEG and Neuroscience, 42*(1), 59-61.

Walker, J. E. (2012a). Remediation of enuresis using QEEG-guided neurofeedback training. *Biofeedback, 40*(3), 109-112. doi:10.5298/1081-5937-40.3.04

Walker, J. E. (2012b). QEEG-guided neurofeedback for remediation of dysgraphia. *Biofeedback 40*(3), 113–114. doi:10.5298/1081-5937-40.3.03

Walker, J. E. (2013). QEEG-guided neurofeedback for anger/anger control disorder, *Journal of Neurotherapy, 17*(1), 88-92. doi:10.1080/10874208.2012.705767

Walker, J. E., Norman, C. A., & Weber, R. K. (2002). Impact of qEEG-guided coherence training for patients with a mild closed head injury, *Journal of Neurotherapy, 6*(2), 31-43. doi:10.1300/J184v06n02_05

Wigton, N. L. (2008, September). *4-channel z-score neurofeedback – A single case study*, Poster presented at the 16th Annual Conference of the International Society for Neurofeedback and Research, San Antonio, TX.

Wigton, N. L. (2009). First impressions of Neuroguide real-time z-score training. In J. Demos (Ed.), *Getting started with dynamic z-score training,* (pp. 81-89). Westminster, VT: Neurofeedback of S.VT.

Wigton, N. L. (2010a, September). *Laplacian z-score neurofeedback: A unique option in the realm of multi-channel z-score neurofeedback*. Plenary Session Oral Presentation at the 18th Annual Conference of the International Society for Neurofeedback and Research, Denver, CO.

Wigton, N. L. (2010b, September). *Case studies overview of multi-channel z-score neurofeedback.* Poster presented at the 18th annual International Society for Neurofeedback and Research Conference, Denver, CO.

Wigton, N. L. (2013). Clinical perspectives of 19-channel z-score neurofeedback: Benefits and limitations. *Journal of Neurotherapy, 17*(4), 259-264. doi: 10.1080/10874208.2013.847142

Wigton, N. L., & Krigbaum, G. (2012, September). *Insights gained from over 3 years of 19-channel z-score neurofeedback: Towards new paradigms*. Poster presented at the 20th annual International Society for Neurofeedback & Research Conference, Orlando, FL.

Wyricka, W., & Sterman, M. B. (1968). Instrumental conditioning of sensorimotor cortex EEG spindles in the waking cat. *Physiology and Behavior, 3*(5), 703-707.

Appendix A

Test Distribution Limitations

Copies of the commercially available BRIEF and DSMD psychometric test instruments cannot be provided due to copyright protections. The publisher of the BRIEF, PAR Incorporated, states on the *Permissions* page of their website (www4.parinc.com/ProRes/permissions.aspx) that permission to include copies of an entire test will not be granted, for any publication, to include dissertations. The publisher of the DSMD, Pearson Education Incorporated, states in its *Terms and Conditions* page of their website (www.pearsonclinical.com/psychology/legal/termsofsale.html) that reproducing test items/scales is strictly prohibited by law as well as the terms and conditions for their products. The IVA is a computerized performance test, and as such, is only accessible by running the program on a computer. Therefore, a printed copy of this test is not available for inclusion in an appendix.

Appendix B

IRB Letter: Determination of Exempt Status

GRAND CANYON UNIVERSITY™

3300 West Camelback Road, Phoenix Arizona 85017 602.639.7500 Toll Free 800.800.9776 www.gcu.edu

DATE:	February 6, 2014
TO:	Nancy Wigton, M.A.
FROM:	Grand Canyon University Institutional Review Board
STUDY TITLE:	[561491-1] Evaluating 19-Channel Z-score Neurofeedback: Addressing Efficacy in a Clinical Setting
IRB REFERENCE #:	
SUBMISSION TYPE:	New Project
ACTION:	DETERMINATION OF EXEMPT STATUS
DECISION DATE:	
REVIEW CATEGORY:	Exemption category # *7.4*

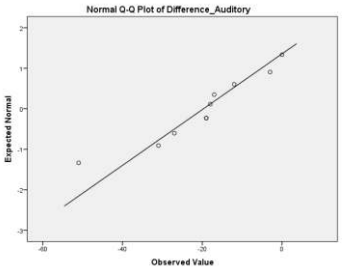
Thank you for your submission of New Project materials for this research study. Grand Canyon University Institutional Review Board has determined this project is EXEMPT FROM IRB REVIEW according to federal regulations.

