NeuroRegulation

The Official Journal of

ISNR
INTERNATIONAL SOCIETY FOR NEUROFEEDBACK & RESEARCH

Volume 2, Number 3, 2015

NeuroRegulation

Volume 2, Number 3

2015

Contents

NeuroRegulation

Editorial – Volume 2, Number 3

Citation: Cannon, R. L. (2015). Editorial – Volume 2, Number 3. *NeuroRegulation*, *2*(3), 110. http://dx.doi.org/10.15540/nr.2.3.110

***Address correspondence to**: Rex L. Cannon, PhD, BCN, Neural Potential, 683 Edgebrook Ln., West Palm Beach, FL 33411, USA. Email: rexcannon@neuralpotential.com

Welcome to *NeuroRegulation* Volume 2, Number 3. In this issue we have several research papers and a review of quantitative electroencephalogram (qEEG) guided neurofeedback. Dr. Svetlana Malyutina and Dr. Dirk-Bart den Ouden examine the effects of transcranial direct current stimulation (tCDS) in aphasic disturbances and conclude this modality may prove useful in language dysfunctions. Dr. Jon Frederick, Kelli Dunn, and Dr. Thomas Collura provide pilot data on the relationship between EEG state discrimination and neurofeedback procedures and potential corresponding mechanisms inherent to these two processes. Stephanie Dreis, Angela Gouger, Edward Perez, Michael Russo, Michael Fitzsimmons, and Dr. Mark Jones provide pilot data examining the impact of using individualized qEEG-guided protocols to reduce symptoms of anxiety. The authors discuss the findings and implications for further study. Finally, Dr. Nancy Wigton and Dr. Genomary Krigbaum provide a review of qEEG-guided neurofeedback data. The authors examine trends and directions needed for future study. We thank all authors for their contribution to the scientific advancement of neurofeedback.

It is nearly time for our International Society for Neurofeedback and Research (ISNR) 23rd annual conference in Denver, Colorado. We would like to encourage all members and presenters to submit their research to *NeuroRegulation*. There are usually quite a large number of exceptional studies and theoretical constructs presented at our conference and we welcome all topics for *NeuroRegulation*. We would also like to encourage students to submit their research as well. We look forward to seeing you all in Denver.

Rex L. Cannon, PhD, BCN
Editor-in-Chief
Email: rexcannon@gmail.com

Published: October 8, 2015

NeuroRegulation

High-Definition tDCS of Noun and Verb Retrieval in Naming and Lexical Decision

Svetlana Malyutina[1,2*] *and Dirk-Bart den Ouden*[1,2]

[1]University of South Carolina, Department of Communication Sciences and Disorders, Columbia, South Carolina, USA
[2]National Research University Higher School of Economics (HSE), Moscow, Russia

Abstract

High-definition transcranial direct current stimulation (HD-tDCS) is a novel brain stimulation method that has high potential for use in language therapy for speakers with aphasia, due to its safety and focality. This study aimed to obtain foundational data on using HD-tDCS to modulate language processing in healthy speakers. Participants received stimulation either of Broca's area or of the left angular gyrus (20 min of anodal, cathodal, and sham stimulation on separate days), followed by naming and lexical decision tasks with single-word verb and noun stimuli. We found that cathodal stimulation over both Broca's area and the left angular gyrus increased naming speed for both verbs and nouns, challenging the traditional view of cathodal stimulation as suppressive or leading to decreased performance. The effect did not extend to the lexical decision task. Additionally, effects of specific stimulation types depended on the order of their administration, suggesting possible physiological carry-over and/or task novelty effects. These results are relevant to the application of HD-tDCS to enhance and direct neural plasticity in patients with neurogenic language disorders.

Keywords: HD-tDCS; brain stimulation; language; word retrieval; Broca's area; angular gyrus

Citation: Malyutina, S., & den Ouden, D-B. (2015). High-Definition tDCS of Noun and Verb Retrieval in Naming and Lexical Decision. *NeuroRegulation*, *2*(3), 111–125. http://dx.doi.org/10.15540/nr.2.3.111

***Address correspondence to**: Svetlana Malyutina, Staraya Basmannaya Street 21/4, Room 510,105066 Moscow, Russia. Email: smalyutina@hse.ru

Edited by:
Genomary Krigbaum, PsyD, Marian University, College of Osteopathic Medicine, Indiana, USA

Reviewed by:
Scott Decker, PhD, University of South Carolina, Department of Psychology, South Carolina, USA
Marvin W. Sams, ND, The Sams Center, Texas, USA

Introduction

It has been suggested that effects of behavioral speech-language therapy for neurogenic language disorders such as aphasia may be enhanced through brain stimulation (Holland & Crinion, 2012). Preliminary data show that several types of brain stimulation are capable of modulating language processing. Transcranial magnetic stimulation (TMS), a method that applies magnetic fields through a metal coil to induce electric current in focal brain areas and thus to cause neurons to fire action potentials, has been shown to modulate speech production and perception, as well as lexical, syntactic, and semantic processing, both in neurologically healthy subjects (Devlin & Watkins, 2007) and in persons with aphasia (PWA; Naeser et al., 2012). Similarly, transcranial direct current stimulation (tDCS), a method that delivers constant low current through electrodes on the scalp and, unlike TMS, is believed to modulate cortical excitability rather than directly cause neurons to fire (Stagg, 2014), has been shown to modulate verbal fluency, picture naming, grammar learning, and other language functions in healthy subjects (Cattaneo, Pisoni, & Papagno, 2011; de Vries et al., 2010; Fertonani, Rosini, Cotelli, Rossini, & Miniussi, 2010; Holland et al., 2011; Iyer et al., 2005; Sparing, Dafotakis, Meister, Thirugnanasambandam, & Fink, 2008; Wirth et al., 2011; for a review, see Price, McAdams, Grossman, & Hamilton, 2015) and PWA (Baker, Rorden, & Fridriksson, 2010; Fiori et al., 2011; Flöel et al., 2011; Fridriksson, Richardon, Baker, & Rorden, 2011; Vines, Norton, & Schlaug, 2011; You, Kim, Chun, Jung, & Park, 2011; for a review, see Monti et al., 2013).

Recently, a high-definition innovation to tDCS (HD-tDCS) has also begun to be applied towards the modulation of language processing (Price, Bonner,

Hamilton, Peelle, & Grossman, 2015). Like conventional tDCS, HD-tDCS is based on applying electric current in order to modulate neuronal excitability by acting on the resting membrane potential, affecting sodium and calcium channels, as well as NMDA receptors, and by possibly modulating synaptic activity (Stagg & Nitsche, 2011). However, an important advantage of HD-tDCS over conventional "sponge" tDCS is its increased focality. HD-tDCS is capable of inducing more intensive electric fields at smaller target locations while leaving others relatively unaffected (Datta et al., 2009), which may potentially be more effective than stimulating broader brain areas. Moreover, there is evidence of high safety and tolerability of HD-tDCS (Borckardt et al., 2012). These characteristics suggest a high potential for routine clinical use and warrant more research on how language processing can be modulated with HD-tDCS, both in the healthy population and ultimately in PWA and people with other language disorders. Similarly to traditional tDCS, HD-tDCS can be anodal (injecting positive charge into the target area and supposedly lowering the neural activation threshold through depolarization) or cathodal (injecting negative charge into the target area and supposedly increasing the neural activation threshold through hyperpolarization), although the dichotomy of anodal versus cathodal stimulation may be an oversimplification, especially for HD-tDCS (Garnett, Malyutina, Datta, & den Ouden, 2015).

The present study applied HD-tDCS in neurologically healthy participants and targeted a specific aspect of language processing: lexical retrieval, tested with an overt naming task and a lexical (word/non-word) decision task. Lexical retrieval is the crucial prerequisite to performing both tasks, although they also involve other processes that are not shared (e.g., naming but not lexical decision requires articulation, etc.). We tested two grammatical classes of words: verbs and nouns. In one group of participants, we aimed to nonspecifically modulate lexical retrieval of both verbs and nouns, by targeting Broca's area, traditionally associated with a very wide range of speech and language functions (Hagoort, 2005; Thompson-Schill, 2005). In the other group of participants, we aimed to specifically modulate verb processing by targeting posterior temporal/inferior parietal cortex.

Verbs have received special attention in many language treatments for aphasia (Bazzini et al., 2012; Thompson, Riley, den Ouden, Meltzer-Asscher, & Lukic, 2013) because they determine what other words appear in a sentence, in what semantic roles, and in what order and grammatical form (such information is referred to as verb argument structure). For example, it is the argument structure of the verb *to give* that determines that a sentence containing this verb should include three arguments (participants): *Anna gave a book to John*. Thus, improvement of verb processing may contribute to general improvement of sentence production and comprehension, and it is of our interest to investigate whether brain stimulation can be used to specifically enhance verb processing, which is more complex than noun processing at a variety of levels (for a review, see Vigliocco, Vinson, Druks, Barber, & Cappa, 2011). Enhancement is particularly relevant for verbs with more complex argument structure, such as a greater number of arguments, since those have been shown to present the greatest challenge both for healthy individuals (e.g., in grammatical class judgment in Rodríguez-Ferreiro, Andreu, & Sanz-Torrent, 2014, and lexical decision, Rodríguez-Ferreiro et al., 2014, and Shapiro, Brookins, Gordon, & Nagel, 1991; see, however, Thompson et al., 2007) and for PWA (e.g., in naming, Kim & Thompson, 2000; and sentence production, Kiss, 2000, and Thompson, Lange, Schneider, & Shapiro, 1997). Aiming to specifically modulate verb processing, we targeted the left angular gyrus, since previous neuroimaging research suggests that this area is involved in the processing of verb argument structure. For example, den Ouden, Fix, Parrish, and Thompson (2009) showed activation of the left angular gyrus for naming two-argument verbs (e.g., *to kiss, to stir*) compared to one-argument verbs (e.g., *to jump, to cough*), and Thompson, Bonakdarpour, and Fix (2010) showed activation of the left angular gyrus in a lexical decision task when comparing three-argument verbs (e.g., *to give, to invite*) to one-argument verbs.

Previous studies have already made first attempts to specifically modulate verb processing with brain stimulation. Cappa, Sandrini, Rossini, Sosta, and Miniussi (2012) applied rTMS in healthy participants and found that rTMS over the left (but not right) prefrontal cortex specifically resulted in faster naming of verbs, but not of nouns, presumably due to the role of this region in the processing of action semantics. However, Fertonani et al. (2010) failed to replicate this effect with tDCS in healthy participants: in their study, anodal (but not cathodal) stimulation over the left dorsolateral prefrontal cortex resulted in faster naming of both actions and objects. Marangolo et al. (2013) applied tDCS in PWA and investigated the potential of left frontal

versus temporal regions to modulate verb processing. They found that anodal tDCS over Broca's area improved verb-naming accuracy compared to anodal tDCS over Wernicke's area and sham; however, the study did not test noun naming and thus cannot exclude that the improvement was not specific to verbs. The present study aimed to add to this line of research by stimulating the left angular gyrus, a region that has received a lot of attention in neuroimaging research of verb processing in relation to the complexity of verb argument structure representations (see above) but has not yet been addressed by brain stimulation research, and by testing performance on both verbs and nouns in order to investigate the specificity of effects, which was not possible in Marangolo et al. (2013).

To summarize, we hypothesized that stimulation of the left angular gyrus would specifically affect verb processing, in particular for more linguistically complex verbs (verbs with a greater number of arguments), whereas stimulation of Broca's area would nonspecifically affect lexical retrieval of both verbs and nouns. We hypothesized that effects would occur at the level of retrieval of lexical items from the mental lexicon and thus be present in both naming and lexical decision tasks, rather than depend on whether the task involves overt speech/articulation processes (naming) or not (lexical decision). In the case of task-specific effects, it is likely that processes other than lexical retrieval are affected by the stimulation; that is, visual processing, object recognition, conceptual/semantic processing, phonological and articulatory planning, and speech motor execution in the naming task, versus reading, response selection, and hand-motor execution in the lexical decision task. Since HD-tDCS is a novel method, an additional purpose of the study was to obtain more data on its safety and tolerability.

Materials and Methods

Participants

Twenty-seven healthy volunteers participated in the study (14 females; mean age 22.1, *SD* 3.2, range 18–31 years; mean number of years of formal education 15.6, *SD* 2.8, range 12–24). All participants were right-handed and native speakers of American English. All participants had normal or corrected-to-normal vision and hearing and no reported history of neurological, psychiatric, speech, or language disorders. The study was approved by the University of South Carolina Institutional Review Board. Informed consent was obtained from all individual participants included in the study.

Design

The stimulation target was left-hemisphere Broca's area in 14 participants (7 females; mean age 22.9, *SD* 4.0, range 19–31 years; mean number of years of formal education 15.4, *SD* 2.9, range 12–23; 12 white Caucasians, 1 white Hispanic, 1 Native Hawaiian) and the left angular gyrus in 13 participants (one dropped out of the study due to adverse effects after first session; thus, data of 13 participants were analyzed: 7 females; mean age 21.2, *SD* 1.8, range 18–29; mean number of years of formal education 15.8, *SD* 2.9, range 13–24; 12 white Caucasians, 1 Asian). Participants were randomly assigned to one of the two groups. Each participant received anodal, cathodal, and sham stimulation, at similar times of day, with a minimum interval between stimulation sessions of 24 hr (mean number of days 2.8, *SD* 3.0, range 1–10). The order of stimulation sessions was maximally counterbalanced across participants. Participants were blinded to their stimulation target and the order of stimulation types. All participants whose data are included into analysis completed all three sessions.

Procedure

Before stimulation, participants were familiarized with the experimental tasks: they received instructions and examples and completed a short practice session of each task (lexical decision, action naming, and object naming in the order in which they were to be administered after stimulation in the participant; practice stimuli did not overlap with experimental stimuli). Feedback was given during practice. Participants were then administered 20 min of stimulation, during which they performed a non-language distraction task (silently working on a jigsaw puzzle). Participants were asked to verbally rate pain and unpleasantness separately at 0.5, 10, and 19.5 min after stimulation onset on a scale from 1 (no pain/no unpleasantness) to 10 (very strong pain/very strong unpleasantness). Two experimental language tasks were performed immediately after the stimulation; the order of tasks was maximally counterbalanced across participants.

Stimulation setup

HD-tDCS was administered through an MxN HD-tDCS stimulator (Soterix Medical, New York, NY) with Ag/AgCl HD electrodes (Minhas et al., 2010) placed into an MCN-system (10–10) EEG cap (size 56 or 58, EASYCAP, Herrsching, Germany) and stabilized with HD electrode holders (Soterix Medical, New York, NY) filled with conducting gel

(SignaGel, Parker Laboratories, Fairfield, NJ). A small amount of benzocaine gel (Lanacane brand) was applied to the skin to reduce any uncomfortable sensations. Lanacane has been routinely used to reduce discomfort during HD-tDCS (e.g., Guleyupoglu et al., 2014). Although it cannot be completely ruled out that any systemic effects caused by the mechanism of action of Lanacane can potentially interfere with effects of HD-tDCS (Scholz, 2002), the dosing in the present study was sufficiently low to consider this possibility as unlikely.

Electrode montages were chosen using HD-Targets™ and HD-Explore™ software (Soterix Medical, New York, NY) and are presented in Table 1 and Figure 1. For the left Broca's area stimulation, we targeted MNI coordinates (-49, 16, 24), which correspond to average peak coordinates in a meta-analysis of sentence processing (Vigneau et al., 2006). For the left angular gyrus stimulation, we targeted MNI coordinates (-27, -57, 51), corresponding to the peak activation associated with complex verb processing in (den Ouden et al., 2009).

Table 1
Electrode configurations.

Stimulation target site	Stimulation type	Electrode configuration (MCN system)	Resulting intensity at target coordinates in the Broca's area	Resulting intensity at target coordinates in the angular gyrus
Broca's area	Anodal	FC5 (+2 mA), AFz (-2 mA)	0.61 V/m	0.16 V/m
	Cathodal	FC5 (-2 mA), AFz (+2 mA)	0.61 V/m	0.16 V/m
	Sham	AFz (+1 mA), FPz (-1 mA), FC5 (+1 mA), FC3 (-1 mA)	0.11 V/m	0.01 V/m
Angular gyrus	Anodal	CP5 (+2.0 mA), POz (-2.0 mA)	0.19 V/m	0.52 V/m
	Cathodal	CP5 (-2.0 mA), POz (+2.0 mA)	0.19 V/m	0.52 V/m
	Sham	CP5 (+1 mA), TP7 (-1 mA)	0.05 V/m	0.06 V/m

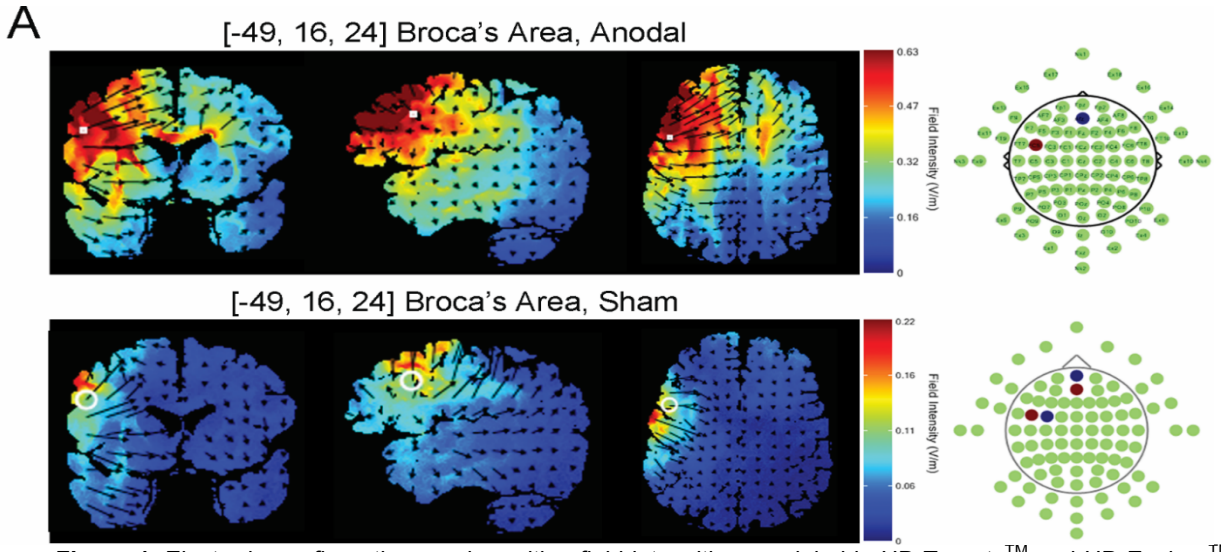

Figure 1. Electrode configurations and resulting field intensities, modeled in HD-Targets™ and HD-Explore™ software. Configurations for cathodal stimulation (not shown) were the same as for anodal stimulation but with reversed polarities of electrodes.

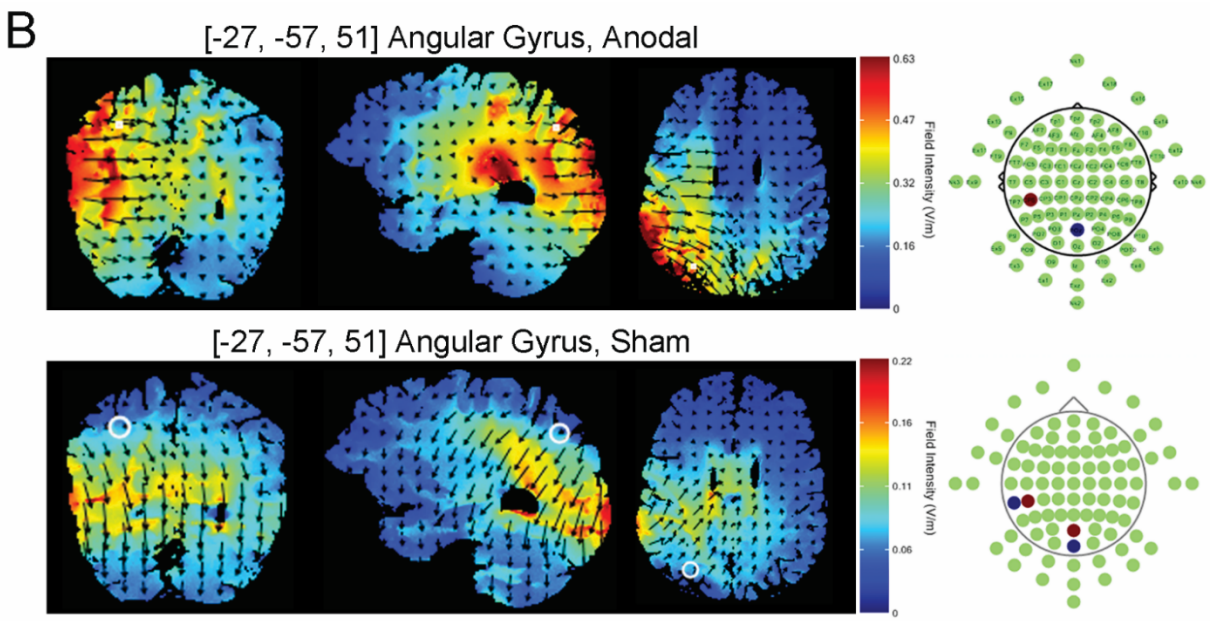

Figure 1. Electrode configurations and resulting field intensities, modeled in HD-Targets™ and HD-Explore™ software. Configurations for cathodal stimulation (not shown) were the same as for anodal stimulation but with reversed polarities of electrodes.

For sham, stimulation was also administered for the entire 20 min but in a montage where the current was modeled to bypass the cortex (Davis, Gold, Pascual-Leone, & Bracewell, 2013; Garnett & den Ouden, 2015; Kessler, Turkeltaub, Benson, & Hamilton, 2012; Richardson et al., 2014). Electrodes were placed in proximal pairs so that the current was flowing in and out at adjacent electrodes. To better disguise sham by having equal numbers of electrodes on the participant's scalp across stimulation types, anodal and cathodal setups included two additional electrodes that did not administer any current.

Tasks and Stimuli

Each participant performed a naming task and a lexical decision task. Order of task administration was counterbalanced across participants. E-Prime software (Psychology Software Tools, Sharpsburg, PA) was used for stimuli presentation and data recording.

Naming task. Participants were shown black-and-white line drawings on a computer screen and were instructed to name them overtly in single words. Drawings were presented for 3 s with an interstimulus fixation cross (1.5 s). Stimuli included 60 pictures of objects to elicit nouns (e.g., *shirt, harp*) and 60 pictures of actions to elicit verbs, which differed on their argument structure: 30 among them had one argument (i.e., referred to actions not requiring an object, e.g., *to laugh, to pray*), and 30 had two arguments (i.e., referred to actions requiring an object, e.g., *to chase, to grill*).

Pictures were taken from a Russian database of action pictures (Akinina et al., 2015) and their English names were normed in a preliminary survey, completed by 61 participants without reported history of neurological, psychiatric, speech, or language disorders (43 females; mean age 33.5, *SD* 12.8, range 17–66 years). Only items with name agreement greater than 70% were included in the experimental materials (mean 91.6%, *SD* 8.0%, range 70–100%).

Nouns, one-argument verbs, and two-argument verbs were matched on their length in phonemes and syllables, name agreement, as well as lexical frequency and familiarity ratings obtained from the MRC Psycholinguistic Database (Coltheart, 1981). One-argument and two-argument verbs were also matched on the number of objects present in the picture (to ensure that any differences between the two conditions were linguistic/representational, rather than perceptual in nature) and on the percentage of their verb use (to account for any effects of verb-noun homonymy, such as *to hammer – a hammer*), estimated by manually counting types of use in 100 random contexts from the Corpus of Contemporary American English (http://corpus.byu.edu/coca/).

Stimuli were split, to be used in three sessions (20 nouns, 10 one-argument verbs, and 10 two-argument verbs in each), balanced on the same parameters. Noun and verb stimuli were presented separately, to ensure clarity of whether an action or an object needed to be named. The order of noun/verb naming was counterbalanced and the order of items within verbs/nouns was randomized. In each session, the naming task took approximately three minutes to complete.

Lexical decision task. Participants were presented with strings of letters on a computer screen and instructed to press one button for real English words and another button for non-words. Strings of letters were presented for 1.2 s with an interstimulus fixation cross (0.5 s). Stimuli included 90 nouns (e.g., *price, word*), 90 verbs (verbs differed on their argument structure: 30 had one argument, e.g., *swear, wait*; 30 had two arguments, e.g., *produce, conduct*; and 30 had three arguments, that is, referred to an action requiring a direct and an indirect object, e.g., *send, provide*) and 360 pronounceable non-words (e.g., *mipe, assect*). Non-words were constructed by recombining pronounceable segments of real words.

Nouns and the three verb types were matched on lexical frequency and imageability obtained from the MRC Psycholinguistic Database (Coltheart, 1981), orthographic neighborhood (Medler & Binder, 2005) and length in letters and syllables. Words and non-words were matched on length in letters and syllables and orthographic neighborhood. The three verb types were matched on the percentage of their verb use (to account for the effects of verb-noun homonymy).

Stimuli were split, to be used in three sessions (30 nouns, 10 verbs of each type, and 120 non-words in each), balanced on the same parameters. The order of stimuli was randomized within each session. In each session, the lexical decision task took approximately five minutes to complete.

Data Analysis

Statistical analysis was performed in the SPSS 22.0 software (SPSS Inc., Armonk, NY). Pain and unpleasantness ratings were analyzed in two repeated-measures ANOVAs with Stimulation Site as a between-subjects variable and Timepoint and Stimulation Type as within-subject variables.

For the naming task, the analyzed outcomes were the percentage of accurate responses (responses were scored as accurate if given for a picture by at least two respondents in preliminary surveys), number of self-corrections and average reaction time (RT; i.e., time until response onset, based on visual analysis of the speech signal in Praat software [Boersma, 2001]). For the lexical decision task, the outcome measures were the percentage of accurate responses and average RT. Since accuracy and reaction times are typically skewed, they were log-transformed to approach the normal distribution, as is widely conventional for both of these outcome measures (Bartlett, 1947; Baayen & Milin, 2010; Hoyle, 1973).

For each outcome measure, analysis was performed over average values per condition per participant (i.e., data were aggregated across individual trials). We used the "linear mixed-effects" (MIXED) procedure in SPSS 22.0 with Subject as a random factor, one between-subject variable: Stimulation Site (Broca's area, the left angular gyrus), and three within-subject variables: Linguistic Condition (noun, one-argument verb, two-argument verb, and three-argument verb, where present), Stimulation Type (anodal, cathodal, sham) and Session Number (1, 2, 3), with subsequent pairwise comparisons for significant effects. Session number was included into the model in order to statistically account for any practice effects (such as participants getting better at the task across sessions due to experience) and to allow us to include the maximum number of participants whose data were collected, rather than to have to limit the analysis to exactly equal sample sizes assigned to each order of session administration (since the numbers of participants completing stimulation in each order were maximally counterbalanced but not exactly the same).

Since the above analysis consistently revealed interactions between Session Number and Stimulation Type (see section 3.2 below), we concluded that there may exist carry-over effects between sessions. Therefore, we conducted a post hoc and complementary exploratory analysis limited to the data from each participant's first session only, which cannot be subject to any carry-over effects. Due to the smaller number of observations in this exploratory analysis, we performed separate Kruskal-Wallis tests for each linguistic condition with Stimulation Type as a between-subjects factor, followed-up by Bonferroni-corrected Mann-Whitney U tests for pairwise comparisons.

Results

Safety, Tolerability and Sham Masking

Mean pain and unpleasantness ratings are presented in Table 2. Repeated-measures ANOVAs revealed that pain, $F(2,24) = 18.26$, $p < .001$, and unpleasantness, $F(2,25) = 37.23$, $p < .001$, subsided within session, with significant reduction both from first to second and from second to third time points. Neither pain nor unpleasantness was significantly affected by stimulation target or type.

Table 2
Pain and unpleasantness ratings, mean (SD).

		Pain			Unpleasantness		
Stimulation target	Stimulation type	0.5 min after start	10 min after start	19.5 min after start	0.5 min after start	10 min after start	19.5 min after start
Broca's area	Anodal	2.54 (1.57)	1.00 (0.00)	1.00 (0.00)	2.61 (1.42)	1.07 (0.27)	1.00 (0.00)
	Cathodal	2.04 (1.15)	1.07 (0.27)	1.00 (0.00)	2.68 (1.88)	1.29 (0.83)	1.14 (0.36)
	Sham	2.93 (1.73)	1.14 (0.36)	1.00 (0.00)	2.93 (1.69)	1.21 (0.43)	1.07 (0.27)
Angular gyrus	Anodal	2.27 (0.99)	1.31 (0.48)	1.00 (0.00)	2.38 (1.06)	1.38 (0.51)	1.08 (0.28)
	Cathodal	2.35 (1.31)	1.23 (0.44)	1.00 (0.00)	2.65 (1.25)	1.31 (0.48)	1.08 (0.28)
	Sham	2.35 (1.25)	1.15 (0.38)	1.00 (0.00)	2.42 (1.08)	1.42 (0.49)	1.08 (0.28)

Eight out of 27 participants correctly guessed which of the three sessions was sham, with chance being 9/27. One participant reported possible side effects (headache, fatigue and nausea) that started several hours after cathodal stimulation over Broca's area (participant's first session) and lasted several hours, although there was no direct evidence of these symptoms being related to the stimulation.

Task Results

Naming task. Accuracy of naming was at ceiling (mean 96.4%, *SD* 4.6%, range 72.5–100.0%). The analysis did not capture significant effects of any factors on accuracy. The number of self-corrections was low (mean 1.1%, *SD* 2.2%, range 0.0–12.5%), with no significant effects of any factors either.

Mean RT was 1103 ms (*SD* 117 ms, range 921–1312 ms). The analysis revealed main effects of Stimulation Site, $F(1, 205.42) = 23.86$, $p < .001$, with participants in the Broca's area group having slower reaction times than participants in the angular gyrus group; Linguistic Condition, $F(2, 134.57) = 85.11$, $p < .001$, with two-argument verbs having slower reaction times than one-argument verbs, $p < .001$, and those in turn having slower reaction times than nouns, $p < .001$ (see Figure 2); and Stimulation Type, $F(2, 137.95) = 6.59$, $p = .002$, with anodal and sham stimulation yielding slower reaction times than cathodal stimulation, $p = .001$ and $p = .008$ respectively (see Figure 3). The analysis also revealed the following interactions: a two-way Stimulation Type by Session Number interaction, $F(4, 136.87) = 2.88$, $p = .025$, and a three-way Stimulation Site by Stimulation Type by Session Number interaction, $F(6, 67.82) = 3.58$, $p = .004$, indicating that the effects of stimulation types were modulated by the order in which they were administered, and that this modulation varied between stimulation sites (see Figure 4a-b); no other interactions were significant.

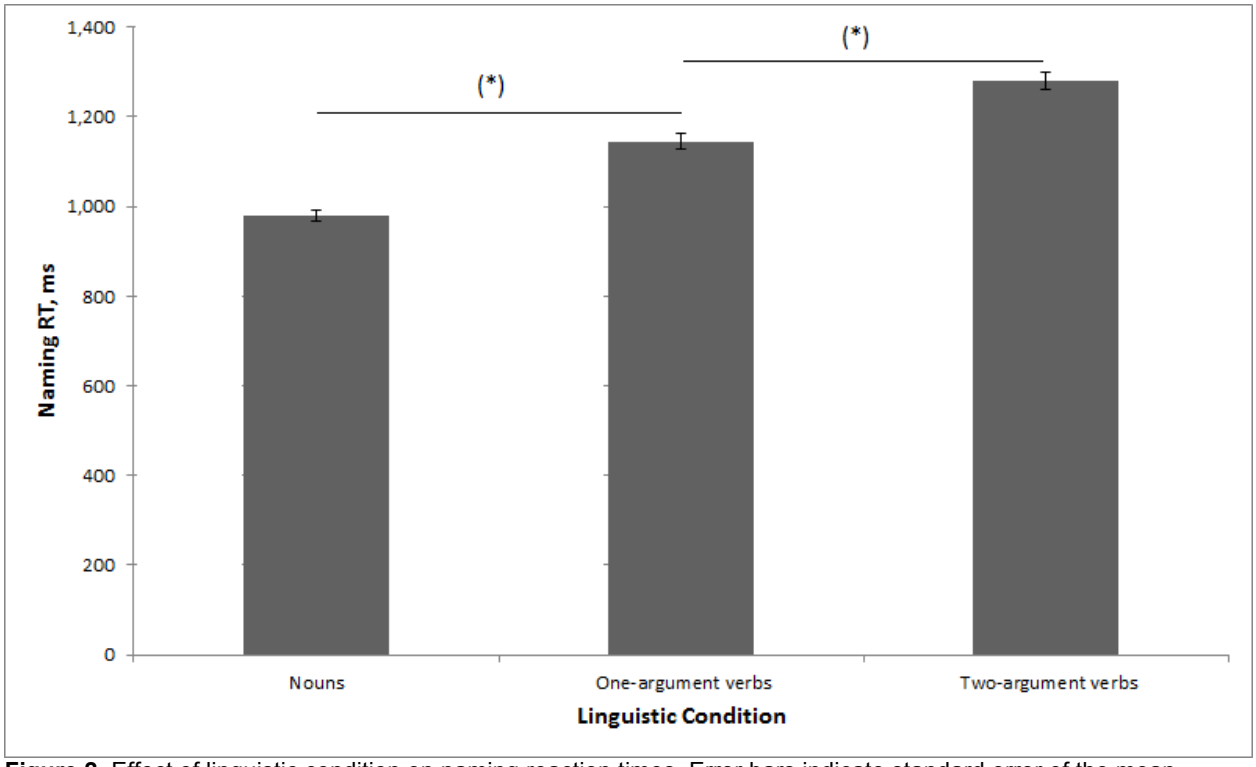

Figure 2. Effect of linguistic condition on naming reaction times. Error bars indicate standard error of the mean.
* indicates significant contrasts ($p < .05$).

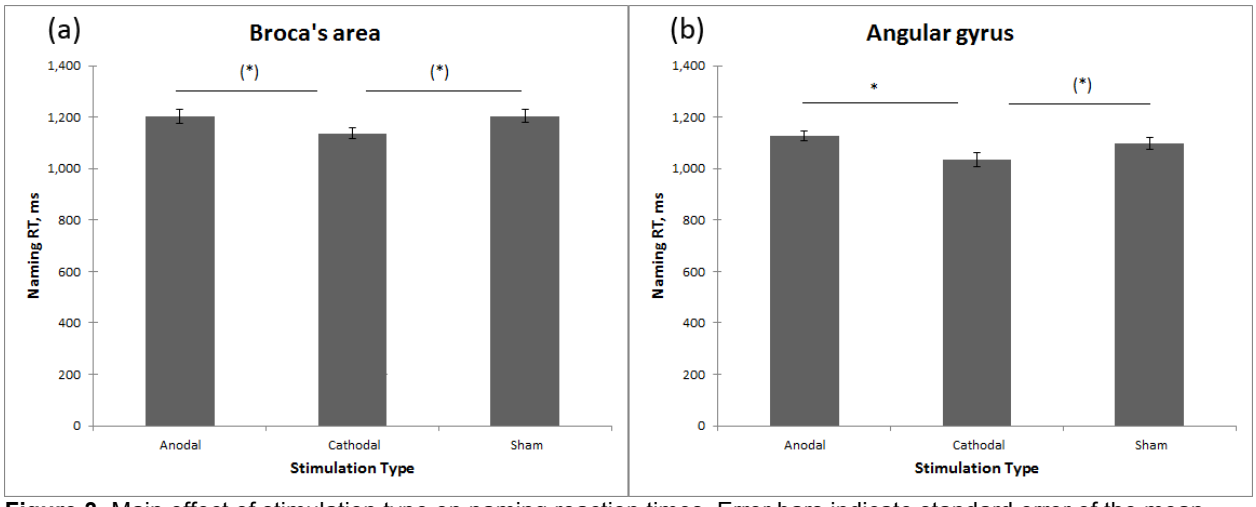

Figure 3. Main effect of stimulation type on naming reaction times. Error bars indicate standard error of the mean.
* indicates significant contrasts ($p < .05$), (*) indicates statistical trends ($.1 < p < .05$)

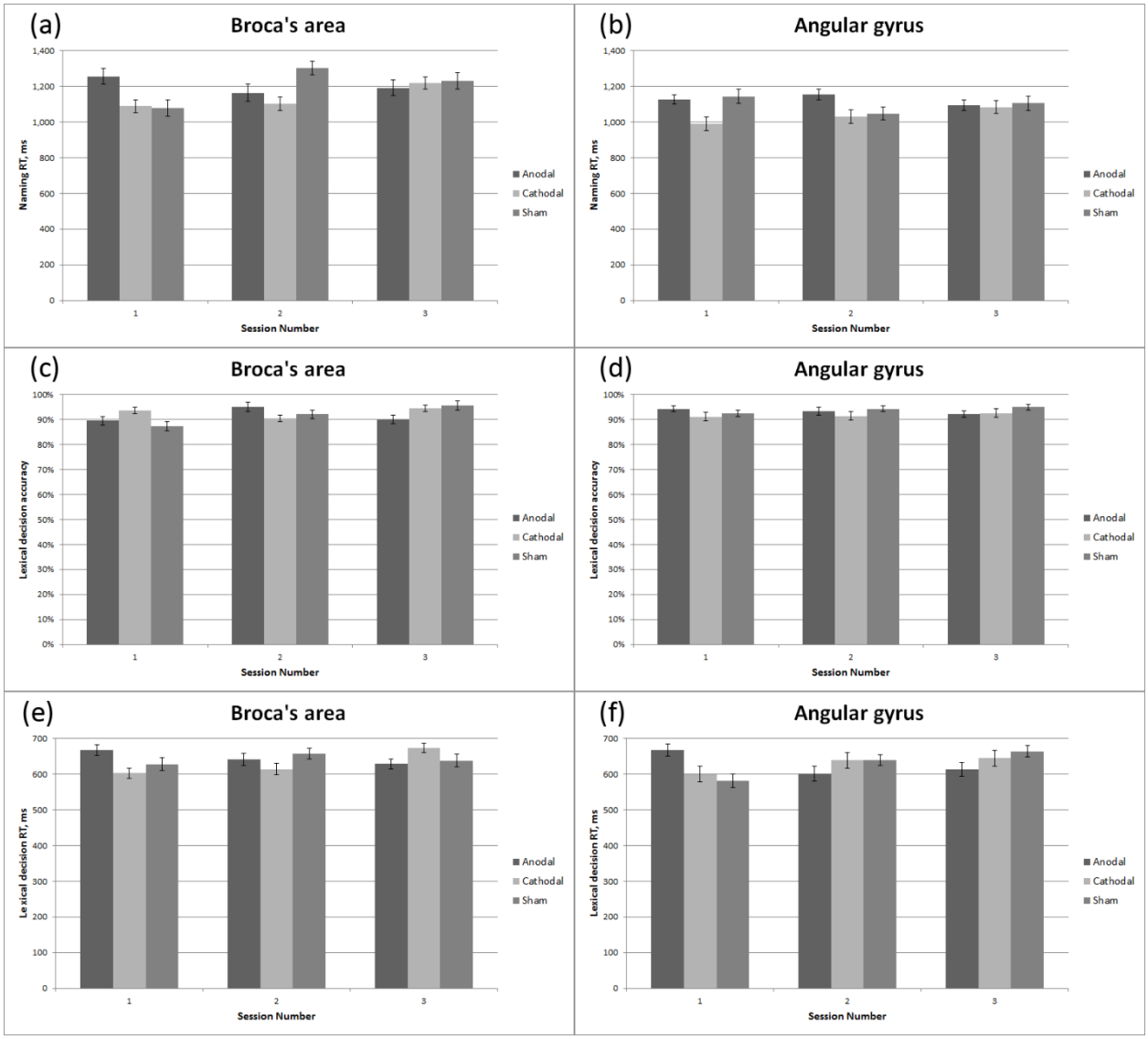

Figure 4. Naming reaction times (a, b), lexical decision reaction accuracy (c, d) and lexical decision reaction times (e, f) across stimulation sessions. Error bars indicate standard error of the mean.

To follow up on the significant three-way Stimulation Site by Stimulation Type by Session Number interaction, we "sliced" the interaction by Stimulation Site and performed separate analyses of naming reaction times in the Broca's area group and the angular gyrus group. In the Broca's area group, we found a main effect of Linguistic Condition, $F(2, 68.72) = 44.88$, $p < .001$, with two-argument verbs having slower reaction times than one-argument verbs, $p = .019$, and those in turn having slower reaction times than nouns, $p < .001$; a trend for an effect of Session Number, $F(2, 67.84) = 2.49$, $p = .088$, with reaction times becoming slower across sessions; a trend for an effect of Stimulation Type, $F(2, 71.74) = 2.52$, $p = .087$, with anodal and sham stimulation having slower reaction times than cathodal stimulation, $p = .061$ and $p = .059$ respectively (see Figure 3a); and a two-way Stimulation Type by Session Number interaction, $F(4, 71.35) = 5.63$, $p = .001$, indicating that the effects of stimulation type were modulated by the order in which they were administered (see Figure 4a); no other factors or interactions were significant. In the angular gyrus group, we found a main effect of Linguistic Condition, $F(2, 52.55) = 46.36$, $p < .001$, with two-argument verbs having slower reaction times than one-argument verbs, $p < .001$, and those in turn having slower reaction times than nouns, $p < .001$; and a main effect of Stimulation Type, $F(2, 63.34) = 4.80$, $p = .011$, with anodal stimulation and,

at the level of a statistical trend, sham stimulation having slower reaction times than cathodal stimulation, p = .003 and p = .058 respectively (see Figure 3b); no other factors or interactions were significant. For lack of statistical power, we did not follow up any further on the interactions between Stimulation Type and Session Number within the two stimulation sites, but Figure 4a-b provide a visual illustration of these effects, which were not uniform between the two sites.

In order to assess the effects of stimulation without interference from potential carry-over effects, we conducted a complementary analysis, limited to the participants' first sessions and separately for the two stimulation sites. For these data, a Kruskal-Wallis test did not reveal an effect of Stimulation Type on the accuracy or number of self-corrections in nouns or verbs in either the Broca's area or the angular gyrus group. For mean RTs in the Broca's area group (see first group of bars in Figure 1a), a Kruskal-Wallis test revealed trends towards an effect of Stimulation Type for both nouns, $\chi^2(2) = 5.23$, p = .073 (driven by anodal stimulation having slower RTs than cathodal stimulation, p = .026), and verbs, $\chi^2(2) = 5.57$, p = .062 (driven by anodal stimulation having slower RTs than sham, p = .033). For mean RTs in the angular gyrus group (see first group of bars in Figure 1b), there was a trend towards an effect of Stimulation Type for nouns, $\chi^2(2) = 5.09$, p = .078 (driven by cathodal stimulation having faster RTs than anodal stimulation, p = .026); the speed-up of RTs for cathodal stimulation compared to sham did not reach significance, p = .134. There was no effect of Stimulation Type for verbs, $\chi^2(2) = 1.53$, p = .465.

Lexical decision task. One participant was excluded from analysis due to non-compliance with the task. After excluding him, the accuracy on the lexical decision task was 93.4% (*SD* 4.7%, range 72.8–99.4%). The analysis revealed a main effect of Linguistic Condition, $F(4, 128.23) = 2.71$, p = .033, driven by nouns having higher accuracy than one-argument, p = .040, two-argument, p = .090, and three-argument, p = .009, verbs and by non-words having higher accuracy than three argument verbs, p = .045; and a two-way Stimulation Type by Session Number interaction, $F(4, 144.05) = 3.99$, p = .004, indicating that the effects of stimulation type were modulated by the order in which they were administered (see Figure 4c-d); no other factors or interactions were significant.