Appendix C

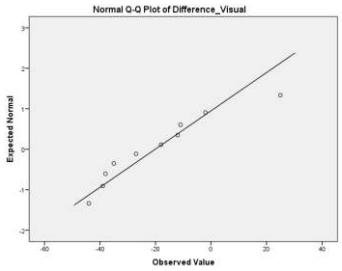
Q-Q Plots of Difference Scores

IVA Group

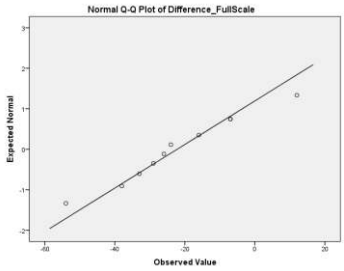
Auditory Scale



Visual Scale

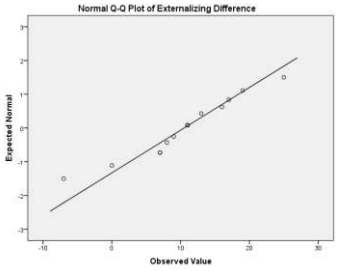


Full Scale

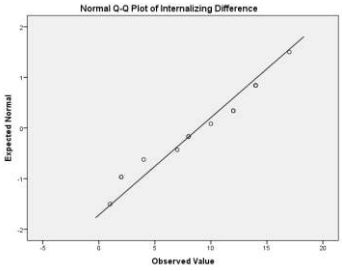


DSMD Group

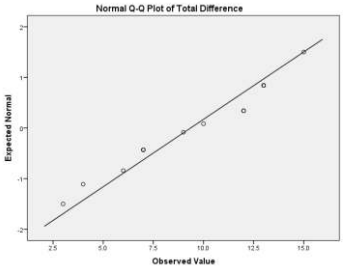
Externalizing Scale



Internalizing Scale

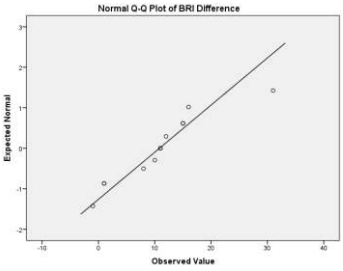


Total Scale

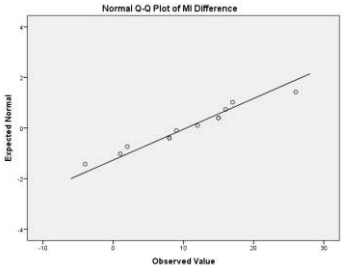


BRIEF Group

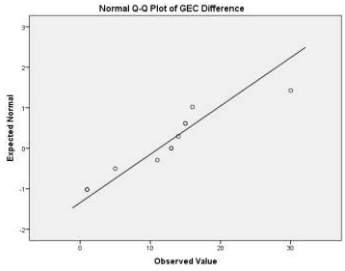
BRI Scale



MI Scale

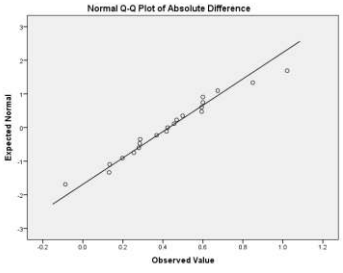


GEC Scale