Mean reaction time was 647 ms (*SD* 81 ms, range 491–826 ms). The analysis revealed a two-way Stimulation Type by Session Number interaction, $F(4, 264.85) = 6.93$, $p < .001$, indicating that the effects of stimulation type were modulated by the order in which they were administered (see Figure 4e-f); no other factors or interactions were significant.

For data from the first session only (see first groups of bars in Figures 4c-f), Kruskal-Wallis tests revealed no effects of Stimulation Type on accuracy or mean RTs in any linguistic condition in either the Broca's area or the angular gyrus group.

Discussion

The present study is one of the first to apply the novel HD-tDCS method to modulate language processing, and specifically the lexical retrieval of nouns and verbs of varied argument structure complexity. We administered anodal, cathodal, and sham stimulation, targeting Broca's area in one group of neurologically healthy participants and the left angular gyrus in the other group, followed by a naming and a lexical (word/non-word) decision task.

Consistent with the existing psycholinguistic literature, we found that linguistic characteristics of the stimuli affected participants' performance. In the naming task, participants were slower in naming two-argument verbs than one-argument verbs, which in turn were named more slowly than nouns; in the lexical decision task, participants were more accurate for nouns than for verbs, especially the most complex (three-argument) verbs. This adds to the evidence of higher complexity of verb than noun processing (Vigliocco et al., 2011) and of higher complexity of verbs with a greater number of arguments (Kim & Thompson, 2000; Kiss, 2000; Rodríguez-Ferreiro et al., 2014; Shapiro et al., 1991; Thompson et al., 1997). Importantly, these findings indicate that the experimental task was valid and drew on the processes of lexical retrieval in an expected way. From a psycholinguistic perspective, these results indeed confirm that the complexity effects are generated at the level of lexical retrieval, as that is the process that is shared between the two tasks.

However, contrary to our original hypotheses about the effects of stimulation on particular linguistic conditions, we did not find that the stimulation over the left angular gyrus specifically affected processing of verbs and/or showed a greater effect for verbs with more complex argument structure: the analysis revealed no significant interactions of stimulation type or site with linguistic condition.

Thus, the present study failed to demonstrate that activation of the left angular gyrus during verb argument structure processing in previous neuroimaging research is due to the area being *necessary* for verb processing, rather than being merely involved but not crucial for it, which is the limit of what functional neuroimaging research has the potential to show. Lack of verb-specific effects is in line with the results by Marangolo et al. (2013) and Fertonani et al. (2010), who also demonstrated effects of brain stimulation to extend across verbs and nouns. To the best of our knowledge, the rTMS experiment by Cappa et al. (2002) remains the only study that demonstrated any effects specific to verb processing in the context of brain stimulation.

Instead, the primary finding with regard to main effects of stimulation was that cathodal stimulation both over Broca's area and, at the level of a statistical trend in the pairwise comparison to sham, over the left angular gyrus made participants faster on the naming task across both nouns and verbs of varied argument structure complexity; no significant differences were found between anodal stimulation and sham. Quite similarly, in the exploratory analysis of the data subset limited to each participant's first session and thus free from any carry-over effects, anodal stimulation over Broca's area led to slower naming relative to both sham and cathodal stimulation (at the level of a statistical trend). Cathodal stimulation over the left angular gyrus increased the naming speed relative to anodal stimulation (at the level of a statistical trend; an increase in speed relative to sham was similar in size but did not reach significance). This challenges the conclusions of previous tDCS literature with regard to directions of behavioral effects of anodal and cathodal stimulation. With tDCS, anodal stimulation over language-related areas has been traditionally associated with increases in language performance in healthy individuals (Cattaneo et al, 2011; Fertonani et al., 2010; Iyer et al., 2005; Wirth et al., 2011), while cathodal stimulation over language-related areas has been associated with decreased performance (Liuzzi et al., 2010), contrary to our findings. Nonetheless, other effects (e.g., positive effect of cathodal stimulation or lack of effect of anodal stimulation, as in the present study) have been demonstrated before. For example, in a pattern identical to our findings, Filmer et al. (2013) demonstrated improved multitasking performance following cathodal tDCS over the left posterior lateral prefrontal cortex, whereas performance following its anodal stimulation did not differ significantly from sham. Likewise, Pirulli, Fertonani, and Miniussi (2014) demonstrated improved performance in a visual task following cathodal tDCS over the primary visual cortex. Such reports are less numerous but this may partially be due to reporting bias and lack of publishing of null results.

To account for the apparent discrepancy noted above, we point out that facilitatory or detrimental effects of anodal and cathodal stimulation (or lack thereof) may greatly depend on the specific task, stimulation target, electrode montage, current intensity and stimulation duration (Garnett et al., 2015). For example, Batsikadze, Moliadze, Paulus, Kuo, and Nitsche (2013) show that the direction of tDCS effects in the motor domain (excitatory versus inhibitory) may vary depending on current intensity and also suggest that effects of stimulation duration may be nonlinear for intensities greater than 1 mA, as in the present experiment. Pirulli et al. (2014) also show effects of duration and intensity on the behavioral effect of anodal versus cathodal polarity in the visual domain. Thus, it is worth noting that some of the previous reports of enhancing language performance with anodal tDCS or decreasing language performance with cathodal stimulation over language-related areas used either smaller current intensities (1 mA in Liuzzi et al., 2010, and de Vries et al., 2010; 1.5 mA in Wirth et al., 2011) or other stimulation durations (8 or 10 min in Fertonani et al., 2010; 30 min in Wirth et al., 2011) than in the present study. The discrepancy may also be due to specific tasks (word learning in Liuzzi et al., 2010, verbal fluency in Cattaneo et al., 2011, and Iyer et al., 2005), stimulation target (left motor cortex in Liuzzi et al., 2010) and electrode montages (such as positioning of the reference electrode), as well as, importantly, to a possible difference in effects of stimulation administered through conventional sponge tDCS and HD-tDCS, the latter having more focal targeting.

Another factor that may have affected the outcome of stimulation is the nature of the task administered during stimulation. Contrary to TMS, tDCS is not considered to induce an action potential, but rather to bring depolarization closer to or further away from the threshold of neural firing. That is, it is considered to "prime" neurons for activation, rather than to induce activation directly. For that reason, it might be expected that task-related involvement of neurons during their stimulation with tDCS may be less relevant than in the case of TMS, where actual neural firing patterns may be reinforced in association with a particular task. This notion, together with the fact that we preferred to keep our outcome-measure tasks "fresh" and thus to avoid ceiling effects or fatigue, led us to use a nonlinguistic

task during stimulation that was deliberately unrelated to our outcome measures (similarly to, e.g., Cattaneo et al., 2011). Nevertheless, recent advances suggest that tDCS may indeed preferentially modulate neural networks that are active during stimulation (Bikson & Rahman, 2013; Gill, Shah-Basak, & Hamilton, 2015). It is possible, therefore, that stimulation effects might be modified if we had used a task that was more closely related to our object of investigation, that is, lexical retrieval and production. In addition, the strength of the behavioral effects of HD-tDCS may have been affected by the time point at which outcome measures were tested, relative to stimulation offset. Recently, Hoy et al. (2013) found that effects of anodal tDCS over the left prefrontal cortex were greater at 40 min post-stimulation compared to 0 min, which was the testing point in the present study. Contrary to that report, however, other evidence suggests that effects may be the strongest immediately post-stimulation (e.g., Fujiyama et al., 2014). So, while the test-timing issue must be considered unresolved at this point, it should be noted that it may have been a factor affecting our behavioral data. These observations again suggest that a larger body of comparable basic research is needed in order to account for apparent inconsistencies in findings (Garnett et al., 2015; Horvath, Forte, & Carter, 2015).

The effects of anodal and cathodal stimulation on naming reaction times were the same for stimulation over Broca's area and the left angular gyrus. We did find that participants in the Broca's area group were generally slower in naming than participants in the left angular gyrus group across stimulation types. However, since no interactions were observed between the effects of stimulation type and stimulation site, slower naming in the Broca's area group is likely due to individual variability, which occurred by chance even though the two groups had similar ages and education levels, rather than to differential stimulation effects.

Main effects of stimulation were limited to the naming task and were not observed in the lexical decision task, in either the primary analysis of the full data set or the exploratory analysis limited to data from participants' first sessions only. One account for this is that lexical decision may simply be a less "natural" task than naming, yielding higher individual variability and thus providing less power to detect any group effects. However, also given the absence of interactions between stimulation (type or site) and linguistic categories, it is more likely that the difference in outcomes is due to the qualitatively different cognitive and linguistic processes involved in the two tasks. Compared to lexical decision, naming involves the additional components of visual processing, object recognition, conceptual/semantic processing, phonological and articulatory planning, and speech motor execution—modulation of any of which could have contributed to the observed effect of stimulation. Many previous studies into the impact of tDCS on language processing have used naming tasks similar to that used in the present study, so it may be considered a fairly standard outcome measure for language studies (Fertonani et al., 2010; Holland et al., 2011; Sparing et al., 2008; Wirth et al., 2011; etc.). However, while naming is certainly an ecologically valid and functionally relevant task that has a potential both as an outcome measure and as a target for aphasia treatment, investigation of effects of stimulation on more diverse linguistic tasks could shed more light on the nature of observed effects (i.e., which specific linguistic sub-processes are modulated by stimulation of focal brain regions).

Besides the increase in naming speed under conditions of cathodal stimulation over Broca's area and the left angular gyrus, another significant finding of the study was the interaction between session number and stimulation type, indicating that the effect of anodal and cathodal stimulation varied depending on whether these stimulation types were administered in participant's first, second, or third session. We originally introduced the factor of session number into the statistical model in order to account for any main effects of task practice over time. No such main effects of session number were observed, but it did consistently modulate the effect of stimulation across outcome measures and tasks (in naming reaction times, lexical decision accuracy, and lexical decision reaction times). Since our study was not originally designed to investigate any such interactions, it did not have enough power to more specifically explore how stimulation effects differed in the first, second, and third session. More research is warranted that would investigate this question by having large sample sizes in stimulation order groupings. One potential account for the emerging interaction of stimulation type and order may be that stimulation effects are modulated by task novelty, consistent with findings by Dockery, Hueckel-Weng, Birbaumer, and Plewnia (2009) who showed that effects of tDCS in a planning task (the Tower of London) were specific to the training phase. However, in the present study the interaction was found not only in the lexical decision task, which was likely novel to most or all participants in their first session, but also in the more "natural" naming

task. Thus, a more likely possibility is that the interactions reflect physiological carry-over effects between stimulation types. So far, prolonged effects of tDCS have typically been reported after administration of multiple stimulation sessions (Brunoni et al., 2012; Olma et al., 2013). However, it is possible that multiple sessions are only necessary in order for behavioral effects to last without further stimulation, whereas if a different type of stimulation is administered as a follow-up, its effects may be modified even by a single previous administration of stimulation. These findings have important implications for the field: first, they call for further research on the duration of HD-tDCS effects; but also, they have implications for design of further research that is not focused on long-lasting effects. In this case, between-subjects designs where each participant only receives one stimulation type may possibly be a better choice than within-subject (e.g., cross-over) designs, until more is known about the effects of different stimulation polarities on one another.

Since HD-tDCS is a novel method, an additional aim of the study was to obtain more information on the method itself. The findings add to the evidence of high safety and tolerability of HD-tDCS (Borckardt, 2012; Garnett & den Ouden, 2015). Participants gave relatively low pain and unpleasantness ratings and these subsided within a 20-min session at 2 mA. Only one participant possibly experienced short-term side effects. This suggests good tolerability and thus a high potential for routine clinical use of HD-tDCS.

Conclusions

We found that cathodal stimulation over both Broca's area and the left angular gyrus increased naming speed for both verbs and nouns, challenging the traditional view of cathodal stimulation as suppressive or leading to decreased performance. The effect did not extend to the lexical decision task. Additionally, effects of specific stimulation types depended on the order of their administration, suggesting possible physiological carry-over and/or task novelty effects. These results are relevant to the application of HD-tDCS to enhance and direct neural plasticity in patients with neurogenic language disorders.

Acknowledgements

The authors thank Julius Fridriksson, Paul Fillmore, Marom Bikson, and Abhishek Datta for their help with technology; and Kelly Marlowe, Ece Bayram, and Keiko Bridwell for their help with data analysis. Research was supported by ASPIRE-I grant #11560-13-32582 (DdO) of the University of South Carolina. The article was partly prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

References

Akinina, Y., Malyutina, S., Ivanova, M., Iskra, E., Mannova, E., & Dragoy, O. (2015). Russian normative data for 375 action pictures and verbs. *Behavior Research Methods*, *47*(3), 691–707. http://dx.doi.org/10.3758/s13428-014-0492-9

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*(2), 12–28.

Baker, J. M., Rorden, C., & Fridriksson, J. (2010). Using transcranial direct-current stimulation to treat stroke patients with aphasia. *Stroke, 41*(6), 1229–1236. http://dx.doi.org/10.1161/STROKEAHA.109.576785

Bartlett, M. S. (1947). The use of transformations. *Biometrics, 3*(1), 39–52.

Batsikadze, G., Moliadze, V., Paulus, W., Kuo, M.-F., & Nitsche, M. A. (2013). Partially non-linear stimulation intensity-dependent effects of direct current stimulation on motor cortex excitability in humans. *Journal of Physiology, 591*(7), 1987–2000. http://dx.doi.org/10.1113/jphysiol.2012.249730

Bazzini, A., Zonca, G., Craca, A., Cafforio, E., Cellamare, F., Guarnaschelli, C., … Luzzatti, C. (2012). Rehabilitation of argument structure deficits in aphasia. *Aphasiology, 26*(12), 1440–1460. http://dx.doi.org/10.1080/02687038.2012.690023

Bikson, M., & Rahman, A. (2013). Origins of specificity during tDCS: Anatomical, activity-selective, and input-bias mechanisms. *Frontiers in Human Neuroscience, 7*, 688. http://dx.doi.org/10.3389/fnhum.2013.00688

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9–10), 341–345.

Borckardt, J. J., Bikson, M., Frohman, H., Reeves, S. T., Datta, A., Bansal, V., … George, M. S. (2012). A pilot study of the tolerability and effects of high-definition transcranial direct current stimulation (HD-tDCS) on pain perception. *Journal of Pain, 13*(2), 112–120. http://dx.doi.org/10.1016/j.jpain.2011.07.001

Brunoni, A. R., Nitsche, M. A., Bolognini, N., Bikson, M., Wagner, T., Merabet, L., … Fregni, F. (2012). Clinical research with transcranial direct current stimulation (tDCS): challenges and future directions. *Brain Stimulation, 5*(3), 175–195. http://dx.doi.org/10.1016/j.brs.2011.03.002

Cappa, S. F., Sandrini, M., Rossini, P. M., Sosta, K., & Miniussi, C. (2002). The role of the left frontal lobe in action naming: rTMS evidence. *Neurology, 59*(5), 720–723. http://dx.doi.org/10.1212/WNL.59.5.720

Cattaneo, Z., Pisoni, A., & Papagno, C. (2011). Transcranial direct current stimulation over Broca's region improves phonemic and semantic fluency in healthy individuals. *Neuroscience, 183*, 64–70. http://dx.doi.org/10.1016/j.neuroscience.2011.03.058

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology, 33*(4), 497–505. http://dx.doi.org/10.1080/14640748108400805

Datta, A., Bansal, V., Diaz, J., Patel, J., Reato, D., & Bikson, M. (2009). Gyri-precise head model of transcranial direct current stimulation: Improved spatial focality using a ring electrode

versus conventional rectangular pad. *Brain Stimulation, 2*(4), 201–207. http://dx.doi.org/10.1016/j.brs.2009.03.005

Davis, N. J., Gold, E., Pascual-Leone, A., & Bracewell, R. M. (2013). Challenges of proper placebo control for non-invasive brain stimulation in clinical and experimental applications. *European Journal of Neuroscience, 38*(7), 2973–2977. http://dx.doi.org/10.1111/ejn.12307

De Vries, M. H., Barth, A. C. R., Maiworm, S., Knecht, S., Zwitserlood, P., & Flöel, A. (2010). Electrical stimulation of Broca's area enhances implicit learning of an artificial grammar. *Journal of Cognitive Neuroscience, 22*(11), 2427–2436. http://dx.doi.org/10.1162/jocn.2009.21385

Den Ouden, D.-B., Fix, S., Parrish, T. B., & Thompson, C. K. (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics, 22*(2), 196–215. http://dx.doi.org/10.1016/j.jneuroling.2008.10.004

Devlin, J. T., & Watkins, K. E. (2007). Stimulating Language: Insights from TMS. *Brain, 130*(3), 610–622. http://dx.doi.org/10.1093/brain/awl331

Dockery, C. A., Hueckel-Weng, R., Birbaumer, N., & Plewnia, C. (2009). Enhancement of planning ability by transcranial direct current stimulation. *Journal of Neuroscience, 29*(22), 7271–7277. http://dx.doi.org/10.1523/JNEUROSCI.0065-09.2009

Fertonani, A., Rosini, S., Cotelli, M., Rossini, P. M., & Miniussi, C. (2010). Naming facilitation induced by transcranial direct current stimulation. *Behavioural Brain Research, 208*(2), 311–318. http://dx.doi.org/10.1016/j.bbr.2009.10.030

Filmer, H. L., Mattingley, J. B., & Dux, P. E. (2013). Improved multitasking following prefrontal tDCS. *Cortex, 49*(10), 2845–2852. http://dx.doi.org/10.1016/j.cortex.2013.08.015

Fiori, V., Coccia, M., Marinelli, C. V., Vecchi, V., Bonifazi, S., Ceravolo, M. G., … Marangolo, P. (2011). Transcranial direct current stimulation improves word retrieval in healthy and nonfluent aphasic subjects. *Journal of Cognitive Neuroscience, 23*(9), 2309–2323. http://dx.doi.org/10.1162/jocn.2010.21579

Flöel, A., Meinzer, M., Kirstein, R., Nijhof, S., Deppe, M., Knecht, S., & Breintenstein, C. (2011). Short-term anomia training and electrical brain stimulation. *Stroke, 42*(7), 2065–2067. http://dx.doi.org/10.1161/STROKEAHA.110.609032

Fridriksson, J., Richardson, J. D., Baker, J. M., & Rorden, C. (2011). Transcranial direct current stimulation improves naming reaction time in fluent aphasia: a double-blind, sham-controlled study. *Stroke, 42*(3), 819–821. http://dx.doi.org/10.1161/STROKEAHA.110.600288

Fujiyama, H., Hyde, J., Hinder, M. R., Kim, S.-J., McCormack, G. H., Vickers, J. C., & Summers, J. J. (2014). Delayed plastic responses to anodal tDCS in older adults. *Frontiers in Aging Neuroscience, 6*, 115. http://dx.doi.org/10.3389/fnagi.2014.00115

Garnett, E. O., Malyutina, S., Datta, A., & den Ouden, D.-B. (2015). On the use of the terms anodal and cathodal in high-definition transcranial direct current stimulation: A technical note. *Neuromodulation: Technology at the Neural Interface*. Advance online publication. http://dx.doi.org/10.1111/ner.12320

Garnett, E. O., & den Ouden, D.-B. (2015). Validating a sham condition for use in high definition transcranial direct current stimulation. *Brain Stimulation*, *8*(3), 551–554. http://dx.doi.org/10.1016/j.brs.2015.01.399

Gill, J., Shah-Basak, P. P., & Hamilton, R. (2015). It's the thought that counts: Examining the task-dependent effects of transcranial direct current stimulation on executive function. *Brain Stimulation, 8*(2), 253–259. http://dx.doi.org/10.1016/j.brs.2014.10.018

Guleyupoglu, B., Febles, N., Minhas, P., Hahn, C., & Bikson, M. (2014) Reduced discomfort during high-definition transcutaneous stimulation using 6% benzocaine. *Frontiers in Neuroengineering, 7*, 28. http://dx.doi.org/10.3389/fneng.2014.00028

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences, 9*(9), 416–423. http://dx.doi.org/10.1016/j.tics.2005.07.004

Holland, R., & Crinion, J. (2012). Can tDCS enhance treatment of aphasia after stroke? *Aphasiology, 26*(9), 1169–1191. http://dx.doi.org/10.1080/02687038.2011.616925

Holland, R., Leff, A. P., Josephs, O., Galea, J. M., Desikan, M., Price, C. J., … Crinion, J. (2011). Speech facilitation by left inferior frontal cortex stimulation. *Current Biology, 21*(16), 1403–1407. http://dx.doi.org/10.1016/j.cub.2011.07.021

Horvath, J. C., Forte, J. D., & Carter, O. (2015). Evidence that transcranial direct current stimulation (tDCS) generates little-to-no reliable neurophysiologic effect beyond MEP amplitude modulation in healthy human subjects: A systematic review, *Neuropsychologia, 66*, 213–236. http://dx.doi.org/10.1016/j.neuropsychologia.2014.11.021

Hoy, K. E., Emonson, M. R. L., Arnold, S. L., Thomson, R. H., Daskalakis, Z. J., & Fitzgerald, P. B. (2013). Testing the limits: Investigating the effect of tDCS dose on working memory enhancement in healthy controls. *Neuropsychologia, 51*(9), 1777–1784. http://dx.doi.org/10.1016/j.neuropsychologia.2013.05.018

Hoyle, M. H. (1973). Transformations—An introduction and a bibliography. *International Statistical Review, 41*(2), 203–223.

Iyer, M. B., Mattu, U., Grafman, J., Lomarev, M., Sato, S., Wassermann, E. M. (2005). Safety and cognitive effect of frontal DC brain polarization in healthy individuals. *Neurology, 64*(5), 872–875. http://dx.doi.org/10.1212/01.WNL.0000152986.07469.E9

Kessler, S. K., Turkeltaub, P. E., Benson, J. G., & Hamilton, R. H. (2012). Differences in the experience of active and sham transcranial direct current stimulation. *Brain Stimulation, 5*(2), 155–162. http://dx.doi.org/10.1016/j.brs.2011.02.007

Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: Implications for lexical organization. *Brain and Language, 74*(1), 1–25. http://dx.doi.org/10.1006/brln.2000.2315

Kiss, K. (2000). Effect of verb complexity on agrammatic aphasics' sentence production. In: R. Bastiaanse & Y. Grodzinsky (Eds.). *Grammatical disorders in aphasia: A neurolinguistic perspective* (pp. 123–151). London: Whurr.

Liuzzi, G., Freundlieb, N., Ridder, V., Hoppe, J., Heise, K., Zimerman, M., … Hummel, F. C. (2010). The involvement of the left motor cortex in learning of a novel action word lexicon. *Current Biology, 20*(19), 1745–1751. http://dx.doi.org/10.1016/j.cub.2010.08.034

Marangolo, P., Fiori, V., Di Paola, M., Cipollari, S., Razzano, C., Oliveri, M., & Caltagirone, C. (2013). Differential involvement of the left frontal and temporal regions in verb naming: A tDCS treatment study. *Restorative Neurology and Neuroscience, 31*(1), 63–72. http://dx.doi.org/10.3233/RNN-120268

Medler, D. A., & Binder, J. R. (2005). MCWord: An On-Line Orthographic Database of the English Language. http://www.neuro.mcw.edu/mcword/

Minhas, P., Bansal, V., Patel, J., Ho, J. S., Diaz, J., Datta, A., & Bikson, M. (2010). Electrodes for high-definition transcutaneous DC stimulation for applications in drug delivery and electrotherapy, including tDCS. *Journal of Neuroscience Methods, 190*(2), 188–197. http://dx.doi.org/10.1016/j.jneumeth.2010.05.007

Monti, A., Ferrucci, R., Fumagalli, M., Mameli, F., Cogiamanian, F., Ardolino, G., & Priori, A. (2013). Transcranial direct current stimulation (tDCS) and language. *Journal of Neurology, Neurosurgery and Psychiatry, 84*(8), 832–842. http://dx.doi.org/10.1136/jnnp-2012-302825

Naeser, M. A., Martin, P. I., Ho, M., Treglia, E., Kaplan, E., Bashir, S., & Pascual-Leone, A. (2012). Transcranial magnetic

stimulation and aphasia rehabilitation. *Archives of Physical Medicine and Rehabilitation, 93*(Suppl. 1), S26–S34. http://dx.doi.org/10.1016/j.apmr.2011.04.026

Olma, M. C., Dargie, R. A., Behrens, J. R., Kraft, A., Irlbacher, K., Fahle, M., & Brandt, S. A. (2013). Long-term effects of serial anodal tDCS on motion perception in subjects with occipital stroke measured in the unaffected visual hemifield. *Frontiers in Human Neuroscience, 7*, 314. http://dx.doi.org/10.3389/fnhum.2013.00314

Pirulli, C., Fertonani, A., & Miniussi, C. (2014). Is neural hyperpolarization by cathodal stimulation always detrimental at the behavioral level? *Frontiers in Behavioral Neuroscience, 8*(226). http://dx.doi.org/10.3389/fnbeh.2014.00226

Price, A., Bonner, M., Hamilton, R., Peelle, J., & Grossman, M. (2015). Modulating language comprehension using HD-tDCS. Poster presented at the NYC Neuromodulation Conference, New York, January 9–11, 2015.

Price, A. R., McAdams, H., Grossman, M., & Hamilton, R. H. (2015). A Meta-analysis of Transcranial Direct Current Stimulation Studies Examining the Reliability of Effects on Language Measures. *Brain Stimulation*. Advance online publication. http://dx.doi.org/ 10.1016/j.brs.2015.06.013

Richardson, J. D., Fillmore, P., Datta, A., Truong, D., Bikson, M., & Fridriksson, J. (2014). Toward development of sham protocols for high-definition transcranial direct current stimulation (HD-tDCS). *NeuroRegulation, 1*(1), 62–72. http://dx.doi.org/10.15540/nr.1.1.62

Rodríguez-Ferreiro, J., Andreu, L., & Sanz-Torrent, M. (2014). Argument structure and the representation of abstract semantics. *PLoS ONE, 9*(8). http://dx.doi.org/10.1371/journal.pone.0104645

Scholz, A. (2002). Mechanisms of (local) anaesthetics on voltage-gated sodium and other ion channels. *British Journal of Anaesthesia, 89*(1), 52–61. http://dx.doi.org/10.1093/bja/aef163

Shapiro, L. P., Brookins, B., Gordon, B., & Nagel, N. (1991). Verb effects during sentence processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*(5), 983–996. http://dx.doi.org/10.1037/0278-7393.17.5.983

Sparing, R., Dafotakis, M., Meister, I. G., Thirugnanasambandam, N., & Fink, G. R. (2008). Enhancing language performance with non-invasive brain stimulation—a transcranial direct current stimulation study in healthy humans. *Neuropsychologia, 46*(1), 261–268. http://dx.doi.org/10.1016/j.neuropsychologia.2007.07.009

Stagg, C. J. (2014). The physiological basis of brain stimulation. In: R. C. Kadosh (Ed.). *The stimulated brain: Cognitive enhancement using non-invasive brain stimulation* (pp. 145–177). London: Elsevier Academic Press.

Stagg, C. J., & Nitsche, M. A. (2011). Physiological basis of transcranial direct current stimulation. *Neuroscientist, 17*(1), 37–53. http://dx.doi.org/10.1177/1073858410386614

Thompson, C. K., Bonakdarpour, B., Fix, S. C., Blumenfeld, H. K., Parrish, T. B., Gitelman, D. R., & Mesulam, M.-M. (2007). Neural correlates of verb argument structure processing. *Journal of Cognitive Neuroscience, 19*(11), 1753–1767. http://dx.doi.org/10.1162/jocn.2007.19.11.1753

Thompson, C. K., Bonakdarpour, B., & Fix, S. F. (2010). Neural mechanisms of verb argument structure processing in agrammatic aphasic and healthy age-matched listeners. *Journal of Cognitive Neuroscience, 22*(9), 1993–2011. http://dx.doi.org/10.1162/jocn.2009.21334

Thompson, C. K., Lange, K. L., Schneider, S. L., & Shapiro, L. P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology, 11*(4–5), 473–490.

Thompson, C. K., Riley, E. A., den Ouden, D.-B., Meltzer-Asscher, A., & Lukic, S. (2013). Training verb argument structure production in agrammatic aphasia: Behavioral and neural recovery patterns. *Cortex, 49*(9), 2358–2376. http://dx.doi.org/10.1016/j.cortex.2013.02.003

Thompson-Schill, S. L. (2005). Dissecting the language organ: A new look at the role of Broca's area in language processing. In: A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 173–189). Hillsdale NJ: Lawrence Erlbaum Associates.

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews, 35*(3), 407–426. http://dx.doi.org/10.1016/j.neubiorev.2010.04.007

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., … Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage, 30*(4), 1414–1432. http://dx.doi.org/10.1016/j.neuroimage.2005.11.002

Vines, B. W., Norton, A. C., & Schlaug, G. (2011). Non-invasive brain stimulation enhances the effects of melodic intonation therapy. *Frontiers in Psychology, 2*, 230. http://dx.doi.org/10.3389/fpsyg.2011.00230

Wirth, M., Rahman, R. A., Kuenecke, J., Koenig, T., Horn, H., Sommer, W., & Dierks, T. (2011). Effects of transcranial direct current stimulation (tDCS) on behaviour and electrophysiology of language production. *Neuropsychologia, 49*(14), 3989–3998. http://dx.doi.org/10.1016/j.neuropsychologia.2011.10.015

You, D. S., Kim, D.-Y., Chun, M. H., Jung, S. E., & Park, S. J. (2011). Cathodal transcranial direct current stimulation of the right Wernicke's area improves comprehension in subacute stroke patients. *Brain and Language, 119*(1), 1–5. http://dx.doi.org/10.1016/j.bandl.2011.05.002

Received: July 31, 2015
Accepted: September 18, 2015
Published: October 8, 2015

NeuroRegulation

Interactions Between Discrimination and Control of EEG Alpha

Jon A. Frederick[1], Kelli N. Dunn[1], and Thomas F. Collura[2,3]*

[1]Middle Tennessee State University, Murfreesboro, Tennessee, USA
[2]BrainMaster Technologies Inc., Bedford, Ohio, USA
[3]Brain Enrichment Center, Bedford, Ohio, USA

Abstract

The relationship between discrimination and control of physiological states is largely unexplored, although it is often suggested that this relationship is important for the mechanism of action of biofeedback. This pilot study examined 6 participants given seven sessions of alpha discrimination training combined with standard neurofeedback "control" training. Four subjects achieved five criterion (binomial $p < .05$) sessions in the discrimination task. The discrimination task performances correlated significantly with performance in the amplitude control task. Evidence that some subjects can use the intertrial interval (ITI) to predict the correct responses in the discrimination task led to an examination of how ITIs were distributed with respect to success (correct or incorrect) and type of trial (same or different from previous) in these and 40 additional subjects from archival data (Frederick, 2012). This analysis found that some information about the correct responses was conveyed by the ITI, but participants made relatively little use of this information. However, the criterion discrimination sessions showed dramatic changes in the distribution of ITIs in the present (but not the archival) study, suggesting that participants were controlling their electroencephalogram (EEG) during these sessions. These findings provide preliminary evidence of generalization of skills between these two tasks.

Keywords: discrimination learning; perceptual motor processes; consciousness states; electroencephalography; biofeedback; neurotherapy

Citation: Frederick, J. A., Dunn, K. N., & Collura, T. F. (2015). Interactions Between Discrimination and Control of EEG Alpha. *NeuroRegulation*, *2*(3), 126–136. http://dx.doi.org/10.15540/nr.2.3.126

***Address correspondence to**: Dr. Jon A. Frederick, Middle Tennessee State University, P.O. Box 87, 1301 East Main Street, Murfreesboro, TN 37132-0001, USA. Email: Jon.Frederick@mtsu.edu

Edited by:
John Davis, PhD, McMaster University, Ontario, Canada

Reviewed by:
John Davis, PhD, McMaster University, Ontario, Canada
Larry Stevens, PhD, Northern Arizona University, Flagstaff, Arizona, USA

Introduction

One of the first demonstrations of operant conditioning of the electroencephalogram (EEG) was a study that showed how human subjects could be trained to discriminate high from low alpha (8–12 Hz) amplitude states (Kamiya, 1962, 1968, 2011). In an operant discrimination task, the experimenter uses a random schedule to decide whether the next trial will be high or low, and provides a prompt when the EEG spontaneously exceeds a threshold. Subjects respond *high* or *low* and are immediately informed whether their response is correct. Kamiya (1968) noticed how participants trained to discriminate alpha subsequently performed better than naive subjects at controlling their alpha amplitude in a standard neurofeedback task—where subjects were rewarded when alpha amplitude exceeded a threshold. This observation was important, because it suggested a relationship between the different psychological constructs trained by the two tasks.

Frederick (2012, *in press*) argued that EEG operant discrimination generally trains observation or awareness of brainwave states, while standard neurofeedback (or "control training") trains volition or control of these states. It is commonly argued that increasing awareness of subtle subjective correlates of physiological states is central to the mechanism of action of biofeedback (Brener, 1974; Congedo & Joffe, 2007; Frederick, *in press*; Olson, 1987; Plotkin, 1981). However, despite a half a century of evolution in neurofeedback since Kamiya's

discovery, there have been very few studies of EEG state discrimination (e.g., Cinciripini, 1984; Kotchoubey, Kubler, Strehl, Flor, & Birbaumer, 2002). Kamiya's initial result was only recently replicated (Frederick, 2012).

Thus, the relationship between the skills trained by EEG state discrimination training and control training remains poorly understood. Control training was observed to facilitate subsequent discrimination training for the sensorimotor rhythm (12–15 Hz; Cinciripini, 1984), and slow cortical potentials (Kotchoubey et al., 2002). However, the only known demonstration of discrimination training facilitating learning of a physiological control task was seen in a peripheral vasomotor response by Fudge and Adams (1985). Kamiya's (1968) report of facilitation of alpha control by prior discrimination training was anecdotal and did not include quantitative data.

The present report describes a pilot study to explore the relationship between alpha discrimination and control. Initially, the hypothesis was that dividing session time equally between state discrimination training and standard neurofeedback training would result in greater learning of both tasks than training of either task alone. This began with a preliminary study of the two-tasks combined condition. With limited resources to run subjects and sessions, it was reasoned that before running the individual tasks separately as controls, a high level of learning should first be seen in the combined condition. In fact, the level of learning was only moderate, prompting a redesigned study that is currently in progress. While facilitation of learning was not directly tested by the pilot study design, the study design did test a more general hypothesis, that there would be a correlation in performance of the two tasks. While such a correlation is not sufficient to demonstrate generalization between the skills involved in the two tasks, it is necessary. It was also predicted that subjects would show a learning curve for both tasks, with better performance on the final sessions than the early sessions.

The present study also sought to resolve technical issues in the design of the state discrimination protocol. A review of the data from an earlier discrimination study (Logsdon, Cox, West, & Frederick, 2013) found that one subject had spuriously achieved significant performance by repeating the previously correct response on trials with very short (4–8 s) intertrial intervals (ITIs). Orne and Wilson (1978) warned about how, with fixed thresholds, the length of the interval between prompts could convey some information about the stimulus, allowing subjects to respond correctly without truly discriminating their physiological states. For instance, if a subject's EEG amplitude changes relatively slowly, then a brief time between trials could be used by subjects to correctly infer that a given trial would be the same (high or low) as the last one. In the Logsdon et al. (2013) data, this advantage was only seen for intervals of less than 8 s between trials (Jon Frederick, unpublished observations). Therefore, in the present study, the minimum ITI was set for 8.1 s, and the null hypothesis was that there would be no advantage to perseverative responding on short ITI trials. To allow further comparison, the same hypothesis was examined in archival data (Frederick, 2012) from 40 subjects.

Method

Participants

With approval of the Middle Tennessee State University Institutional Review Board, 6 participants (3 female; age 18–44) were recruited from students at Middle Tennessee State University and the surrounding community. Participants were compensated $10 per session after their final session. Participants were required to have a peak alpha frequency (peak median alpha amplitude between 8–12 Hz; PAF) evident in a 60-s baseline recording.

Measurements and Apparatus

Skin at the recording sites was prepared to bring impedance below 10 kΩ, with no site greater than twice the others. Taking into account both the comfort of participants and modern amplifier input impedances (Feree, Luu, Russell, & Tucker, 2001), impedances up to 15 kΩ were sometimes accepted if repeated preparations did not bring them lower. Tin electrodes were attached to the parietal midline (Pz). Left and right ears were randomly assigned to reference and ground each session.

EEG was recorded with a BrainMaster Atlantis amplifier and BrainMaster 3.7i software using the default settings (Butterworth filter order 6; default passband 0.5–64 Hz; peak-to-peak amplitude scale; 60 Hz input notch filter; 256 samples per second). Rewards were controlled with a BrainMaster Event Wizard protocol where the alpha signal had a damping factor of 10, with sustain rewarded criterion and refractory period set to zero.

The alpha band was defined as each subject's peak alpha frequency (PAF) plus or minus 2 Hz (Klimesch, 1999). For example, if the PAF were 11

Hz, the alpha band was then defined as 9–13 Hz. Clear alpha peaks were seen in all subjects. If a consistent peak alpha frequency were seen across three or more sessions, that value was used in subsequent sessions even if the participant deviated from that value on a subsequent day.

For the alpha amplitude "control" task, the experimenter watched a 25-s event trend window of the filtered alpha amplitude and was allowed to change the reward threshold in real time. BrainMaster 3.0 displayed a running average of the percent time in reward for the most recent 60 s. The experimenter attempted to maintain the participant's percent time in reward close to 30%. About every 20 seconds, the threshold was adjusted by 0.5 to 1.0 μV if the percent time in reward was less than 10% or greater than 40%. Effort was taken to avoid changing the threshold when the subject's alpha amplitude was close to it.

Alpha amplitudes from the control task were visually examined as 1-min averages alongside averages in the delta (0–3 Hz) and hibeta (20–30 Hz) bands. One-minute segments with excessively high delta or hibeta amplitudes where assumed to result from muscle artifact and excluded from the analysis.

For the discrimination task, Fourier-transformed amplitudes for each 1 Hz band from 1–32 Hz were passed to a dynamic link library using shared memory, and then sampled in 10 times per second by custom real-time software (*Introspect*, written in C++), which recorded both EEG and task responses. The sum of amplitudes in lodelta (0.5–2 Hz) and hibeta (23–32 Hz) were each continuously monitored as artifact channels, and recording and task were automatically suspended (and an artifact warning tone played) whenever either value exceeded a threshold.

Frederick (2012) identified several signal parameters that resulted in significantly better discrimination of EEG alpha. Subjects discriminated absolute amplitudes better than relative amplitudes; 5 Hz bandwidths surrounding the peak alpha frequency better than 1 Hz bandwidths; 2- and 4-s stimulus durations (EEG smoothing averages) better than 1-s durations; and stimulus magnitudes far from the median (below the 10th and above the 90th percentile) better than more moderate stimulus magnitudes (near the 30th or 70th percentile). However, longer intertrial intervals are required when waiting for signals with extreme durations or magnitudes. This results in a trade-off between signal quality and the number of opportunities for learning (trials per minute) that can be administered during a session. While the use of absolute amplitude and 5 Hz bandwidths were clearly indicated, the use of 2-s stimulus durations and moderate stimulus magnitudes (30th and 70th percentiles) were needed to achieve a desired 3–4 prompts per minute.

The dimensions of the two tasks were made as similar as possible. For instance, the target of 30% time in reward for the control task corresponded to the 30th and 70th percentile thresholds to trigger a prompt in the discrimination task. A 60-s baseline was used in the discrimination task because 60 s was the maximum sliding baseline for the percent time calculations provided by the BrainMaster event wizard.

Procedure

After obtaining informed consent, participants sat in a reclining chair with eyes closed in a dimly lit, sound-attenuated room. Participants were instructed about the nature of muscle artifact and strategies to relax and minimize it.

Each session consisted of 20 min of alpha control training followed by about 20 minutes of alpha discrimination training. The alpha control training started with 5 min during which high alpha amplitude was rewarded (the "enhance" condition), then 5 min during which low alpha amplitude was rewarded (the "inhibit" condition. These two conditions were then repeated. To control for spontaneous shifts in baseline alpha amplitude, the measure of performance in this task was defined as the percent amplitude difference between the high and low amplitude conditions.

Each run of 5 min in the control task was prefaced by saying, "Now you will be rewarded for increasing (or decreasing) alpha. Are you ready?" After at least one run of a condition, the threshold from the most recent run was re-used, and then adjusted as needed to bring the reward percentage near 30%.

Before the discrimination task, participants were provided with a set of written instructions as described (Frederick, 2012). The instructions explained that EEG alpha usually means a relaxed but alert state with eyes closed; that high alpha might be increased by clearing, emptying, or quieting the mind, or disconnecting from mental contents. Low alpha was described as the opposite, the presence of imagination, attention to sensory details, thought, intention, or inner speech. It was emphasized that their own experiences before and

after each prompt were equally or more important, as these instructions were only rough guidelines. A strategy was suggested that if no discrimination prompt was received in a long time (e.g., 30 s), to try changing their mental state and see if that evoked a prompt.

Alpha state discrimination training consisted of a median of 40 (min 30, max 60) trials. A 60-s eyes-closed baseline EEG was recorded each session. During the task, each EEG epoch was ranked among a percentile distribution of alpha amplitudes of the most recent 60 s initially derived from the baseline recording. The baseline was updated with each response, or whenever the experimenter pressed the pause button. Triggering of new prompts was suspended for 8 s after each prompt or after resuming from a pause. A random number generator determined in advance whether each trial would be high or low. A prompt tone then sounded whenever the alpha band amplitude exceeded a critical threshold difference from the median of the baseline. Recording was then suspended until the subject responded.

The critical threshold for triggering a prompt was set at the 30th percentile for low alpha trials, and the 70th percentile for high alpha trials. Subjects responded “high” or “low” with a keypress response, and received immediate verbal feedback after each trial whether the response was correct. Software was programmed to exclude runs of six or more of the same (high or low) trial type, although participants were not informed of this constraint.

Analysis

Performance in the EEG alpha amplitude control task was analyzed as the percent difference, or the amplitude difference between the high and the low conditions, divided by the overall average amplitude. The choice of a percent difference rather than a raw amplitude difference served to control for variance in factors such as skull thickness, which may have masked real differences in achievement in the task.

Performance in the EEG alpha discrimination task was analyzed in terms of probability under the binomial theorem, where a criterion performance was defined as a significant number correct with binomial $p < .05$. Four subjects achieved a total of five criterion sessions in the discrimination task.

Results

There was a strong association between performance in the discrimination task and percent difference between the high and low conditions of the amplitude control task. Among seven sessions, the five criterion performances on the discrimination task all occurred on sessions with the first, second, or third (median second, of 7) highest amplitude difference in the control task (Table 1).

Table 1
Percent Amplitude Difference Between High and Low Conditions in the Control Task, and Percent Correct in the Discrimination Task, Over Seven Sessions

		Session						
Subject	**Task**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
mt003	control	MD	MD	-10	-16	**22**	7	**11**
	discrim	36	50	50	58	**90****	48	**60***
mt004	control	-4	-7	15	0	**8**	-2	1
	discrim	51	43	52	49	**62.5***	52	57
mt005	control	7	-6	-14	20	**15**	6	13
	discrim	58	55	53	35	**65***	53	40
mt007	control	-6	7	21	-2	8	24	**18**
	discrim	51	50	55	56	48	60	**69****
average	control	-1	-2	3	0	13	9	11
	discrim	49	49	52	50	66	53	57

Note. *Denotes binomial $p < .05$; **Denotes binomial $p < .01$; MD, missing data.