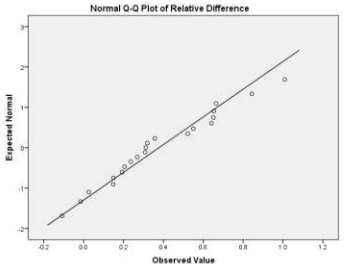


QEEG Group

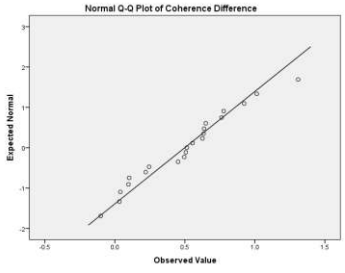
Absolute Power



Relative Power



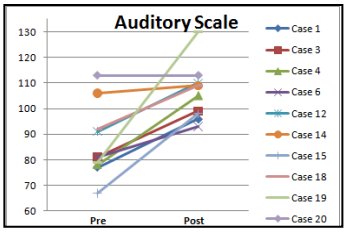
Coherence

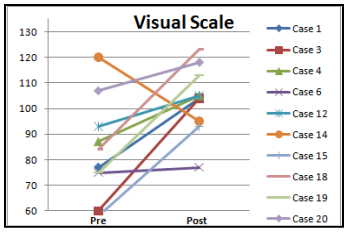


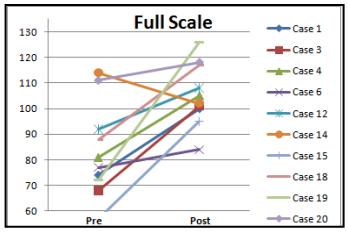
Appendix D

Line Graphs of Individual Pre-Post Scores

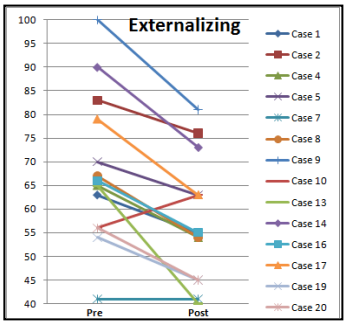
IVA Group

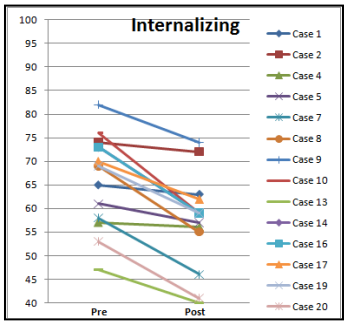


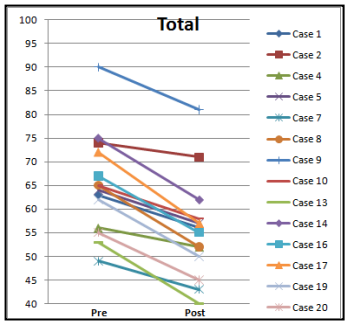




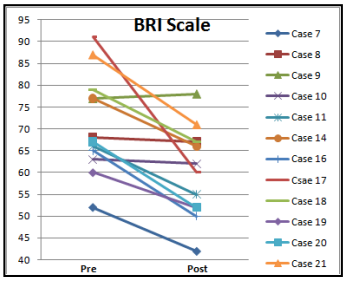
DSMD Group

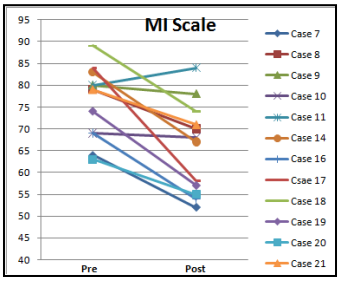


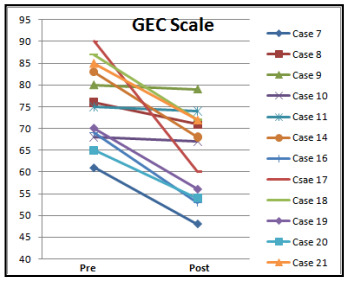




BRIEF Group







QEEG Group