The covariation in performance between the two tasks is clearly seen when the mean discrimination performance is plotted as a difference from 50 percent alongside the control task performance (Figure 1). Quantitatively, the mean within-subject correlation between the two tasks had Pearson's r = .34, which was not significantly different from zero, $t(3)$ = 1.49, p = .12, d = 0.744. However, when these data were analyzed categorically (reducing the contribution of random error from below chance scores on the two tasks), the mean point-biserial correlation between criterion performances on the discrimination task and above average performance on the amplitude control task was r = .57, $t(3)$ = 3.96, p = .014, d = 1.98.

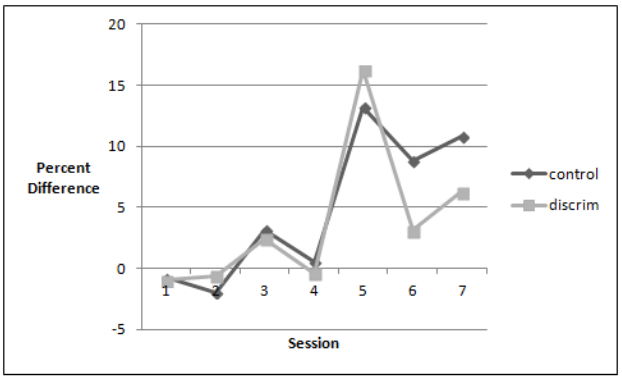

Figure 1. Covariation in performance of alpha control task (mean difference in percent amplitude between enhance and inhibit conditions) and alpha discrimination task (difference from 50% correct). Each line represents the mean of n = 4 subjects.

There was a strong learning curve effect, where performances in both tasks were nonrandom with respect to session number. The mean within subject correlation between session number and above average task performance for the control task was r = .43, $t(3)$ = 3.77, p = .016, d = 1.89. The mean correlation between session number and criterion performance for the discrimination task was r = .40, $t(3)$ = 2.91, p = .031, d = 1.45. Thus, it is possible that there was no causal relationship between the two performance variables, but an independent effect of learning caused both variables to covary over time.

However, in three cases performance declined in sessions after the first criterion performance (after session 5; see Table 1). In these cases, the mean point-biserial correlation between the two tasks across sessions 5–7 was r = 0.67, $t(2)$ = 4.00, p = .029, d = 2.31.

As previously reported (Frederick, 2012), shorter times between sessions appeared to improve discrimination performance. The five criterion sessions had a mean 4.4 days since the previous session, compared to 8.1 for the remaining sessions, $t(31)$ = 1.806, p = .040, d = 0.877. A similar effect was not seen for the control task.

Information Conveyed by Intertrial Intervals

The distribution of intertrial intervals (ITIs) and same versus different trials (from previous trial types) was examined to determine whether ITIs were conveying information about the type of trial (Orne & Wilson, 1978). The 34 non-criterion sessions were first studied as a control condition. Not including first trials when the ITI was undefined, there were 1486 discrimination trials in these sessions. The mean ITI across all trials was 16.6 s (*SD* 10.8 s). There were 726 or 48.9% of trials that were the same versus 759 or 51.1% that were different from the previous (high or low) trial type. This difference was expected because runs of six or greater identical trial types were prevented by the Introspect software.

The pattern of same and different trials with respect to ITI was examined to see if there was an advantage to responding *same* as the previous correct trial type on short ITI trials, or to responding *different* on long ITI trials. To give all subjects equal weight (despite differing numbers of trials), the percentage of trials in each ITI was first computed within each subject and then the mean percentage was computed across subjects. This allowed for accurate degrees of freedom and valid statistical comparisons (n = 6 subjects, not n = 1,486 trials).

Consistent with the concern that information about the type of trial was conveyed by the ITI, the percent of same trials (% Same) was significantly higher in the ≤ 8.1 s category than the > 45.0 s category, one-tailed $t(5)$ = 4.28, p= .0039, d = 1.75 (54.3 vs. 26.6% of trials, Table 2). The advantage (% Adv) of using this information was significantly different from zero for ITIs of 8.1 s, $t(5)$ = 2.09, p = .046, d = 0.85; 30.1–45.0 s, $t(5)$ = 2.02, p = .049, d = 0.83; and > 45.0 s, $t(5)$ = 3.65, p = .0074, d = 1.49. One-tailed tests were used for these comparisons because first, there were theoretical reasons to predict an advantage to *same* responding on the shortest (8.1 s) ITI trials, and *different* on the longest (> 45.0 s) ITI trials (Cott, Pavloski, & Black, 1981; Orne & Wilson, 1978). Since it was unclear where the boundary between short and long was between these extremes, it was reasoned that an effective p = .10 alpha-level was justified if a one-tailed test was not, because type II error for these comparisons

amounts to type I error for the validity of the discrimination paradigm.

Including all three categories, then, the ITI may have conveyed useful information on 28.2% of trials. The total weighted advantage across all ITIs was 3.8%, meaning that a subject with perfect information and a perfect sense of timing could on have scored 53.8% using ITI information alone.

However, this information could only contribute to scores on the task if participants had a large response bias in the correct direction. The response bias (% Bias) is quantified as the difference between success on same trials (Same S) and different trials (Diff S). Overall, participants tended to show a bias toward *same* responding, where the average score on same trials was 57.1% and the average score on different trials was 41.8%. To use ITI information to score effectively in the task, % Bias must approach 100%. For instance, a subject who responded 100% *same* on 8.1-s ITI trials where % Adv was 4.3% could expect an average score of 54.3% on those trials (100% correct on same trials, 0% on different trials). Subjects showed a significant response bias for ITIs of 8.1 s, one-tailed $t(5) = 2.16$, $p = .042$, $d = 0.88$; 8.2–9.0 s, $t(5) = 2.17$, $p = .041$, $d = 0.88$; 9.1–15.0 s, $t(5) = 2.69$, $p = .022$, $d = 1.10$; and 15.1–30.0 s, $t(5) = 2.15$, $p = .042$, $d = 0.88$. However, the percent points gained (% Gain) from use of the ITI is found by multiplying % Bias by **%** Adv times % Obs. As would be expected from noncriterion sessions where the overall mean score was 49.5%, % Gain was near zero for all ITIs. Table 2 could be summarized by saying that subjects had a large bias where it gained them no advantage, and no bias where the advantage was large.

Table 2

Information Conveyed by the Time Between Prompts and Discrimination Task Success for Noncriterion Sessions

ITI, s	% Obs	% Same	% Adv	Same S	Diff S	Avg S	% Bias	% Gain
8.1	17.8	54.3	4.3*	59.5	37.5	49.8	22.1*	0.2
8.2–9.0	7.5	51.1	1.1	57.2	30.7	44.2	26.6*	0.0
9.1–15.0	33.3	51.5	1.5	56.5	39.5	48.1	17.1*	0.1
15.1–30.0	31.0	47.7	-2.3	57.0	45.8	50.8	11.2*	-0.1
30.1–45.0	6.6	37.7	-12.3*	60.2	53.8	54.9	6.4	-0.1
> 45.0	3.8	26.6	-23.4**	46.1	50.0	50.9	-3.8	0.0

Note. 6 subjects; 1,486 trials. ITI, intertrial interval in seconds; % Obs, percent of trials in ITI category; % Same, percent of trials same as the correct response for the previous trial; % Adv, possible increase in score on these trials or advantage by using ITI—negative sign indicates advantage for responding *different*; Same S, percent success on same trials; Diff S, percent success on different trials; Avg S, average success; % Bias, difference between Same S and Diff S; % Gain, total percent points resulting from % Bias given %Adv and % Obs. *Denotes one-tailed *t*-test $p < .05$. **Denotes one-tailed *t*-test $p < .01$.

The overall percentages of same versus different trials were similar for the five criterion sessions (195 trials; 48.7% same, 51.3% different; Table 3). The mean ITI across all 195 trials was 19.1 s (*SD* 12.9 s). A much greater difference of % Same between short (8.1–9.0 s) and longer (> 9.0 s) ITIs was seen in the criterion sessions. The difference in % Same between 8.1 s and > 45 s was highly significant, $t(3) = 10.64$, $p < .001$, $d = 5.32$. There was a significant advantage to responding *same* on 8.1-s trials (33.6%), $t(3) = 5.80$, $p = .0051$, $d = 2.90$; and *different* for 15.1- to 30.0-s trials (11.9%), $t(3) = 5.80$, $p = .0051$, $d = 2.40$. The advantage approached significance for the 8.2- to 9.0-s trials (27.1%), $t(3) = 1.72$, $p = .092$, $d = 0.86$; and of 10.4% for the > 45-s trials (10.4%), $t(3) = 1.67$, $p = .097$, $d = 0.83$. These four categories included 59% of trials.

The distribution of % Same was also significantly different from the noncriterion sessions for 8.1 s (83.6 vs. 54.3%), two-tailed between groups $t(8) = 5.56$, $p = .00053$, $d = 3.59$; and 15.1–30.0 s (38.1 vs. 47.7%), $t(8) = 2.54$, $p = .034$, $d = 1.64$; and approached significance for 8.2–9.0 s (77.1 vs. 51.1%), $t(8) = 2.04$, $p = .075$, $d = 1.32$. The effect size was moderate to large ($d = 0.57$ to $d = 0.89$) for the others.

Given the distribution of % Same in Table 3, the total weighted advantage for responding based on ITIs alone was 13.6%. A participant with perfect information and perfect timing could, then, score 63.6% using ITI information alone.

Subjects in the criterion sessions showed a strong overall response bias toward perseverative or *same*

responding. Whereas the overall average score was 67.0% for the criterion sessions, the mean score was 82.7% on same trials (Same S) versus 52.8% on different trials (Diff S). The % Bias was significant for 9.1–15 s, one-tailed $t(3) = 2.36$, $p = .049$, $d = 1.18$; and approached significance for 8.1–9.0 s, $t(3) = 1.98$, $p = .070$, $d = 0.99$. Gain scores were computed for individual subjects. Three subjects had no Diff S trials for either the 8.1 or 8.2–9.0 categories. After combining these categories all subjects had at least one trial, allowing for computation of % Bias and % Gain. When summed across all ITIs, the mean % Gain across 4 subjects was 2.0%, which was not significant, $t(3) = .56$, one-tailed $p = .31$, $d = 0.28$.

One subject, however, had a % Gain of 10.3 (explaining more than half of his overall score of 69%) of which 9.6 points were earned from positive advantage and positive bias from 8.1- to 15.0-s ITIs.

Table 3

Information Conveyed by the Time Between Prompts and Discrimination Task Success for Criterion Sessions

ITI, s	% Obs	% Same	% Adv	Same S	Diff S	Avg S	% Bias	% Gain
8.1	18.3	83.6	33.6**	79.9				
8.2–9.0	5.1	77.1	27.1	100.0				
8.1–9.0	23.4	80.2	30.2	83.7	37.5	74.5	46.2	3.3
9.1–15.0	31.7	43.1	-6.9	80.7	32.4	50.1	48.3*	-1.1
15.1–30.0	27.2	38.1	-11.9**	80.0	66.5	71.1	13.5	-0.4
30.1–45.0	9.3	53.0	3.0	79.2	71.7	72.6	7.5	0.0
> 45.0	8.3	39.6	-10.4	75.0	93.8	89.6	-18.8	0.2

Note. 4 subjects; 195 trials. ITI, % Obs, % Same, % Adv, Same S, Diff S, Avg S, % Bias, % Gain, see Table 2 caption. *Denotes one-tailed *t*-test $p < .05$. **Denotes one-tailed *t*-test $p < .01$.

It was of interest whether the distribution of ITIs had conveyed information about the stimuli in a previous study with larger sample size (Frederick, 2012). All noncriterion sessions were first examined from subjects who had achieved criterion in this archival data set. Data from one subject and 28 sessions were excluded because they had the minimum ITI set higher than 4.1 s, leaving 38 subjects, 152 sessions, and 14,279 trials (Table 4). The mean ITI for these sessions was 15.5 s (*SD* 16.1 s) and the average score was 50.1%.

These data had 6,969 or 48.8% same trials versus 7,310 or 51.2% different trials. These data also showed a pattern of greater % Same trials for short ITIs and lower percent same trials for longer ITIs. Although small, the difference in % Same between the shortest (4.1 s) and longest (> 45.0 s) ITIs bordered on significance with $n = 38$ (54.7 vs. 46.1%), $t(37) = 1.68$, $p = .051$, $d = 0.48$. There was a significant advantage for responding based upon ITI for 4.1 s, $t(37) = 1.92$, $p = .031$, $d = 0.31$; 5.1–8.0 s, $t(37) = 1.93$, $p = .031$, $d = 0.31$; 8.1–15.0 s, $t(37) = 3.18$, $p = .0015$, $d = 0.52$; and 15.1–30.0 s, $t(37) = 2.46$, $p = .0092$, $d = 0.40$.

The total weighted advantage was 3.1%, meaning that information conveyed by ITIs could be used to score up to 53.1% correct.

The % Bias was significant for 4.1 s, $t(35) = 6.64$, $p < .001$, $d = 1.106$; 4.2–5.0 s, $t(34) = 3.40$, $p < .001$, $d = 0.57$; 5.1–8.0 s, $t(37) = 7.61$, $p < .001$, $d = 1.23$; 8.1–15.0 s, $t(37) = 5.87$, $p < .001$, $d = 0.95$; 15.1–30.0 s, $t(37) = 3.76$, $p < .001$, $d = 0.61$; and approached significance for > 45.0 s, $t(33) = 1.64$, $p = .055$, $d = 0.28$. However, owing to small % Adv requiring 100% bias to fully take advantage, this response bias resulted in near zero % Gain in score for these noncriterion sessions.

Table 4
Information Conveyed by the Time Between Prompts and Discrimination Task Success for Noncriterion Sessions

ITI, s	% Obs	% Same	% Adv	Same S	Diff S	Avg S	% Bias	% Gain
4.1	10.3	54.7	4.7*	63.5	33.8	51.4	28.6**	0.1
4.2–5.0	6.0	53.3	3.3	62.0	40.5	51.1	24.9**	0.0
5.1–8.0	23.1	52.8	2.8*	64.4	38.6	52.7	20.8**	0.3
8.1–15.0	27.9	46.9	-3.1**	58.7	41.4	49.2	17.3**	-0.2
15.1–30.0	21.3	47.2	-2.8**	54.6	43.3	48.4	11.3**	-0.1
30.1–45.0	6.3	48.0	-2.0	52.2	48.7	49.0	4.9	-0.1
> 45.0	5.0	46.1	-3.9	45.7	55.9	50.8	-9.7	0.1

Note. Archival data, n = 38 subjects; 14,279 trials (Frederick, 2012). ITI, % Obs, % Same, % Adv, Same S, Diff S, Avg S, % Bias, % Gain, see Table 2. *Denotes one-tailed t-test $p < .05$. **Denotes one-tailed t-test $p < .01$. % Bias and % Gain were computed only from subjects who had at least one same and one different trial in a given ITI category, unlike %Obs, %Adv, Same S, or Diff S. Thus the % Bias and % Gain reported in Tables 4 and 5 are not exactly equal to those computed from other columns in the table.

The distribution of ITIs versus same/different trials and performance was also examined in the criterion sessions from the same archival subject group (Frederick, 2012; Table 5). Two sessions and 23 sets (parts of sessions) were excluded because the minimum ITI was set higher than 4.1 s, leaving 76 sets and 16 sessions from 36 subjects and 4,532 trials. The mean ITI was 14.7 s (*SD* 14.8 s), and the mean score was 69.2%. The difference in % Same between the shortest (4.1 s) and longest (> 45.0 s) ITIs approached significance (54.2 vs. 46.2%), $t(28) = 1.34$, $p = .094$, $d = 0.25$.

Unlike the four criterion subjects from the present study, in the archival subjects there were no significant differences in % Same between the criterion and noncriterion data. Only the 15.1–30.0 s ITIs showed a significant difference from 50% (45.5%), $t(36) = 1.77$, $p = .042$, $d = 0.30$.

The total weighted advantage was 3.3%. A subject with perfect information and a perfect sense of timing could thus score 53.3% using ITI information alone.

Participants showed a strong bias toward *same* responding, scoring 74.3% overall on same trials and 61.6% on different trials. % Bias was significant for 4.1 s, one-tailed $t(31) = 3.03$, $p = .0025$, $d = 0.53$; 4.2–5.0 s, $t(25) = 3.24$, $p = .0017$, $d = 0.63$; and 5.1–8.0 s, $t(33) = 3.60$, $p = .00052$, $d = 0.62$; and approached significance for 30.1–45.0 s, $t(23) = 1.49$, $p = .075$, $d = 0.30$.

The % Gain was significantly different from zero for 4.1 s, $t(31) = 2.44$, $p = .010$, $d = 0.43$; and 4.2–5.0 s, $t(25) = 2.01$, $p = .027$, $d = 0.39$. The overall mean % Gain was 0.9%, which was significant, $t(35) = 1.69$, $p = .0499$, $d = 0.28$. Among 36 gain scores, there were an equal number (18) positive and negative. If the distribution of negative gain scores is assumed to represent random variation, the absolute value of scores beyond the 90th percentile (-2.23%) could be used as a one-tailed $p < .05$ test for suspicious positive gain scores. Six subjects had gain scores in this category: 3.0, 3.6, 4.6, 4.8, 11.3, and 11.9% (by contrast, the six most negative gain scores were -1.4, -1.7, -2.1, -2.2, -2.2, and -2.7%). Among the six subjects with suspicious positive gain scores, subtracting the positive gain would have resulted in performances below a binomial $p = .05$ for two subjects and between $p = .05$ and $p = .01$ for two subjects.

The 11.9% score belonged to a subject who scored 94.1% overall (explaining at most 27% of his score; $p < .001$ after removing this effect). However, the 11.3% gain score could explain as much as 48% of that subject's score of 73.5%. Nearly all of the 11.3% was earned from positive advantage and positive bias on short (4.1–8.0 s) ITI trials. Subtracting this gain resulted in binomial $p = .09$.

Table 5
Information Conveyed by the Time Between Prompts and Discrimination Task Success for Criterion Sessions

ITI, s	% Obs	% Same	% Adv	Same S	Diff S	Avg S	% Bias	% Gain
4.1	14.9	54.2	4.2	78.2	56.8	71.6	23.3**	0.5*
4.2–5.0	6.1	56.8	6.8	80.0	54.9	71.0	26.0**	0.3*
5.1–8.0	21.0	52.2	2.2	79.7	59.6	71.2	19.3**	0.1
8.1–15.0	26.0	52.3	2.3	69.7	63.4	67.4	7.3	0.1
15.1–30.0	22.2	45.5	-4.5*	69.5	65.4	66.8	3.6	0.0
30.1–45.0	6.2	50.3	0.3	77.9	63.2	71.0	11.2	-0.1
> 45.0	3.6	46.2	-3.8	72.5	64.0	69.9	3.1	0.1

Note. Archival data, n = 36 subjects; 76 sets, 16 sessions; 4,532 trials. ITI, % Obs, % Same, % Adv, Same S, Diff S, Avg S, % Bias, % Gain, see Table 2. *Denotes one-tailed t-test $p < .05$. **Denotes one-tailed t-test $p < .01$.

Discussion

This study showed a strong association between performance in an EEG alpha control task and an EEG alpha discrimination task over the course of seven sessions. The best performances on the discrimination task tended occur on the same days when participants achieved the greatest differences between the enhance and inhibit conditions in the control task.

However, performance also showed a significant correlation with session number for both tasks, suggesting a learning curve effect. This learning curve effect might have been a confound for interpreting some generalization of skills between the two tasks. However, results for several subjects after the peak of the learning curve served as a control, actually showing a higher correlation when there was no learning effect.

The use of categorical data for computing correlations (criterion performance on the discrimination task and above average performance on the amplitude control task) rather than raw interval data could be argued to be a "cherry-picking" of analytical methods, since the correlation between raw performance scores was not significant. The rationale was that discrimination scores below 50% and amplitude differences below 0% are functionally equal, so nominalizing these data reduces error variance. However, all statistics from a pilot study with n = 4 are to be interpreted with caution and only as suggestions for further research.

The overall performance on the discrimination task in this subject group seemed to be less successful that observed in Frederick (2012). Whereas 56% of subjects achieved $p < .01$ performance by the seventh session in that study, only 2 out of 6 subjects achieved $p < .01$ in this study.

There were several differences between the two studies that may have contributed to reduced performance in the discrimination task. Most notably, only half rather than the entire session time was spent practicing the discrimination task in this study. Further, Frederick (2012) varied the prompt threshold within the 0–30 and 70–100 percentile ranges (mean difference from 50th percentile, 34.3, *SD* 8.9), whereas they were set at constant 30th and 70th percentiles in the present study (mean difference 28.7, *SD* 9.5).

However, several other factors were expected to improve performance in this study. The present study used absolute (not relative) amplitude for all trials rather than about half of trials, and 5-Hz alpha bandwidths for all trials rather than about one third of trials. This study also consistently used the same parameters for absolute amplitude, bandwidth, stimulus duration, and location, whereas varying these parameters in Frederick (2012) might have confused participants.

Another factor that may have influenced performance was that participants received verbal rather than automated feedback on the discrimination task in this study. Modeling the importance of the client-therapist relationship, it was reasoned that having the experimenter say whether a response was correct would might improve motivation by conveying that a person was in the room who cared. However, it was noted that this verbal feedback often took 200 or more milliseconds to initiate after the response; and created opportunities for variance in how this information

was perceived. Sherlin et al. (2011) noted the importance of minimizing the delay of reinforcement after the behavior, citing studies showing how learning can be adversely affected by latencies exceeding 250–350 ms (Felsinger, Gladstone, Yamaguchi, & Clark, 1947; Grice, 1948). The motivational advantage of personal feedback may, then, be outweighed by the need to present feedback quickly and consistently.

The observation that ITIs conveyed significant information about the types of trials in both the present and the archival study—the tendency for % Same to be greater for shorter ITIs and lower for the longer ITIs—represents a flaw in the design of both studies. Frederick (2012) argued that the use of the sliding baseline avoided this complication. The relatively small (3.1 to 3.3%) advantage of using this information in that study suggests that the sliding baseline reduced, but did not eliminate this problem. The advantages seen in the present data set (3.8 and 13.6%) show that increasing the minimum ITI to 8.1 s did not reduce the problem. Subjects in both studies tended to have the greatest response biases for *same* responding for the shortest ITIs and the greatest response biases for *different* responding for the longest ITIs. However, these biases where rarely large enough to result in gains in score that were both significant and substantial relative to the total score.

Future research should take steps to reduce the possible advantage of responding based on the ITI. For instance, our discrimination task software has now been revised to automatically increase the minimum ITI whenever the difference in number of trials between same and different trial types is greater than one for 3.1–5.0, 5.1–10.0, and 10.1–15.0 s intervals. One benefit to this analysis has been the elimination of the 8.1 minimum ITI, allowing for more trials per minute or a higher prompt threshold.

An unexpected discovery in this study was the significant difference in the distribution of same versus different trials between the criterion and noncriterion sessions (Tables 2 and 3). Shorter ITIs had a much higher % Same and longer ITIs had much lower % Same. This difference was not seen in the archival study, which had not included control task training (Tables 4 and 5). One possible interpretation is that participants were successfully stabilizing and controlling their EEG—generalizing their control task skill to the discrimination session. Participants in the criterion sessions tended to perseverate, not only in reporting the previously correct state (for all ITIs), but also in maintaining it (for up to 9.0 s).

The overall bias toward *same* responding might also reflect an honest strategy for guessing one's internal state in the absence of a clear perception. Thus, if one's subjective perceptions are no different on one trial compared to the previous one, it is reasonable to assume that one's alpha amplitude hasn't changed either.

This study has provided preliminary evidence of generalization between the skills involved in EEG state discrimination and standard neurofeedback control tasks, both in the correlation of task performances and in participants' tendency to stabilize their alpha amplitude between trials in the criterion discrimination sessions (greater % Same on short ITI trials).

Future studies in this laboratory will explore more directly whether there is a causal interaction between the skills involved in these two tasks. Currently, we are assessing whether the skills generalize or transfer, by measuring performance on one task after seven training sessions on the other task.

It is argued that awareness or explicit processing is important to early stages of learning (Fitts & Posner, 1967; Gentile, 2000), but can actually decrease performance on highly practiced tasks (Beilock & Carr, 2001). It stands to reason, then, that adding discrimination training to the early stages of standard neurofeedback might increase this explicit processing and enhance learning of the standard neurofeedback task. Thus, an additional subject group in our study will receive both tasks in the same session—similar to this pilot study—to assess whether combining the two types of training results in better learning than either task alone.

References

Beilock S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130*(4), 701–725. http://dx.doi.org/10.1037/0096-3445.130.4.701

Brener, J. (1974). A general model of voluntary control applied to the phenomena of learned cardiovascular change. In P. Obrist, A. H. Black, J. Brener, & L. V. DiCara (Eds.), *Cardiovascular psychophysiology*. Chicago: Aldine.

Cinciripini, P. M. (1984). Discrimination of sensorimotor EEG (12–15 Hz) activity: A comparison of response, production, and no-feedback training conditions. *Psychophysiology*, *21*(1), 54–62. http://dx.doi.org/ 10.1111/j.1469-8986.1984.tb02317.x

Congedo, M., & Joffe, D. (2007). Multichannel tomographic neurofeedback: Wave of the future? In J. R. Evans (Ed.),

Handbook of neurofeedback. New York: The Hayworth Medical Press.

Cott, A., Pavloski, R. P., & Black, A. H. (1981). Operant conditioning and discrimination of alpha: Some methodological limitations inherent in response-discrimination experiments. *Journal of Experimental Psychology: General, 110*(3), 398–414. http://dx.doi.org/10.1037/0096-3445.110.3.398

Felsinger, J. M., Gladstone, A. I., Yamaguchi, H. G., & Clark, L. (1947). Reaction latency (StR) as a function of the number of reinforcements (N). *Journal of Experimental Psychology, 37*(3), 214–228. http://dx.doi.org/10.1037/h0055587

Feree, T. C., Luu, P., Russell, G. S., & Tucker, D. M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clinical Neurophysiology, 112*(3), 536–544. http://dx.doi.org/10.1016/S1388-2457(00)00533-2

Fitts, P. M., & Posner, M. I. (1967). *Human performance.* Oxford, England: Brooks and Cole.

Frederick, J. A. (2012). Psychophysics of EEG alpha state discrimination. *Consciousness and Cognition, 21*(3), 1345–1354. http://dx.doi.org/10.1016/j.concog.2012.06.009

Frederick, J. A. (*in press*). EEG state discrimination and the phenomenal correlates of brainwave states. In T. F. Collura & J. A. Frederick (Eds.), *Clinician's Companion to QEEG and Neurofeedback* (annotated and with an introduction by J. Kiffer). New York: Taylor & Francis.

Fudge, R., & Adams, H. D. (1985). The effects of discrimination training on voluntary control of cephalic vasomotor activity. *Psychophysiology, 22*(3), 300–306.

Gentile, A. M. (2000). Skill acquisition: Action, movement, and neuromotor processes. In J. Carr & R. Shepherd (Eds.), *Movement science: Foundations for physical therapy in rehabilitation* (2nd ed., pp. 111–187). Gaithersburg, MD: Aspen.

Grice, G. R. (1948). The relation of secondary reinforcement to delayed reward in visual discrimination learning. *Journal of Experimental Psychology, 38*(1), 1–16. http://dx.doi.org/10.1037/h0061016

Kamiya, J. (1962). *Conditioned discrimination of the EEG alpha rhythm in humans.* Proceedings of the Western Psychological Association, San Francisco, California.

Kamiya, J. (1968). Conscious control of brain waves. *Psychology Today, 1,* 57–60.

Kamiya, J. (2011). The first communications about operant conditioning of the EEG. *Journal of Neurotherapy, 15*(1), 65–73. http://dx.doi.org/10.1080/10874208.2011.545764

Klimesch, W. (1999). Full-length review: EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews, 29*(2–3), 169–195. http://dx.doi.org/10.1016/S0165-0173(98)00056-3

Kotchoubey, B., Kübler, A., Strehl, U., Flor, H., & Birbaumer, N. (2002). Can humans perceive their brain states? *Consciousness and Cognition, 11*(1), 98–113. http://dx.doi.org/10.1006/ccog.2001.0535

Logsdon, H., Cox, C., West, K., & Frederick, J. A. (2013). First evidence of 14–18 Hz EEG state discrimination at Fz and anterior cingulate in human subjects. Presented at ISNR, Dallas, TX.

Olson, P. (1987). Definitions of biofeedback. In M. S. Schwartz (Ed.), *Biofeedback: A practitioner's guide* (pp. 33–38). New York: The Guilford Press.

Orne, M. T., & Wilson, S. K. (1978). On the nature of alpha feedback training. In G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 2). New York: Plenum Press.

Plotkin, W. B. (1981). A rapprochement of the operant-conditioning and awareness views of biofeedback training: The role of discrimination in voluntary control. *Journal of Experimental Psychology: General, 110*(3), 415–428. http://dx.doi.org/10.1037/0096-3445.110.3.415

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. http://dx.doi.org/10.1080/10874208.2011.623089

Received: July 30, 2015
Accepted: September 20, 2015
Published: October 8, 2015

NeuroRegulation

Using Neurofeedback to Lower Anxiety Symptoms Using Individualized qEEG Protocols: A Pilot Study

*Stephanie M. Dreis, Angela M. Gouger, Edward G. Perez, G. Michael Russo, Michael A. Fitzsimmons, and Mark S. Jones**

The University of Texas at San Antonio, San Antonio, Texas, USA

Abstract

Introduction: Anxiety disorders affect approximately 40 million Americans ages 18 and over (NIMH, 2015). Although qualitative and small-scale quantitative neurofeedback (NF) studies show reduction in anxiety symptoms, large-scale studies and quantitative electroencephalogram (qEEG) driven protocols are non-existent. This retrospective pilot study intended to assess whether qEEG guided amplitude NF is viable in symptom reduction of anxiety. Methods: Nineteen clients were assessed for anxiety, 14 were included in the data. Demographics include age ranges from 11–61 (*M* = 31.71, *SD* = 16.33), 9 male and 5 female; six identified as Caucasian, five as Hispanic/Latino, and three Caucasian/Hispanic ethnicity. Pre- and post-assessments included the Zung Self-Rating Anxiety Scale, Screen for Child Anxiety Related Disorders (SCARED), and the Achenbach System of Empirically Based Assessment (ASEBA). Clients received 30-min qEEG guided NF treatment sessions, twice a week. The range of attended session was 7–28 (*M* = 12.93, *SD* = 6.32). Results: Enhancement in clients' well-being was evidenced by statistically significant improvement in symptom measures scores. Although improvements for the two most anxiety-related categories on the ASEBA were not significant, other anxiety-related categories did show significant improvement. Yet, qEEG findings were not statistically significant. Directions for future research are discussed.

Keywords: anxiety; anxiety symptoms; qEEG guided amplitude neurofeedback; neurofeedback; *z*-scores

Citation: Dreis, S. M., Gouger, A. M., Perez, E. G., Russo, G. M., Fitzsimmons, M. A., & Jones, M. S. (2015). Using Neurofeedback to Lower Anxiety Symptoms Using Individualized qEEG Protocols: A Pilot Study. *NeuroRegulation*, *2*(3), 137–148. http://dx.doi.org/10.15540/nr.2.3.137

***Address correspondence to**: Dr. Mark Jones, Department of Counseling, The University of Texas at San Antonio, 501 Cesar Chavez Blvd., Durango Building 3.304E, San Antonio, TX 78207, USA. Email: mark.jones@utsa.edu

Edited by:
Rex Cannon, PhD, Neural Potential, Florida, USA

Reviewed by:
John Davis, PhD, McMaster University, Ontario, Canada
Randall Lyle, PhD, Mount Mercy University, Iowa, USA

Introduction

According to the National Institute of Mental Health (NIMH), anxiety disorders rank as the top leading diagnosis by clinicians within the mental health field. Anxiety disorders affect approximately 18% of the United States population, or 40 million individuals within a given year (NIMH, 2015). While the majority of Americans experience stress periodically within their lifespan, individuals diagnosed with anxiety have severe pervasive symptoms that interfere with their daily lives. Three of the most commonly diagnosed types of anxiety disorders are: generalized anxiety disorder, 6.8 million adult Americans; panic disorder, 6 million adult Americans; and social phobia, 15 million adult Americans (NIMH, 2015). Psychotherapy, cognitive behavioral therapy (CBT), exposure-based treatment, stress management techniques, meditation, and aerobic exercise are various therapeutic modalities that may or may not be used in conjunction with medication in the treatment of anxiety disorders (NIMH, 2015).

With the onset frequently developing during childhood, many anxiety disorders can be persistent if not treated and present more frequently in women at a 2:1 ratio (American Psychiatric Association, 2013). A variety of symptoms are reported by individuals with anxiety disorders including: trouble falling asleep and staying asleep, fatigue, headaches, and muscle tension (NIMH, 2015). More severe symptoms can include sudden and repeated attacks of fear, pounding and racing heart,

and purposely excluding oneself from certain people or places.

Literature Review

Various biofeedback modalities have been implemented by clinicians in the treatment of anxiety including: electromyography (EMG), peripheral temperature, and electrodermal response (EDR) prior to neurofeedback's (NF) popularization (Price & Budzynski, 2009). NF, a subcategory of biofeedback, is a method of self-regulation which uses a brain-computer interface to promote neural plasticity, by providing feedback to an individual about their brain's electrical activity at a specific scalp location in a specified frequency range (Cannon, 2015). NF has been used to lower anxiety symptoms in a variety of populations, as addressed throughout the following reviewed literature.

A study by Kerson, Sherman, and Kozlowski (2009) illustrates how the various modalities of earlobe temperature training, alpha suppression, and alpha symmetry training were used in eight adults who either were diagnosed with generalized anxiety disorder or presented with multiple anxious behaviors. Participants were assessed for high alpha frequency at the International 10–20 Electrode system sites Fp1, Fp2, F3, F4, F7, and F8. A 5-min baseline electroencephalogram (EEG) of the participants was recorded with their eyes open for the initial measurement and with their eyes closed for the secondary measurement. Post-baseline measures were also recorded 1 week after the last NF training occurred. The initial six sessions were used to increase the participant's earlobe temperature. The following 6–16 sessions consisted of decreasing alpha magnitude by 10% in the anterior lobes for 30 or more minutes. Once alpha was suppressed, the protocol shifted to improvement of alpha symmetry by a 15% increment for 30 minutes or more during 8–32 sessions. All sessions were conducted on a biweekly basis. Continued assessment of participants was conducted throughout the study by means of The State-Trait Anxiety Inventory (STAI; Spielberger, 1983) in which a significant improvement in scores resulted. The pre- and post-mean change in EEG was 1.41 *z*-scores towards the mean. Limitations mentioned within the study include: a limited amount of participants, lack of variance in protocols, and the lack of a control group.

A study conducted by Cheon et al. (2015) researched NF implemented on 77 adults diagnosed with various psychiatric disorders within a psychiatric setting. The following disorders are listed in order of prevalence according to the research: depressive disorders, anxiety disorders, sleep disorders, somatoform disorders, adjustment disorders, bipolar disorder, schizophrenia, attention-deficit/hyperactivity disorder, alcohol dependence, game addiction, and impulse control disorder. Protocols were designed depending on the participant's chief complaint (e.g., anxiety, emotional instability, lethargy, etc.), the opinion of the attending psychiatrist, neuropsychiatric evaluation results, and the subjective-symptom-rating scale. The clinical Global Impression-Severity Scale (CGI-S; Busner & Targum, 2007) and the Hill-Castro (2002) checklist were also implemented on a weekly basis as a measure of treatment effectiveness. NF protocols included training sensorimotor rhythm (SMR), beta, and/or also contained alpha-theta training. The various frequency bandwidths which were rewarded during training, included: SMR from 12 to 15 Hz, beta from 15 to 18 Hz, theta from 5 to 8 Hz, and alpha between 8 and 12 Hz. The individualized site locations in which training was implemented included: Fp1, Fp2, F3, F4, F7, F8, T3, T4, C3, C4, P1, P2, O1, O2, and Oz based on the International 10–20 Electrode system. Alpha-theta training was conducted at the PZ site location. Protocols were evaluated and finalized during weekly NF meetings, which included a team of three psychiatrists trained in NF, as well as a trained NF therapist. The number of appointments for a client's training ranged from 1 to 20 or more sessions. The Hill-Castro Checklist score showed an improvement in multiple symptom areas including anxiety (p = .0001). The pre- and post-CGI score showed a significant reduction in the severity of symptoms (p < .001). Limitations mentioned within the study included having a heterogeneous group and no control group, as well as not utilizing the quantitative electroencephalography (qEEG) to determine protocols.

Singer (2004) used NF on two female dancers, 27 and 52 years of age, who had persistent levels of performance anxiety. A STAI assessment was taken by each participant before a NF session and before each of their major dance performances. The course of NF treatment included 20 sessions at the time interval of 30 min per session. Sensors were placed on site locations T3 and T4 and thresholds were adjusted during each session dependent upon the participant's response. Post assessments indicated a significant decrease in anxiety symptoms associated with performance. The trait anxiety portion of the first participant's assessment indicated a decrease in score from 59 to 43.5, while the state

portion underwent a decrease in score of 66 to 44. The trait anxiety portion of the second participant's assessment indicated a decrease in score as well from 52 to 36, while the state portion underwent a decrease in score of 56 to 30. Limitations to this study included: a small sample size, lack of individualized protocols, and no control group.

Walker (2009) implemented a study based upon whether NF could lower anxiety symptoms for 19 clients diagnosed with post-traumatic stress disorder (PTSD). Four clients, who were originally diagnosed with PTSD and in the NF group, but had dropped out after the qEEG, were included in the control group. Each client received a qEEG using the NeuroGuide software. Results were compared to the Lifespan Normative database. Excessive high frequency beta (21–30 Hz) was then downtrained for five to seven sessions for each site that presented excessive high frequency beta; 10 Hz activity was uptrained at the same sites. The sites were in various and multiple areas depending on where the excessive beta was located, as protocols were determined by a qEEG. A self-rated anxiety Likert scale from 1 to 10 was also used to determine the presence of anxiety symptoms each participant had felt. The number of sessions per individual ranged from five to seven. Participants who had NF training had a significant reduction in self-rated anxiety with a pre-treatment score of 5/10 to 7/10, to a post-treatment score of 0/10 to 2/10, and 1 month after NF training the scores remaining between 0/10 to 2/10. Subjects who did not have NF training had little or no reduction in self-rated anxiety 3 months after their qEEG. Limitations with this study include using a self-rating scale for anxiety rather than an evidence-based assessment.

A study by Scheinost et al. (2013) evaluated 10 subjects with contamination anxiety to undergo functional magnetic resonance imaging (fMRI) NF training and compared their neural connectivity with real-time functional magnetic resonance imaging (rt-fMRI). A matched control group of 10 subjects that received sham fMRI-NF (SNF) of their matched pair was used. Subjects had an initial fMRI to localize their activity in the orbitofrontal cortex (OFC) from contamination anxiety. They then met with a psychologist to discuss strategies for manipulating brain activity that could later be refined during fMRI-NF. There were eight sessions total where subjects were shown contamination-related photos and asked to rate their anxiety on a scale of 1 to 5. The first and the last session consisted of subjects being asked to implement the personal coping mechanisms, which they would typically use to try to lessen their anxiety. The middle six sessions consisted of 90 min of fMRI-NF. The fMRI-NF sessions consisted of subjects receiving cues of when to increase activity their OFC area, when to decrease activity, and when to rest based on their OFC output. Resting cues included a neutral image. Between-group differences in fMRI's were identified using Wilcoxon's rank-sum test. The fMRI-NF group reported greater self-reported reduction in anxiety ($p = 0.02$) compared to the SNF group ($p = 0.45$). The fMRI-NF group had significant ($p < 0.05$) neural changes compared to the SNF group as recorded by the last fMRI taken several days after the last fMRI-NF session. The fMRI-NF group had significant decrease in connectivity for the brain regions associated with emotion processing, including: the insula and adjacent regions, the hippocampi, parahippocampal and entorhinal cortex, the right amygdala, the brain stem in the vicinity of the substantia nigra, the temporal pole, superior temporal sulcus, thalamus, and fusiform gyrus. The fMRI-NF group also had an increased degree of connectivity that was seen in prefrontal areas associated with emotion regulation and cognitive control, including: right lateral prefrontal cortex and bilateral portions of Brodmann's area 8. This study illustrated how changes directly resulting from fMRI-NF were possible and how structural changes can last days after a fMRI-NF session. This study also supported the idea of finding and confirming a localized area related to a symptom and using that area for fMRI-NF. Limitations to this study include low number of fMRI-NF sessions and a small sample size.

These studies illustrate how NF can be a viable tool in lowering anxiety symptoms. They each have their strengths and limitations. A substantial limitation is either using the same protocol for each patient and/or using a protocol based on symptoms alone. Protocols based on symptoms alone and/or using the same protocol for each patient bypasses the time, cost, and training of running a qEEG (Thompson & Thompson, 2003). Hammond (2010) expresses the importance of using a qEEG to identify heterogeneity in brain wave patterns, finding comorbidities, and looking for effects from medication.

Krigbaum and Wigton (2014) argue the importance of qEEG guided and *z*-score NF as it allows the clinician to develop a more individualized treatment plan which encompasses a qEEG baseline, history, and clinical status of the client. Wigton and Krigbaum (2015a) further assert how 19-channel *z*-score NF (19ZNF) protocols facilitate identifying the

link between localized cortical dysfunctions and connectivity issues associated with mental health symptoms. In this modality, qEEG metrics are compared to a normative database to create *z*-scores; then, those *z*-scores are incorporated into the NF protocol in real time during the session. This allows for pre-treatment assessment, a helpful tool in measuring progress with the client, and combining real-time assessment with the operant conditioning of NF. Thus, 19ZNF training is used to bring these scores closer to the mean, otherwise known as *normalizing*. Moreover, 19ZNF protocols also reduce the number of sessions, which is more economical for the clients. Wigton and Krigbaum's pilot study used 19ZNF to train the deviant *z*-scores.

Unlike Wigton and Krigbaum (2015a), this research is a pilot study which used single-channel qEEG guided amplitude training, rather than *z*-score training, for three reasons: (1) it is commonly used by many practitioners, (2) it is a straightforward method for students in training to learn before advancing to other modalities, and (3) the numerous one- or two-channel qEEG-guided amplitude training studies which exist in the literature, as reviewed by Wigton and Krigbaum (2015b). Therefore, based on the literature review, this retrospective pilot study sought to assess whether individualized qEEG-guided protocol amplitude NF is viable in symptom reduction of anxiety-related disorders.

Methods

Clients

Clients contacted the Sarabia Family Counseling Center at the University of Texas at San Antonio (UTSA) to receive therapy and NF treatment free of charge. Clients learned about the clinic through community referral sources and/or university media relations. Upon calling, clients were screened by clinically licensed, doctoral-level students in the UTSA Department of Counseling to determine if they met the criteria for anxiety-spectrum disorders. If the individual satisfied the clinical criteria, as well as the required biweekly availability and willingness to complete the treatment requirements on an ongoing basis, the clients were then scheduled to meet with a NF student clinician. Prior to completing any formal assessments of anxiety, student clinicians acquired a comprehensive informed consent from each client. As retrospective research, the study was deemed to be exempt from review by the UTSA Institutional Review Board.

The pilot study started with 19 clients that were seen over a period between one or two semesters; however, the average number of sessions that clients acquired was approximately 12.9 sessions. In order to preserve our sample size we relaxed the inclusion criteria to a minimum of seven sessions per client. Three clients were excluded from the study because they dropped out without completing the full round of sessions or completing the final assessments. The data sets of two clients were excluded from the study; of the two clients that were excluded, one client had previously received a regimen of NF treatment and the other admitted to daily use of cannabis. A total of 14 clients are represented in the data. Of the included clients, demographics consisted of 9 males and 5 females. Clients ranged in age from 11 to 61 years of age with the average age being 31.71 (*SD* = 16.33) years of age. Six clients identified as Caucasian, five as Hispanic/Latino, and three identified as mixed Caucasian and Hispanic ethnicity (see Table 1).

Table 1
Client Demographics

Client #	Age	Gender	Ethnicity	Number of Sessions
1	17	M	Hispanic	14
2	20	F	Hispanic	26
4	48	F	Hispanic	28
6	52	M	Caucasian	12
7	15	F	Caucasian	10
8	50	M	Caucasian	14
10	21	M	Hispanic	8
11	11	M	Hispanic Caucasian Mix	11
12	37	M	Hispanic Caucasian Mix	8
13	26	F	Hispanic	7
14	18	M	Hispanic Caucasian Mix	10
15	25	M	Caucasian	12
16	61	F	Caucasian	11
17	43	M	Caucasian	10

Therapists

The student clinicians consisted of master's-level students within a program certified by the nationally accredited Council for Accreditation of Counseling and Related Education Programs (CACREP). These students are also in the supervision phase of pursuing their Board Certification in NF (BCN); thus, were overseen by a certified and licensed supervisor. Students had previously completed the required didactic coursework that is recognized by The Biofeedback Certification International Alliance (BCIA; http://www.bcia.org).

Measures

A within-subjects research design was implemented, which included the following pre-conditional and post-conditional assessments: the Screen for Child Anxiety-Related Disorders (SCARED) for children and adolescents, the Zung Self-Rating Anxiety Scale for adults, the age-appropriate self-reports for the Achenbach System of Empirically Based Assessment (ASEBA), and qEEG. The symptom measurements were selected on: the bases of their focus on anxiety symptoms, widespread acceptance in the therapeutic community, and standardization.

The qEEG measures assessed deviances from a normative database, which were then used to develop individualized protocols for training. Pre- and post-assessment comparisons were made using *z*-score changes, where improvement is assumed when scores move toward the mean ($z = 0$). Some of the challenges related to this form of measure are discussed below, but *z*-score comparisons provide one form of common reference with which to compare individualized protocols across the treatment group (Wigton & Krigbaum, 2015a).

Instrumentation

The qEEGs were acquired via 19-channel recordings in the eyes-closed and eyes-open conditions in a resting state, using a BrainMaster (BrainMaster Technologies, Inc., Bedford, Ohio) Discovery 24 high-impedance amplifier and NeuroGuide (Applied NeuroScience, Inc., Largo, Florida) software. Recordings utilized correct size Electro-Cap (Electro-Cap International, Inc., Eaton, Ohio) 10–20 electrode appliances, which were fitted as per manufacturer's guidelines and ear-clip leads placed. Preparation of electrodes was performed in a manner adequate to achieve impedance levels of less than 5,000 Ω (Jones, 2015). NF was provided utilizing BrainMaster Atlantis two-channel amplifiers and BioExplorer (CyberEvolution, Inc., Seattle, Washington) software. Electrode site preparation was done by cleaning site, ground, and reference locations with rubbing alcohol and abrading using PCI prep pads and Nuprep. Gold-plated electrodes were attached to the clients using Ten-20 paste. Impedance measurements were taken to insure that interelectrode impedance was less than 5,000 Ω (Jones, 2015).

Protocols

Clients agreed to attend a minimum total number of 15 NF training sessions that were to be held at the same time, twice per week, and free of charge. Participants were instructed to discontinue the consumption of caffeine or any other non-essential substances that may alter the qEEG significantly, such as supplements or medications. At least a 24-hour window prior to the qEEG recording was suggested for clients to restrict consumption for non-essential substances, unless otherwise medically directed. All medically directed substances were factored into qEEG interpretation and protocol development.

Collectively, participants underwent an average of 12.93 sessions of NF with a range of 7 to 28 total sessions. Participants that did not meet our original set threshold of 15 sessions were included due to the aspect of increasing our client size for a sufficient statistical interpretation. A total of 181 sessions were completed between all of the participants (see Table 1). These training protocols consisted of amplitude uptraining and/or downtraining of selected frequency bands based on qEEG findings. Protocol selections were based on current research and reflect markers found to be associated with anxiety issues (Dantendorfer et al., 1996; Demerdzieva & Pop-Jordanova, 2011; Gold, Fachner, & Erkkilä, 2013; Gunkelman, 2006; Gurnee, 2000; Heller, Nitschke, Etienne, & Miller, 1997; Johnstone, Gunkelman, & Lunt, 2005; Machleidt, Gutjahr, Muegge, & Hinrich, 1985; Price & Budzynski, 2009; Savostyanov et al., 2009; Siciliani, Schiavon, & Tansella, 1975; Stern, 2005, p. 196; Tharawadeepimuk & Wongsawat, 2014; Walker, 2009).

Based on the preferences of the clients and clinical judgment of the practitioners, feedback was presented using a variety of formats: games, animations, sounds, and analogical presentations (such as the size of boxes representing the amplitude of the respective bandpass filtered EEG signals). Thresholds were set manually at the beginning of the session based on the aimed percentage of a successful reward rate of approximately 50% of the time. Periodic adjustments were made to the threshold settings

within and between sessions as needed to shape behavior towards the client's specific treatment goals. Records were made for each session, which included: frequency bands, threshold settings, session average amplitude levels, type of feedback utilized, and significant details from client reports and clinician impressions. EEG data was recorded for each session.

Table 2
Training Sites and Frequency Bands for Each Client

Client #	EC/EO	Site	Band1 Decrease	Band2 Increase	Band3 Decrease	Combined Sites
1	EO	Pz		8–12		
2	EO	F2	5–7	10–12	20–25	Fz/F4
4	EO	Pz	7–9		25–29	
6	EO	Pz	7–12		17–22	
7	EO	CPz			21–27	Cz/PZ
8	EO	Cz	7–9	12–15	19–24	
10	EO	Fz	5–9	12–15	25–30	
11	EO	Cz	20–25		25–30	
12	EO	Cz	3–6		25–30	
13	EO	Cz	4–7		18–25	
14	EO	Cz	3–5	12–15	20–25	
15	EO	Cz	1–5	12–15	25–30	
16	EO	Fz	3–5	12–15	8–11	
17	EC	Pz		8–10	25–30	

Note. Combined sites = two 10/20 sites adjacent to selected 10/10 site. Client number column omits clients whose data was excluded.

Statistical Analysis

The statistical analysis for the symptom measure assessments were paired *t*-tests using IBM SPSS Statistics Version 22. Quantitative analysis was performed using NeuroGuide software, which was exported in by topographical and tabular form. Further analysis was done using Microsoft Excel 2010 and IBM SPSS Statistics Version 22. Computations were done for the frequency bands trained for each client. Given sites, number of bands, and frequency range of bands were unique to each client (see Table 6), it was not feasible to compare simple amplitude changes across clients. As such, the absolute values of the positive and negative *z*-scores were used instead as a way to compare a common metric of pre- and post-changes across clients. The process involved calculating *z*-scores using NeuroGuide software, exporting the results in tabular form using 1 Hz bins, transforming the *z*-scores to use absolute value, then averaging the transformed values for the respective frequency band(s) used for each client. If more than one frequency band was trained at a time (such as downtraining and/or uptraining), the *z*-score values for the bands trained were then averaged for each client and the statistical analysis was completed between the pre- and post-assessments as a group using paired *t*-tests. As opposed to merely averaging the absolute power at each of the treatment sites, *z*-score results were used in order to provide a common measure that was applicable across all frequency bands. Due to the 1/frequency characteristic of the EEG spectrum, with typical alpha peaks, power measures are not consistent across the frequency spectrum. In addition, alpha power measures typically vary significantly between eyes-closed and eyes-open recording conditions. For example, if the power of the frequency band of 8–12 Hz changes by 1 μV, such a change may not be comparable to a 1 μV change in the frequency band of 20–25 Hz.

Results

Symptom Measures

All grouped averaged pre-post comparisons of the three assessments resulted in improvements. A cumulative summary of these results are presented in Table 3.

On the Zung Anxiety Scale, for 11 adult clients, the mean of the pre-scores was 46.00 (*SD* = 9.07) and the mean of the post-scores was 38.83 (*SD* = 7.37). The *t*-test yielded a statistically significant improvement, with $t(10) = 4.59$, $p < 0.001$. While nine clients reported a decrease in their scores, 2 of the 11 clients, reported an increase. See Table 4 for the pre-post scores for each client.

For the SCARED, for three minor clients, the mean of the pre-scores was 37.22 (*SD* = 14.47) and the mean of the post-scores was 21.33 (*SD* = 13.65). The *t*-test resulted a statistically significant improvement, with $t(2) = 27.71$, $p < 0.001$. All clients had improved self-report scores. See Table 5 for the individual pre-post scores.

On the ASEBA, for all categories averaged, the mean of the pre-scores was 63.27 (*SD* = 6.51) and the mean of the post-scores was 59.33 (*SD* = 6.35). The results of the *t*-test was a statistically significant improvement, with $t(17) = 8.75$, $p < 0.001$. Moreover, scores on all 18 categories of the ASEBA improved; see Table 6 the pre-post scores for each category. Improvements in the categories most specific to anxiety symptoms, that is, Anxious/Depressed and Anxiety Problems, were not statistically significant. The checklists do, however, assess for symptoms frequently associated with anxiety, such as withdrawal, somatic issues, thought problems, internalizing, and avoidance; and improvements in these areas were statistically significant.

Table 3
Group Averaged Pre-Post Assessment Results

Assessment (*n*)	Pre-scores *M* (*SD*)	Post-scores *M* (*SD*)	*t*(df)	*p*
Zung Anxiety Scale (*n* = 11)	46.00 (9.07)	38.82 (7.37)	4.59(10)	< 0.001
SCARED Scale (*n* = 3)	37.22 (14.47)	21.33 (13.65)	27.71(2)	< 0.001
ASEBA Across All Categories (*n* = 14)	63.27 (4.88)	59.33 (4.67)	8.76(17)	< 0.001

Table 4
Zung Anxiety Scale

Client #	Pre-scores	Post-scores
2	60	51
4	56	39
6	38	30
8	44	36
10	42	33
12	42	33
13	35	37
14	44	45
15	62	52
16	40	34
17	43	37
Mean (*SD*)	**46.00** (9.07)	**38.83** (7.37)

Note. $t(10) = 4.59$, $p < 0.001$.

Table 5
SCARED Scale

Client #	Pre-scores	Post-scores
1	28	12
7	30	15
11	54	37
Mean (*SD*)	**37.22** (14.47)	**21.33** (13.65)

Note. $t(2) = 27.71$, $p < 0.001$.

Table 6
Achenbach Behavior Checklists (ASEBA)

Category	Pre	Post	*t*(df)	*p*
Anxious/Depressed	69.57	66.86	1.212(13)	.247
Withdrawn	66.21	61.64	2.329(13)	**.037**
Somatic Complaints	65.14	60.71	2.74(13)	**.017**
Thought Problems	66.29	57.86	3.042(13)	**.009**
Attention Problems	69.07	63.43	2.112(13)	.055
Aggressive Behavior	61.79	56.93	2.62(13)	**.021**
Rule-breaking Behavior	60.00	55.43	4.738(13)	**< .001**
Intrusive	44.07	43.14	1.153(10)	.276
Internalizing	69.36	64.93	2.174(13)	**.049**
Externalizing	59.71	54.07	2.713(13)	**.018**
Critical Items	52.57	49.14	3.612(10)	**.005**
Total Problems	65.79	60.79	2.557(13)	**.024**
Depressive Problems (DSM)	69.50	68.79	0.306(13)	.764
Anxiety Problems (DSM)	65.36	64.64	0.49(13)	.632
Somatic Problems (DSM)	62.36	59.21	1.717(13)	.110
ADHD Problems (DSM)	66.29	63.00	1.47(13)	.165
Avoidant Personality Problems (DSM)	66.00	61.93	2.194(13)	**.047**
Antisocial Personality Problems (DSM)	59.79	55.36	3.169(13)	**.007**
Category Mean (*SD*)	**63.27(6.50)**	**59.33(6.34)**		

Note. Bolded values are statistically significant.

Quantitative EEG Results

While not all clients realized improvements in *z*-scores, the difference between pre- and post-measurement showed a decrease in absolute *z*-score values, averaged across all cases, from 1.21 (*SD* = 0.73) to 1.10 (*SD* = 0.62). The improvement was not statistically significant, however. Table 7 provides the pre-post average *z*-scores for each client. It should be noted that one-channel amplitude training was employed as the method of NF, not *z*-score training.

Table 7
Results Pre-Post qEEG Z-scores

Client #	Pre-scores *z*-score	Post-scores *z*-score
1	1.51	0.77
2	1.67	2.32
4	0.77	1.29
6	1.33	1.50
7	0.77	1.44
8	0.70	0.70
10	0.84	0.32
11	2.91	0.49
12	0.75	1.08
13	2.54	2.37
14	0.60	0.89
15	1.10	0.90
16	0.64	0.55
17	0.77	0.72
Mean (*SD*)	**1.21 (0.73)**	**1.10 (0.62)**

Note. *Z*-score pre-post difference was not statistically significant.

Discussion

Symptom improvement was shown with various assessments including: the self-report ASEBA, Zung Anxiety Scale, and SCARED. While two of the most anxiety-specific categories of the ASEBA yielded improvements that were not statistically significant, other anxiety-related categories resulted in significant improvement, and overall the improvement in averaged scores across categories were statistically significant. Taken together, the symptom scales present evidence of a significant improvement in the client's sense of wellbeing.

Interestingly, two categories of the ASEBA that showed robust improvement were Rule-Breaking and Antisocial Personality. A number of researchers have examined the comorbidity of anxiety disorders and Antisocial Personality Disorder or Conduct Disorder, with some evidence of a correlation (Galbraith, Heimberg, Wang, Schneier, & Blanco, 2014; Goodwin & Hamilton, 2003; Hodgins, De Brito, Chhabra, & Côté, 2010). This relationship may serve as an added dimension to the ongoing study based on this pilot, or as an additional focus of research.

The parent rating version of the SCARED was administered, but results presented some problems in interpretation. In one instance, the parents rated their child in opposite ways—one parent reported a large improvement, while the other parent reported a large worsening of symptoms. In this case there was significant parental conflict and one parent divulged that they were divorcing. Due to the confounding nature of the parental reports, only self-reports on the assessments were included for analysis. Parental ratings can be included as the size of the sample increases in the future.

A small sample size and the lack of a control group was a roadblock to an effective research design in some aspects of the study. There were also limitations based on clients receiving therapeutic care (as self-reported) and experimenter bias/skill level. This experimenter bias could have resulted in a response-expectancy effect (Kirsch, 2009). Furthermore, some clients experienced confounding life stressors that could have influenced treatment and medication effects that were not present during the pre- and post-qEEG. Treatment was provided to clients who clearly had characteristics that compromised the quality of data that might be gained from them. They included clients who were inconsistent in attendance, exhibited substance abuse issues (data was excluded), experienced significant life events (such as relational or financial crises), or had mental or medical disorders that possibly reduced the effect of the treatment. This may have resulted in spending a portion of the sessions engaged in active listening and numerous client-centered or CBT therapeutic interventions in different ways and to various extents with the clients. The relative merits of various strategies of controlling for these variations in the future are being considered.

Quantitative designs are descriptive or experimental in nature. A descriptive study establishes only associations between variables and an experimental usually establishes causality. Unfortunately, many variables were not accountable or annotatable. One such effect was positive reinforcement. The presentation and style of secondary reinforcers varied based on student-clinician decisions and were not directly addressed in this study. Operant and classical conditioning techniques were employed to make the feedback as much of a positive reinforcement as possible. This included

the selection of feedback type based on client preference. Some clients expressed preferences for one or more of available options or classes of options, which included: games, animations, sounds (including music), or analogical feedback (such as boxes that grow and shrink in size based on which wave analysis was trained). Positive reinforcement was also provided via verbal prompts and coaching. As the study progresses in the future with additional clients, it may be possible to analyze these variations for significant differences in treatment outcomes.

There was variability in the skill and experience levels of the student counselors. Students were at various levels in their studies within their degree program. Some students had significant experience with NF, while most were novices. Student counselors who were taking an advanced NF course, as an elective to their counseling degree program, saw clients in the counseling department's center. In addition to an introductory course, some of the students had completed one or two semesters of advanced practical and theoretical applications in NF. During the previous courses, the students had worked with one or more NF software systems, had practiced performing NF on other students, and had NF procedures designed for themselves, which were based on qEEG analysis. Some of the students had completed counseling skills courses, practicum and internship hours, while others were novices to counseling. In one case, the student had been the counselor for the client they were seeing for NF treatment as part of a counseling practicum course one semester prior. Controls for the effect of student bias and skill level differences were: supervision from the professor who monitored via informal verbal reports from students and clients, session notes, closed-circuit television, and weekly case conferences.

"Neurofeedback training is all about learning. Each person's rate of learning is unique; some respond more quickly than others do" (Demos, 2005, p. 127). As such, a combined client-centered and quantitative approach is best used in the future. In this case, a quasi-experimental approach needs to be designed. Clients would need to previously be scored on self-efficacy, anxiety scores, and education of basic NF principles. If all scales can be quantified, then limitations, placebo effect, and counselor technique can be assessed during the design phase, and several uncontrolled variables can be at least factored. Excluding students from treating clients with whom they have any previous clinical or personal relationship (e.g., previous student and talk therapy clients they may have had in practicum or internship portions of degree path).

Other client variables to control for, as affecting possible treatment outcomes, would include: adjunct therapies (concurrently used or attending), medications, familial/financial/extraneous life stressors and major life events, injuries/illnesses, changes in sleep, and other therapeutic lifestyle changes, that is, diet, exercise, meditation. Future considerations need to assess whether counselor-client therapeutic modalities need to be standardized amongst clinicians to established protocols of breathing techniques, mindfulness, and meditation in hopes of decreasing variability.

A few clients in the study were taking psychotropic medications, such as benzodiazepine-class anxiolytics and SSRIs. While these effects on the EEG were assessed as part of the qEEG analysis, they remain as a confounding variable for treatment outcomes. As the study continues with the addition of more clients each semester, accounting for this variable will make statistical analysis more robust. This will be accomplished by (1) setting up a comparison between medicated and non-medicated clients, and (2) excluding medicated client data.

Training was conducted using amplitude measures and monopolar site placements only. While this was by design, it excluded other forms of NF which may be based on connectivity measures and multiple site placements. As noted above in the results section, while *z*-score calculations were used in the statistical analysis of EEG changes, the training did not utilize *z*-score training, but qEEG-guided protocols. Two clients, for example, were given posterior alpha enhancement training based on qEEGs that reflected the low-amplitude fast phenotype. One of these clients had a fast alpha peak frequency, showing an elevated *z*-score in the 11–12 Hz range with normal *z*-scores for 8–10 Hz. But, the protocol for this client included uptraining 8–10 Hz (and downtraining 25–30 Hz). In this case, it was expected that the absolute *z*-score might actually show an increase, which turned out to be the case. Although the client successfully modified the amplitudes of both frequency bands, with accompanying symptom improvement, these results present a confounding factor in the *z*-score analysis. The study may have also been strengthened by the addition of a learning curve. This will be added in future analyses.

Finally, it is worth emphasizing that the setting of the study is a community counseling center, located on

a university campus, operated as part of a graduate counseling educational program. As such, the prevailing values in the treatment are (1) the well-being and therapeutic needs of clients, and (2) the learning opportunities for students. Students in the NF program are taught an integrative model of NF and psychotherapy; as such, they naturally carried this approach into their sessions with clients. It became obvious to the professor and students that these priorities, at times, took precedence over a purely NF-based research design in ways that may have compromised the acquisition of "clean" data. It is hoped that as the study continues, the ongoing addition of more clients and students will enable the clearer identification of the sole effects of NF. Nonetheless, the study may replicate the common practices of most NF practitioners and hold value in that regard.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Busner, J., & Targum, S. D. (2007). The clinical global impressions scale: Applying a research tool in clinical practice. *Psychiatry (Edgmont)*, *4*(7), 28–37.

Cannon, R. L. (2015). Editorial perspective: Defining neurofeedback and its functional processes. *NeuroRegulation,* *2*(2), 60–69. http://dx.doi.org/10.15540/nr.2.2.60

Cheon, E.-J., Koo, B.-H., Seo, W.-S., Lee, J.-Y., Choi, J.-H., & Song, S.-H. (2015). Effects of neurofeedback on adult patients with psychiatric disorders in a naturalistic setting. *Applied Psychophysiology and Biofeedback*, *40*(1), 17–24. http://dx.doi.org/10.1007/s10484-015-9269-x

Dantendorfer, K., Prayer, D., Kramer, J., Amering, M., Baischer, W., Berger, P., … Katschnig, H. (1996). High frequency of EEG and MRI brain abnormalities in panic disorder. *Psychiatry Research: Neuroimaging*. *68*(1), 41–53. http://dx.doi.org/10.1016/S0925-4927(96)03003-X

Demerdzieva, A., & Pop-Jordanova, N. (2011). Alpha asymmetry in QEEG recordings in young patients with anxiety. *Prilozi, 32*(1), 229–244.

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

Galbraith, T., Heimberg, R. G., Wang, S., Schneier, F. R., & Blanco, C. (2014). Comorbidity of social anxiety disorder and antisocial personality disorder in the National Epidemiological Survey on Alcohol and Related Conditions (NESARC). *Journal of Anxiety Disorders*, *28*(1), 57–66. http://dx.doi.org/10.1016/j.janxdis.2013.11.009

Gold, C., Fachner, J., & Erkkilä, J. (2013). Validity and reliability of electroencephalographic frontal alpha asymmetry and frontal midline theta as biomarkers for depression. *Scandinavian Journal of Psychology*, *54*(2), 118–126. http://dx.doi.org/10.1111/sjop.12022

Goodwin, R. D., & Hamilton, S. P. (2003). Lifetime comorbidity of antisocial personality disorder and anxiety disorders among adults in the community. *Psychiatry Research*, *117*(2), 159–166. http://dx.doi.org/10.1016/S0165-1781(02)00320-7

Gunkelman, J. (2006). Transcend the DSM using phenotypes. *Biofeedback*, *34*(3), 95–98.

Gurnee, R. (2000, September). *EEG Based Subtypes of Anxiety (GAD) and Treatment Implications.* [Abstract]. Oral Presentation at the 8th Annual Conference of the International Society for Neurofeedback and Research, St. Paul, MN.

Hammond, D. C. (2010). The Need for Individualization in Neurofeedback: Heterogeneity in QEEG Patterns Associated with Diagnoses and Symptoms. *Applied Psychophysiology and Biofeedback*, *35*(1), 31–36. http://dx.doi.org/10.1007/s10484-009-9106-1

Heller, W., Nitschke, J. B., Etienne, M. A., & Miller, G. A. (1997). Patterns of regional brain activity differentiate types of anxiety. *Journal of Abnormal Psychology*, *106*(3), 376–385. http://dx.doi.org/:10.1037/0021-843X.106.3.376

Hill, R. W., & Castro, E. (2002). *Getting rid of ritalin: How neurofeedback can successfully treat attention deficit disorder without drugs*. Charlottesville, VA: Hampton Roads.

Hodgins, S., De Brito, S. A., Chhabra, P., & Côté, G. (2010). Anxiety disorders among offenders with antisocial personality disorders: A distinct subtype? *Canadian Journal of Psychiatry*, *55*(12), 784–791.

Johnstone, J., Gunkelman, J., & Lunt, J. (2005). Clinical database development: Characterization of EEG phenotypes. *Clinical EEG and Neuroscience*, *36*(2), 99–107. http://dx.doi.org/10.1177/155005940503600209

Jones, M. S. (2015). Comparing DC Offset and Impedance Readings in the Assessment of Electrode Connection Quality. *NeuroRegulation*, *2*(1), 29–36. http://dx.doi.org/10.15540/nr.2.1.29

Kerson, C., Sherman, R. A., & Kozlowski, G. P. (2009). Alpha Suppression and Symmetry Training for Generalized Anxiety Symptoms. *Journal Of Neurotherapy*, *13*(3), 146–155. http://dx.doi.org/10.1080/10874200903107405

Kirsch, I. (2009). Antidepressants and the placebo response. *Epidemiology and Psychiatric Sciences*, *18*(4), 318–322. http://dx.doi.org/10.1017/S1121189X00000282

Krigbaum, G., & Wigton, N. L. (2014). When Discussing Neurofeedback, Does Modality Matter? *NeuroRegulation, 1*(1), 48–60. http://dx.doi.org/10.15540/nr.1.1.48

National Institute of Mental Health. (2015). What are anxiety disorders? Retrieved from http://www.nimh.nih.gov/health/topics/anxiety-disorders/index.shtml

Machleidt, W., Gutjahr, L., Muegge, L., & Hinrich, A. (1985). Anxiety processes in the EEG. *Electroencephalography and Clinical Neurophysiology*, *61*(3), S118–S119. http://dx.doi.org/10.1016/0013-4694(85)90468-7

Price, J., & Budzynski T. (2009). Anxiety, EEG patterns, and neurofeedback. In T. H. Budzynski, H. K. Budzynski, J. R. Evans, & A. Abarbanel (Eds.), *Introduction to Quantitative EEG and Neurofeedback: Advanced Theory and Applications* (pp. 453–470). Burlington, MA: Elsevier.

Savostyanov, A. N., Tsai, A. C., Liou, M., Levin, E. A., Lee, J.-D., Yurganov, A. V., & Knyazev, G. G. (2009). EEG-correlates of trait anxiety in the stop-signal paradigm. *Neuroscience Letters*, *449*(2), 112–116. http://dx.doi.org/10.1016/j.neulet.2008.10.084

Scheinost, D., Stoica, T., Saksa, J., Papademetris, X., Constable, R. T., Pittenger, C., & Hampson, M. (2013). Orbitofrontal cortex neurofeedback produces lasting changes in contamination anxiety and resting-state connectivity. *Translational Psychiatry*, *3*(4), e250. http://dx.doi.org/10.1038/tp.2013.24

Siciliani, O., Schiavon, M., & Tansella, M. (1975). Anxiety and EEG alpha activity in neurotic patients. *Acta Psychiatrica Scandinavica*, *52*(2), 116–131. http://dx.doi.org/10.1111/j.1600-0447.1975.tb00028.x

Singer, K. (2004). The effect of neurofeedback on performance anxiety in dancers. *Journal of Dance Medicine and Science*, *8*(3), 78–81.

Spielberger, C. D. (1983). *State-Trait Anxiety Inventory for Adults*. Redwood City, CA: Mind Garden, Inc.

Stern, J. (2005). *An Atlas of EEG Patterns*. Philadelphia, PA: Lippincott Williams & Wilkins.

Tharawadeepimuk, K., & Wongsawat, Y. (2014, November). *QEEG evaluation for anxiety level analysis in athletes*. Paper presented at the 2014 7th Biomedical Engineering International Conference (BMEiCON) of IEEE, Fukuoka, Japan. http://dx.doi.org/10.1109/BMEiCON.2014.7017400

Thompson, M., & Thompson, L. (2003). *The neurofeedback book: An introduction to basic concepts in applied psychobiology*. Wheat Ridge, CO: The Association for Applied Psychophysiology and Biofeedback.

Walker, J. E. (2009). Anxiety associated with post traumatic stress disorder—the role of quantitative electro-encephalograph in diagnosis and in guiding neurofeedback training to remediate the anxiety. *Biofeedback*, *37*(2), 67–70.

Wigton, N. L., & Krigbaum, G. (2015a). Attention, executive function, behavior, and electrocortical function, significantly improved with 19-channel *z*-score neurofeedback in a clinical setting: A pilot study. *Journal of Attention Disorders.* Advance online publication. http://dx.doi.org/10.1177/1087054715577135

Wigton, N. L., & Krigbaum, G. (2015b). A review of qEEG-guided neurofeedback. *NeuroRegulation,* *2*(3), 149–155. http://dx.doi.org/10.15540/nr.2.3.149

Received: August 5, 2015
Accepted: September 27, 2015
Published: October 8, 2015

NeuroRegulation

A Review of qEEG-Guided Neurofeedback

Nancy L. Wigton[1,2*] *and Genomary Krigbaum*[3,4]

[1]Grand Canyon University, College of Humanities and Social Sciences, Phoenix, Arizona, USA
[2]Applied Neurotherapy Center, LLC, Scottsdale, Arizona, USA
[3]Marian University, College of Osteopathic Medicine, Indianapolis, Indiana, USA
[4]Grand Canyon University, College of Doctoral Studies, Phoenix, Arizona, USA

Abstract

While there are literature reviews and meta-analytic coverage of neurofeedback (NF) studies that focus on traditional amplitude NF and slow cortical potential NF, the same is not true for quantitative electroencephalographic (qEEG)-guided NF (qNF). To that end, this is a literature review of several qNF research articles. Generally, most are found in clinical settings, address a wide variety of symptoms and diagnoses, use clinical assessments as outcome measures, employ individualized NF protocols based on qEEG findings, and define efficacy in terms of improvement on pre-post outcome measures. However, few report pre-post qEEG metrics as outcome measures. Suggestions for future research are presented.

Keywords: qEEG-guided NF; neurofeedback; qEEG; EEG biofeedback

Citation: Wigton, N. L., & Krigbaum, G. (2015). A Review of qEEG-Guided Neurofeedback. *NeuroRegulation*, *2*(3), 149–155. http://dx.doi.org/10.15540/nr.2.3.149

***Address correspondence to**: Nancy L. Wigton, PhD, Applied Neurotherapy Center, 10020 N 92nd Street, #120, Scottsdale, AZ 85260, USA. Email: Dr.Wigton@appliedneurotherapy.com

Edited by:
Rex Cannon, PhD, Neural Potential, Florida, USA

Reviewed by:
Rex Cannon, PhD, Neural Potential, Florida, USA
Randall Lyle, PhD, Mount Mercy University, Iowa, USA

Introduction

In recent years there has been a rapid surge of articles focused on neurofeedback (NF) in the literature. In this landscape, there exist reviews and meta-analysis studies on traditional amplitude-based NF (Arns, de Ridder, Strehl, Breteler, & Coenen, 2009; Arns, Heinrich, & Strehl, 2014; Brandeis, 2011; Gevensleben, Rothenberger, Moll, & Heinrich, 2012; Lofthouse, Arnold, Hersch, Hurt, & DeBeus, 2012; Niv, 2013; Pigott, De Biase, Bodenhamer-Davis, & Davis, 2013) and a meta-analytic style review of slow cortical potential NF (Mayer, Wyckoff, & Strehl, 2013). In regards to the recent *z*-score NF modalities, a few studies with a quantitative analytic focus have begun to emerge (such as, Hammer, Colbert, Brown, & Ilioi, 2011; Krigbaum & Wigton, 2015; Wigton & Krigbaum, 2015); yet, there are too few to expect a meta-analysis or review summaries. However, to date no meta-analysis or comprehensive review has been found of quantitative electroencephalographic (qEEG)-guided NF (qNF), in spite of its origins dating back to the 1990s.

Even so, that is not to say that qNF is devoid of research. In fact, from 2002 to 2015 there are numerous studies in peer-reviewed literature addressing the qNF model. Unique to this genre of studies, though, is great diversity in the different conditions treated, as well as a greater use of individualized, custom-designed protocols; thus, making meta-analysis of this collection of research less feasible (Krigbaum & Wigton, 2014). Nonetheless, these studies do represent a body of research pointing to the efficacy of qNF. This, then, is intended to review qNF as represented in the literature. While this is not intended to be an exhaustive review of all qNF studies, it is believed to be a representative sample of the literature coverage of this particular NF modality.

Background Information

Historical Perspectives

While understanding of the multiple components to the EEG signal was evident as early as the 1930s, the advent of computer technology was necessary for qEEG advances (Collura, 1995); for example, the incorporation of normative databases in conjunction

with qEEG analysis. Early implementations of qEEG normative database applications date 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 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. As a result, the 1990s brought forth a wider acceptance of qEEG technology 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 NF is 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 brought forth 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.

Theoretical Foundations

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 led to 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 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).

Normalization Model of qNF

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, Drinkenburg, & Kenemans, 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, 2010). 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, 2010).

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 are not occurring by chance, but due to the training 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, 2010). 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; Sürmeli & Ertem, 2009; Walker, 2010). 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 (Sürmeli, Ertem, Eralp, & Kos, 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, Arns, Peters, Giepmans, & Verhoeven, 2010; Collura, 2008; Sürmeli et al., 2012; Sürmeli & Ertem, 2009, 2010; Walker, 2009, 2010, 2011, 2012a).

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.

qNF in the Literature

Arns et al. (2012) conducted a well-designed open-label study of 21 attention deficit/hyperactivity disorder (ADHD) participants 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 3 years earlier (Arns et al., 2009). The outcome measures incorporated were a self-report scale based on the Diagnostic and Statistical Manual-IV (APA, 2000) list of symptoms and the Beck Depression Inventory (Beck, Steer, & Garbin, 1988). The findings of the 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 (*SD* 16.09), 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. Additionally, 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 in scope.

Koberda, Hillier, Jones, Moses, and Koberda (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. The 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*-tests 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. The 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 participants were collected to include pre- and post-scores for outcome measures from a behavioral rating scale (Child Behavior Checklist; CBCL; Achenbach, 1991), and a computerized performance test (Test of Variables of Attention; TOVA; Greenberg, 1987). 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 participants 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 Sürmeli 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 and encompass a timeframe from 2002 to 2015.

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 issues (Walker, 2013). His qNF protocol development centers on tailoring the protocol to the individual clinical and 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 clinical outcome measures and yield benefits from qNF. The outcome measures that Walker typically employs are primarily Likert or percentage-based self-reports, except in the anger control study where the DeFoore (2002) 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 $n = 26$, 84% of the participants reported greater than 50% improvement in symptoms. For the anxiety/post-traumatic stress article, with $n = 19$, 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, in the NF group 54% had complete remission of headaches, 39% had a greater than 50% reduction, and 4% experienced less than 50% reduction in migraines; 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 $p < .001$.

Sürmeli and colleagues reported on Down syndrome (Sürmeli & Ertem, 2007), personality disorders (Sürmeli & Ertem, 2009), intellectual disability (mental retardation; Sürmeli & Ertem, 2010), obsessive-compulsive disorder (Sürmeli & Ertem, 2011), schizophrenia (Sürmeli et al., 2012), and dementia (Sürmeli et al., 2015). Notable in this collection of work are conditions previously not known to respond to NF, such as personality disorders, intellectual disability, 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 Sürmeli et al. studies are a high number of sessions reported for the cases; ranging from an average of 45 to an average of 120 sessions. No particular research design is declared in the Sürmeli et al. studies, but here too, comparisons of pretest-posttest outcome measures are reported, indicating qNF brings about improvements in outcome measures. These studies 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; Kay, Fiszbein, & Opler, 1987), for the obsessive compulsive disorder research they incorporated the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS; Goodman et al., 1989), and for the dementia study the Mini Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975) was the primary outcome measure. For many studies, the computerized performance TOVA was used. Yet, as with Walker's work, in spite of all protocols being individually qEEG-guided, qEEG data is infrequently used or reported as an outcome measure; typically, 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 n = 13, 12 were significantly improved on all outcome measures; with the Symptom Assessment-45 Questionnaire (Riverside Publishing; Rolling Meadows, IL) at p = .002, the Minnesota Multiphasic Personality Inventory (MMPI; University of Minnesota Press; Minneapolis, MN) Psychopathy scale at p = .000, and the TOVA at $p < .05$ on the visual and auditory impulsivity scales. With the article focusing on participants with intellectual disability, including n = 23, for 19 there was improvement on the Wechsler Intelligence Scale for Children-Revised (Wechsler, 1974; 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 study evaluating participants with Down syndrome, 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 participants in the study (n = 7) showed improvement at $p < .02$ on all questionnaire scales. With qNF for obsessive compulsive disorder, with n = 36, 33 showed improvement on the Y-BOCS (Obsession subscale, Compulsion subscale, and Total score all $p < .01$). In the schizophrenia study, with 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. Finally, in the dementia study, with n = 20, all participants' MMSE scores improved with an increase of six points on average ($p < .01$), regardless of dementia type (Alzheimer's disease or Vascular dementia); also qEEG improvements were reported as theta activity decreasing overall ($p < .01$) and a decrease in interhemispheric coherence ($p < .01$).

Conclusion

In summary, studies evaluating qNF typically focus on a wide variety of clinical symptoms and/or mental health diagnoses, and frequently have relatively small sample sizes. With few exceptions, literature presented on qNF 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. 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 report pre-post qEEG metrics, and only three (Arns et al., 2012; Breteler et al., 2010; Sürmeli et al., 2015) incorporated statistical analysis of qEEG metrics as an outcome measure (and that was to a limited degree). More so, none report a measure of overall normalization of the qEEG. 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 qEEG *z*-score mean data in the present qNF research. Therefore, it is important for future qNF studies to incorporate qEEG metrics as outcome measures. Methodologies developed by Krigbaum and Wigton (2015) in a single-subject design, and Wigton and Krigbaum (2015) with group means data, while implemented with the 19-channel *z*-score NF modality, are similarly applicable to qNF studies data as a means of measuring overall normalization of the qEEG.

References

Achenbach, T. M. (1991). Integrative guide to the 1991 CBCL/4-18, YSR, and TRF profiles. Burlington, VT: University of Vermont, Department of Psychiatry.

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

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. http://dx.doi.org/10.1177/155005940904000311

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*(3), 171–180. http://dx.doi.org/10.1007/s10484-012-9191-4

Arns, M., Heinrich, H., & Strehl, U. (2014). Evaluation of neurofeedback in ADHD: The long and winding road. *Biological Psychology, 95*, 108–115. http://dx.doi.org/10.1016/j.biopsycho.2013.11.013

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77–100. http://dx.doi.org/10.1016/0272-7358(88)90050-5

Brandeis, D. (2011). Neurofeedback training in ADHD: More news on specificity. *Clinical Neurophysiology, 122*(5), 856–857. http://dx.doi.org/10.1016/j.clinph.2010.08.011

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*(1), 5–11. http://dx.doi.org/10.1007/s10484-009-9105-2

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*(1), 27–31. http://dx.doi.org/10.1016/j.neulet.2012.04.035

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*(1), 13–23. http://dx.doi.org/10.1007/s10484-009-9102-5

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*(1), 37–38. http://dx.doi.org/10.1007/s10484-009-9108-z

Corsi-Cabrera, M., Galindo-Vilchis, L., del-Río-Portilla, Y., Arce, C., & Ramos-Loyo, 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. http://dx.doi.org/10.1016/j.clinph.2006.08.013

DeFoore, W. G. (2002). Anger scale. In R. diGiuseppe & R. Tatrate (Eds.), Understanding anger disorders (pp. 346–347). New York, NY: Oxford University Press.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198. http://dx.doi.org/10.1016/0022-3956(75)90026-6

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. http://dx.doi.org/10.1586/ERN.12.22

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., ... Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale: I. Development, use, and reliability. *Archives of General Psychiatry, 46*(11), 1006–1011. http://dx.doi.org/10.1001/archpsyc.1989.01810110048007

Greenberg, L. M. (1987). An objective measure of methylphenidate response: Clinical use of the MCA. *Psychopharmacology Bulletin, 23*, 279–282.

Hammer, B. U., Colbert, A. P., Brown, K. A., & Ilioi, E. C. (2011). Neurofeedback for insomnia: A pilot study of *z*-score SMR and individualized protocols. *Applied Psychophysiology and Biofeedback, 36*(4), 251–264. http://dx.doi.org/10.1007/s10484-011-9165-y

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

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. http://dx.doi.org/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. http://dx.doi.org/10.1176/jnp.11.2.190

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*(3–4), 31–52. http://dx.doi.org/10.1300/J184v07n03_02

Kay, S. R., Fiszbein A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin, 13*(2), 261–276. http://dx.doi.org/10.1093/schbul/13.2.261

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. http://dx.doi.org/10.1080/10874208.2012.705770

Krigbaum, G., & Wigton, N. L. (2014). When discussing neurofeedback, does modality matter? *NeuroRegulation, 1*(1), 48–60. http://dx.doi.org/10.15540/nr.1.1.48

Krigbaum, G., & Wigton, N. L. (2015). A methodology of analysis for monitoring treatment progression with 19-channel *z*-score neurofeedback (19ZNF) in a single-subject design. *Applied Psychophysiology and Biofeedback, 40*(3), 139–149. http://dx.doi.org/10.1007/s10484-015-9274-0

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*(5), 351–372. http://dx.doi.org/10.1177/1087054711427530

Mayer, K., Wyckoff, S. N., & Strehl, U. (2013). One size fits all? Slow cortical potentials neurofeedback: A review. *Journal of Attention Disorders, 17*(5), 393–409. http://dx.doi.org/10.1177/1087054712468053

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

Niv, S. (2013). Clinical efficacy and potential mechanisms of neurofeedback. *Personality and Individual Differences, 54*(6), 676–686. http://dx.doi.org/10.1016/j.paid.2012.11.037

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. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology*, (3rd ed., pp. 56–84), Cambridge, England: Cambridge University Press.

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

Sürmeli, T., & Ertem, A. (2009). QEEG guided neurofeedback therapy in personality disorders: 13 case studies. *Clinical EEG and Neuroscience, 40*(1), 5–10. http://dx.doi.org/10.1177/155005940904000107

Sürmeli, 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. http://dx.doi.org/10.1177/155005941004100108

Sürmeli, 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. http://dx.doi.org/10.1177/155005941104200310

Sürmeli, 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*(2), 133–144. http://dx.doi.org/10.1177/1550059411429531

Sürmeli, T., Eralp, E., Mustafazade, I., Kos, H., Özer, G. E., & Surmeli, O. H. (2015). Quantitative EEG neurometric analysis-guided neurofeedback treatment in dementia: 20 cases. How neurometric analysis is important for the treatment of dementia and as a biomarker? *Clinical EEG and Neuroscience*. Advanced online publication. http://dx.doi.org/10.1177/1550059415590750

Thatcher, R. W. (2012). Handbook of quantitative electroencephalography and EEG biofeedback. St. Petersburg, FL: 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. http://dx.doi.org/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.

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*(2), 67–70. http://dx.doi.org/10.5298/1081-5937-37.2.67

Walker, J. E. (2010). 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*(1), 25–27. http://dx.doi.org/10.1007/s10484-009-9107-0

Walker, J. E. (2011). QEEG-guided neurofeedback for recurrent migraine headaches. *Clinical EEG and Neuroscience, 42*(1), 59–61. http://dx.doi.org/10.1177/155005941104200112

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

Walker, J. E. (2012b). QEEG-guided neurofeedback for remediation of dysgraphia. Biofeedback 40(3), 113–114. http://dx.doi.org/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. http://dx.doi.org/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. http://dx.doi.org/10.1300/J184v06n02_05

Wechsler, D. (1974). Wechsler Intelligence Scale for Children-Revised. New York: Psychological Corporation.

Wigton, N. L., & Krigbaum, G. (2015). Attention, executive function, behavior, and electrocortical function, significantly improved with 19-channel *z*-score neurofeedback in a clinical setting: A pilot study. *Journal of Attention Disorders*. Advance online publication. http://dx.doi.org/10.1177/1087054715577135

Received: September 19, 2015
Accepted: September 28, 2015
Published: October 8, 2015